Machine Translated by Google

Enzyklopädie

Autor und Herausgeber

Elzein Salah Elzein

IEEE-Mitglied

- Beiträge zum Thema Elektronik

IEEE Publisher, Großbritannien



Dieser Inhalt ist vertraulich

Machine Translated by Google

Elektronik

MDPI

Artikel

Selektives Greifen komplex geformter Teile mit topologischer Skelettextraktion

Andrea Pennisi [1], Monica Sileo [2], Domenico Daniele Bloisi [1] und Francesco Pierri 2,*

[1] Abteilung für Internationale Geistes- und Sozialwissenschaften, UNINT Internationale Universität Rom, 00147 Rom, Italien; andrea.pennisi@gmail.com (AP); domenico.bloisi@unint.eu (DDB)

[2] Fakultät für Ingenieurwissenschaften, Universität Basilikata, 85100 Potenza, Italien; monica.sileo@unibas.it

* Korrespondenz: francesco.pierri@unibas.it; Tel.: +39-0971-205020

Zusammenfassung: Um die Autonomie und Flexibilität von Robotersystemen zu verbessern, spielt die Fähigkeit, Objekte wahrzunehmen und zu greifen, eine entscheidende Rolle. Genauer gesagt müssen Robotermanipulatoren die Anwesenheit der Objekte in ihrem Arbeitsbereich erkennen, den Greifpunkt identifizieren und eine Flugbahn berechnen, um sich den Objekten mit einer für die Ausführung der Aufgabe geeigneten Pose des Endeffektors zu nähern. Dies können bei komplexen Geometrien, bei denen mehrere Greifpunktkandidaten erkannt werden können, anspruchsvolle Aufgaben sein. In diesem Artikel stellen wir einen neuartigen Ansatz für den Umgang mit komplex geformten Autoteilen vor, der aus einer auf Deep Learning basierenden Methode zur topologischen Skelettextraktion und einem aktiven Greifpositionsauswahlmechanismus besteht. Insbesondere verwenden wir eine modifizierte Version des bekannten Lightweight OpenPose-Algorithmus, um das topologische Skelett realer Autoteile zu schätzen. Das geschätzte Skelett wird verwendet, um die beste Greifposition für das vorliegende Objekt auszuwählen . Unser Ansatz ist so konzipiert, dass er im Vergleich zu anderen vorhandenen **Methoden zur Greifpositionserkennung rechnerisch effizienter ist . Quantitative Experimente, die mit einem 7-DoF-Manipulator an verschiedenen realen Automobilkomponenten durchgeführt wurden, belegen die Wirksamkeit des vorgeschlagenen Ansatzes mit einer Erfolgsrate von 87,04 %.**

Schlüsselwörter: Computer Vision für die Fertigung; Deep Learning beim Greifen und Manipulieren; visuelles Lernen

check for updates

Zitat: Pennisi, A.; Sileo, M.; Bloisi, DD; Pierri, F. Selektives Greifen komplex geformter Teile mittels topologischer Skelettextraktion. Electronics **2024,** 13, **3021. https:// doi.org/ 10.3390/electronics13153021**

Wissenschaftliche Herausgeber: Krzysztof Okarma und Piotr Lech

Erhalten: 4. Juni 2024
Überarbeitet: 23. Juli 2024
Akzeptiert: 29. Juli 2024
Veröffentlicht: 31. Juli 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

In der **Fertigungsindustrie, z. B. in der Automobilindustrie, werden häufig Komponenten mit komplexen geometrischen Formen verwendet .** Der Einsatz von Robotern zur Handhabung komplexer Teile ist **aufgrund von Wahrnehmungs-, Planungs- und Denkproblemen immer noch eine anspruchsvolle Aufgabe. Insbesondere Unsicherheiten in der Position des zu greifenden Objekts und Wahrnehmungsrauschen aufgrund reflektierender Materialien sind häufige Herausforderungen in industriellen S**

Das Greifen von Objekten mit komplexen Geometrien kann grob in modellbasierte Methoden eingeteilt werden, die auf bereits vorhandenen 3D-Modellen basieren, und lernbasierte Techniken, die maschinelles Lernen zur Vorhersage von Greifpunkten einsetzen. Beide Ansätze müssen die äußere Umgebung mithilfe von visuellen Algorithmen wahrnehmen, die auf Kameras und Punktwolken zur Objekterkennung, Segmentierung und Posenschätzung basieren, wie in [1,2], oder taktilen Strategien, wie in [3], die Sensoren zur Kraftmessung, haptischen Rückmeldung und Rutscherkennung erfordern. Hybride Ansätze kombinieren mehrere Methoden zur Robustheit, darunter Multisensorfusion und aktive Wahrnehmung.

In diesem Artikel präsentieren wir eine vollständige Pipeline für die Handhabung komplex geformter Automobilteile mit einem 7-DoF-Robotermanipulator. Insbesondere verwenden wir einen Deep-Learning-basierten Ansatz, um einen Multiobjektdetektor zu entwickeln, der das topologische Skelett des zu greifenden Teils extrahieren soll, um dessen Pose präzise abzuschätzen. Nach der Schätzung der Pose wird die beste Greifpose ausgewählt, um die Chance zu erhöhen, das **Objekt erfolgreich zu greifen .**

Der Beitrag dieser Arbeit ist dreifach.

Machine Translated by Google

1. Der Skelettextraktionsprozess liefert eine Darstellung der Objektposition im 3D- Raum. Daher ermöglicht er auch eine genaue Schätzung der Ausrichtung des Objekts.
2. Wir verwenden MobileNetV3, um das ursprüngliche MobileNetV1-Backbone im Skelettextraktionsnetzwerk zu ersetzen , und passen es an, um das Skelett von Industrieobjekten mit komplexer Geometrie sowohl von vorne als auch von hinten zu erkennen, selbst wenn die Objekte auf beiden Seiten unterschiedliche Formen aufweisen.
3. Die Auswahl der Greifpose erfolgt dynamisch. Das bedeutet, dass ein Fehler bei der Skelettextraktion selbstständig erkannt wird und der Roboter seine Position aktiv ändert, um den Greifvorgang besser durchführen zu können.

Wir haben mehrere Experimente mit echten Autoteilen durchgeführt, um unseren Ansatz zu validieren, der die Schlüsselpunkte des Objekts ohne jegliche Einschränkung aus der kopfstehenden Ansicht erkennen kann .

Der Rest des Dokuments ist wie folgt gegliedert. Abschnitt 2 enthält eine kurze Beschreibung bestehender, verwandter Arbeiten. Unsere Methode wird in Abschnitt 3 vorgestellt. Experimente, die die Wirksamkeit des vorgeschlagenen Ansatzes demonstrieren, werden in Abschnitt 4 gezeigt. Abschließend werden in Abschnitt 5 Schlussfolgerungen und zukünftige Richtungen gezogen .

2. Verwandte Arbeiten

Dank der Verfügbarkeit leistungsstarker GPUs sind Deep-Learning-Methoden in den letzten Jahren für die Greifpunkterkennung geeignet geworden. Sie haben sich als fähig erwiesen, traditionelle analytische Ansätze zu ersetzen, die auf geometrischen Eigenschaften, physikalischen Modellen und Kraftanalysen basieren. In [4] wurde eine Convolutional Neural Network (CNN)-Architektur namens GraspNet vorgestellt , die greifbare Bereiche auf der Oberfläche von Objekten segmentieren kann . In [5] wurde ein CNN in Kombination mit den von einer Tiefenkamera bereitgestellten Informationen verwendet, um die Anwesenheit des Objekts und die beste Greifposition zu erkennen.
Es wurden verschiedene Ansätze vorgeschlagen, um die Genauigkeit von Deep CNN zu verbessern, siehe z. B. [6,7], aber sie erfordern üblicherweise lange Rechenzeiten (d. h. im Sekundenbereich).

Effizientere Ansätze, die nur Tiefenbilder erfordern, wurden in [8,9] vorgeschlagen.
Genauer gesagt wurde in [8] ein Deep Convolutional Neural Network in einer simulierten Umgebung trainiert, um greifrelevante Merkmale zu erlernen und für jedes Objekt eine Einzelgrifflösung zu liefern. In [9] wurde das sogenannte Generative Grasping Convolutional Neural Network (GG-CNN) vorgeschlagen. Es ermöglicht die direkte Bewertung der Greifqualität und -haltung für jedes Pixel in einem Tiefenbild und ist schnell genug, um das Greifen in dynamischen Umgebungen durchzuführen . Die Leistung des GG-CNN wurde durch die Einführung des GG-CNN2 [10] verbessert , einem CNN, das auf der semantischen Segmentierungsarchitektur von [11] basiert.
Ein gemeinsames Merkmal von Deep-Learning-basierten Methoden zur Greifpunkterkennung besteht darin, dass für jeden einzelnen Pixel im vorliegenden Bild der Greifqualitätswert berechnet werden muss, was äußerst zeitaufwändig ist.

Wenn weder Wissen über das Objekt noch die möglichen Greifposen verfügbar sind, kann das Objekt mithilfe von Formprimitiven approximiert werden, z. B. durch Multiview-Messungen [12] oder durch Identifizierung von Merkmalen in Sensordaten [13]. Die in [14] vorgeschlagene Methode besteht darin, mögliche Greifposen auszuwählen, nachdem Bereiche lokalisiert wurden, in denen bereits ein erfolgreicher Griff erfahren wurde. In [1] wird das Greifen teilweise bekannter Objekte in unstrukturierten Umgebungen vorgeschlagen, basierend auf einer Erweiterung der bekannten Technik der Hintergrundsubtraktion auf den industriellen Kontext [15]. Dank der Verbreitung kostengünstiger Tiefensensoren wurden viele 3D-Registrierungsalgorithmen genutzt, um das Problem des Objektgreifens zu bewältigen. So wird in [2] beispielsweise ein Modell des zu greifenden Objekts mithilfe einer Reihe von Punktwolken erzeugt, die aus verschiedenen Positionen aufgenommen wurden, und die nominale Greifpose wird festgelegt. Anschließend wird dieses

In dieser Arbeit schlagen wir einen skelettbasierten Ansatz zur Erkennung der Greifposen vor, der aufgrund der kompakten Darstellung des Objekts durch das Skelett von Natur aus weniger rechenintensiv ist.

Machine Translated by Google

Ein qualitativer und quantitativer Vergleich zwischen unserem Ansatz und den wichtigsten in diesem Abschnitt beschriebenen Arbeiten ist in Tabelle 1 dargestellt. Dieser Vergleich Berücksichtigen Sie nicht nur die Ergebnisse, sondern auch die Einschränkungen jeder Arbeit.

Tabelle 1. Vergleichstabelle zwischen verschiedenen Ansätzen zum Greifen von Objekten.

Methoden	Anwendungen	Quantitative Ergebnisse	Einschränkungen
CNN-Architektur mit einem DDF Modul [4]	Robotergreifen in Echtzeit	90 % Genauigkeit bei Cornell Datensatz erfassen	Fehler bei der Vorhersage der Orientierung für einige Objekte
Struktur basierend auf schneller R-CNN und DACAB [5]	Objektgreifen mit dem Handy Manipulator	86,3 % Erfolgsquote	Ineffiziente Suchmethode
GQ_CNN zur robusten Klassifizierung Greifen [6]	Greifen von Haushaltsgegenständen	99 % Präzision	Lange Rechenzeit
DCNN basierend auf Tiefenbildern zur Vorhersage der Greifposition [8]	Greifen unbekannter Objekte	92 % (70 %) Präzision mit zylindrisch geformt (kastenförmige) Objekte	Erzeugung eines Einzelgriffs Lösung für ein einzelnes Objekt
NN zum Lernen von Prototypen Teile [14]	Greifen gleichartiger Gegenstände	N / A	Erfassen von komplex geformten Objekte mit noch nie gesehenen Merkmale.
Topologisches Skelett Extraktion (diese Arbeit)	Erfassen von komplex geformten Autoteile	87,04 % Erfolgsquote	Benötigt eine gute Kamera Kalibrierung

3. Vorgeschlagene Methode

Abbildung 1 zeigt die funktionale Gesamtarchitektur unseres Ansatzes. Sie besteht aus vier Hauptmodule, nämlich Visual Data Acquisition, Topological Skeleton Extraction, Grasping Posenauswahl und Robotergreifen. Jedes Modul wird unten ausführlich beschrieben.

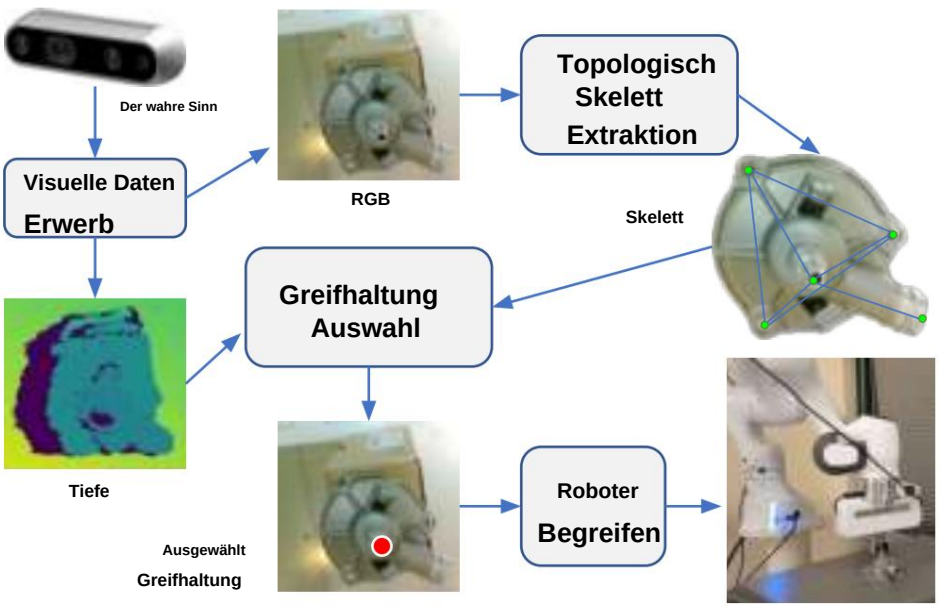


Abbildung 1. Funktionale Architektur des vorgeschlagenen Ansatzes.

Machine Translated by Google

3.1. Visuelle Datenerfassung

Visuelle Daten (sowohl RGB als auch Tiefe) werden mithilfe einer Intel Realsense D435 RGBD-Kamera erfasst , die über eine 3D-gedruckte Halterung in der sogenannten Auge-in-Hand- Konfiguration **am Endeffektor montiert ist . Abbildung 2 zeigt die Referenzrahmen, die am Roboter-Endeffektor** Fe **und der Kamera** Fc angebracht sind.

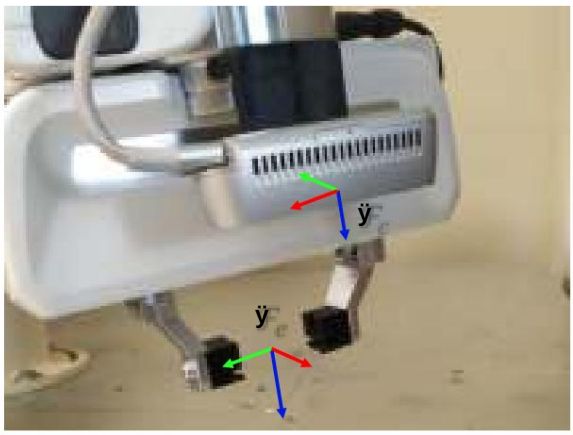


Abbildung 2. Referenzrahmen für Endeffektor und Kamera. ÿ

Die Intel Realsense D435-Kamera verfügt über eine Mindesttiefenentfernung, über die hinaus sie keine Tiefenmessung von etwa 28 cm liefern kann, und die Datenerfassung der Kamera erfordert die Bibliothek realsense-ros, **da die Kommunikation zwischen den Modulen über das Robot Operating System (ROS) erfolgt. Die Kamera wurde zuvor mit 30 Bildern eines 2D-Schachbrettmusters kalibriert. Der Kalibrierungsprozess umfasst sowohl intrinsische als auch extrinsische Kalibrierungen. Die erste zielt darauf ab, die Kameraparameter zu bestimmen, die beschreiben, wie die Kamera die 3D-Koordinaten der Szene in die 2D-Koordinaten des Bildes umwandelt, wie Brennweite, Hauptpunkt und optische Verzerrungen, während die zweite die Parameter liefert, die die starre Transformation beschreiben,** die die 3D-Koordinaten der realen Welt auf die 3D-Koordinaten **des Referenzsystems der Kamera abbildet .** Die von **der VISP-Bibliothek [16] vorgeschlagene Implementierung des Kalibrierungsverfahrens , basierend auf [17,18], wurde unter Verwendung eines Schachbretts übernommen, das aus 9 × 6 Quadraten mit Abmessungen von 0,02645 m besteht.**

Es sei daran erinnert, dass ein gutes Kalibrierungsverfahren für den Erfolg des Greifvorgangs von entscheidender Bedeutung ist, da es eine genaue Wahrnehmung der Umgebung gewährleistet und so eine präzise Identifizierung und Positionierung der Punkte im dreidimensionalen Raum für eine erfolgreiche Manipulation ermöglicht.

3.2. Extraktion des topologischen

Skeletts **Die vorgeschlagene Methode wurde für Objekte entwickelt, die:**

• sind starr, da sie nicht auf verformbare Objekte anwendbar sind; • sind nicht perfekt symmetrisch, denn obwohl es möglich ist, ein nicht-symmetrisches topologisches Skelett zu definieren , kann der Detektor während des Extraktionsprozesses aufgrund symmetrischer Merkmale verwirrt werden.

In dieser Arbeit konzentrieren wir uns auf echte Autoteile, darunter zwei Kurbelgehäuse-Ölabscheiderdeckel aus Gusseisen und Kunststoff sowie ein Luftrohr. Die ausgewählten Objekte weisen einen zunehmenden Schwierigkeitsgrad auf. Das erste Objekt, der Kurbelgehäuse-Ölabscheiderdeckel aus Gusseisen, weist einen hohen Grad an Symmetrie mit mehreren Greifpunkten auf und kann mit einem zylindrischen Teil gegriffen werden, wodurch die Auswirkungen der Ausrichtungsfehler des Roboters um die Achse des Stifts reduziert werden. Das zweite Objekt, der Kurbelgehäuse-Ölabscheiderdeckel aus Kunststoff, weist ebenfalls einen hohen Grad an Symmetrie mit verschie Schließlich weist das Luftrohr eine komplexe Form auf, ist nicht symmetrisch und verfügt nur über zwei Greifpunkte, was die größte Herausforderung für den Roboter darstellt.

Machine Translated by Google

Wir haben uns entschieden, ihre Skelette unter Berücksichtigung einiger Schlüsselpunkte zu modellieren, von denen einige den potenziellen Greifpunkten zum Anheben dieses Objekts mit dem Robotermanipulator entsprechen (siehe oberer rechter Teil von Abbildung 1). Um das topologische Skelett (TS) der zu greifenden Objekte zu ermitteln, verwenden wir Lightweight OpenPose [19], dessen Architektur aus drei Hauptkomponenten besteht: einem Merkmalsextraktor, einem TS-Schätzer und einem Part Affinity Fields (PAF)-Netzwerk.

Im Vergleich zum ursprünglichen OpenPose [20] haben wir uns für Lightweight OpenPose entschieden, da der hohe Rechenaufwand der ersteren Methode es weniger für Echtzeitanwendungen auf Geräten mit geringer Rechenleistung geeignet macht. OpenPose verwendet eine zweizweigige, mehrstufige CNN-Architektur. Der erste Zweig prognostiziert Teilkonfidenzkarten (PCM) für Körperteile und der zweite Zweig prognostiziert Teilaffinitätsfelder (PAF), um die Verbindungen zwischen Körperteilen zu modellieren. Die Architektur umfasst mehrere Stufen von Faltungen, um diese Vorhersagen iterativ zu verfeinern, was zu einer hohen Genauigkeit auf Kosten einer erhöhten Rechenlast führt. Light OpenPose hingegen modifiziert die ursprüngliche Architektur, um die Komplexität zu reduzieren und die Effizienz zu verbessern. Der Ansatz reduziert die Anzahl der Faltungsschichten und -stufen und verwendet tiefenweise trennbare Faltungen anstelle von Standardfaltungen, um die Anzahl der Parameter und Operationen zu reduzieren. Darüber hinaus verwendet das Backbone- Netzwerk MobileNet oder ShuffleNet anstelle des schwereren VGG19 oder ResNet, das im ursprünglichen OpenPose verwendet wurde, und optimiert die Berechnung von Teilaffinitätsfeldern, um ein Gleichgewicht zwischen Genauigkeit und Effizienz zu erreichen.

Merkmalsextraktion. Das ursprüngliche Lightweight OpenPose verwendet ein MobileNetV1-Netzwerk , das für die Merkmalsextraktion in Echtzeit optimiert ist. MobileNet ist eine Familie neuronaler Netzwerkarchitekturen, die für die effiziente Bereitstellung auf mobilen und eingebetteten Geräten mit begrenzten Rechenressourcen entwickelt wurden. Das Hauptmerkmal von MobileNet ist die Verwendung von tiefenweise trennbaren Faltungen, die die Anzahl der erforderlichen Parameter und Berechnungen bei gleichzeitig hoher Genauigkeit erheblich reduzieren können.

MobileNetV1 ist zwar ein hocheffektives neuronales Netzwerk, weist jedoch einige Einschränkungen und Nachteile auf, die berücksichtigt werden sollten. Beispielsweise ist die Genauigkeit begrenzt, da es auf ein Gleichgewicht zwischen Modellgröße und Genauigkeit ausgelegt ist. Es erreicht möglicherweise nicht die gleiche Genauigkeit wie größere und komplexere neuronale Netzwerke, insbesondere bei anspruchsvollen Objekten, bei denen die Schlüsselpunkte (Verbindungen) nicht offensichtlich sind. Die in MobileNetV1 verwendete tiefenweise trennbare Faltungsoperation kann weniger ausdrucksstark sein als herkömmliche Faltungsoperationen und ist möglicherweise nicht in der Lage, alle wichtigen Merkmale eines Bil

Aus den oben genannten Gründen schlagen wir in dieser Arbeit vor, MobileNetV1 **für den Schritt der Merkmalsextraktion** durch **MobileNetV3 [21] zu ersetzen .** MobileNetV3 wurde entwickelt, um **die Einschränkungen von MobileNetV1 zu beheben und gleichzeitig die Effizienz beizubehalten. Die Architektur des in dieser Arbeit verwendeten MobileNetV3-Netzwerks ist in Abbildung 3 dargestellt.**

MobileNet V3 hat zwei Hauptvarianten: (1) MobileNet V3-Large, **das für Anwendungen mit höherer Genauigkeit entwickelt wurde und mehr Schichten und Kanäle hat, und (2)** MobileNet V3-Small, das **für ressourcenbeschränkte Umgebungen optimiert ist** und etwas Genauigkeit zugunsten eines geringeren Rechenaufwands einbüßt . **Wir haben uns für die letztere Variante entschieden. MobileNet V3 führt mehrere neue Komponenten ein, wie z. B. Inverted Residual Blocks zur Aufrechterhaltung eines hohen Effizienzgrads, Squeeze-and-Excitation (SE)-Module zur Verbesserung der Darstellungsleistung des Modells durch Neukalibrierung kanalweiser Merkmalsreaktionen und die H-Swish-Aktivierungsfunktion.**

Die Hard-Swish-Funktion ist eine nichtlineare Aktivierungsfunktion, die effizienter sein soll als herkömmliche Aktivierungsfunktionen wie ReLU. Die Hard-Swish-Funktion ist definiert als

$$\textbf{swish(x)} = x\,\frac{\text{ReLU6}(x+3)}{6}, \quad \textbf{(1)}$$

wobei ReLU6(x) = min(max(x, 0), 6) **eine abgeschnittene ReLU-Funktion ist, die Werte zwischen 0 und 6 ausgibt und immer noch ein nichtlineares Verhalten bietet, während die Rechengeschwindigkeit im Vergleich zur Standard-** ReLU- **Funktion erhöht wird.**

Ein weiteres wichtiges Merkmal von MobileNetV3 ist die Verwendung eines Squeeze-and-Excitation-Moduls (SE). Das SE-Modul ist eine einfache und effiziente Möglichkeit, die Darstellungsleistung zu verbessern

Machine Translated by Google

des Netzwerks (d. h. die Fähigkeit, komplexe Muster und Merkmale zu lernen und darzustellen in der Eingabedaten). Es funktioniert durch das Erlernen kanalweiser Skalierungsfaktoren, die verwendet werden, um selektiv Verbesserung der informativen Funktionen im Netzwerk. Das SE-Modul wird zu jedem Engpass hinzugefügt Block in der MobileNetV3-Architektur, der zur Erhöhung der Genauigkeit in Bezug auf zu MobileNetV1.

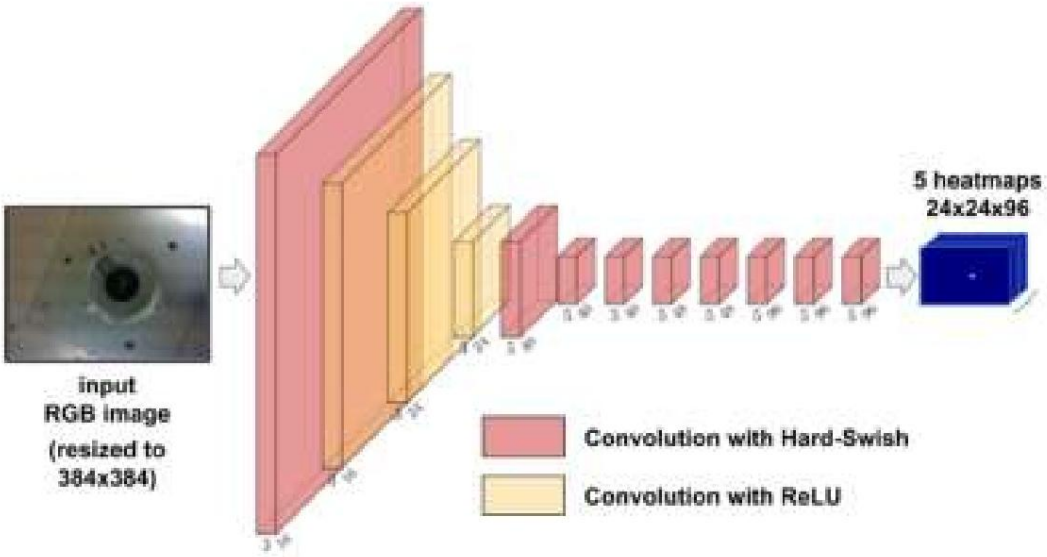


Abbildung 3. MobileNetV3-Architektur.

MobileNetV3 führt außerdem eine neue Technik ein, die als mobiler invertierter Flaschenhals bezeichnet wird. Faltung (MBConv), eine modifizierte Form der tiefenweise trennbaren Faltung wird in MobileNetV1 verwendet. Der MBConv-Block besteht aus drei Arten von Faltungen: einer 1 × 1 Faltung zur Erweiterung der Kanalanzahl, eine Tiefenfaltung zur räumlichen Filterung und eine 1 × 1-Faltung, um die Anzahl der Kanäle wieder auf die ursprüngliche Größe zu reduzieren. Der MBConv-Block enthält auch eine Verknüpfung, die den Farbverlauf fließen lässt direkt von der Eingabe zur Ausgabe. Der MBConv-Block hilft bei der Erhöhung der Ausdruckskraft des Modells in Bezug auf MobileNetV1.

Schließlich enthält MobileNetV3 einen Middle-Flow-Block, der zur Aufrechterhaltung einer hohen Genauigkeit bei gleichzeitiger Minimierung der Anzahl der erforderlichen Berechnungen. Es verwendet auch eine dynamische Faltungsoperation, die sich an die Eingabedaten anpasst. Die Details der Parameter Die für jeden Block verwendeten Werte sind in Tabelle 2 beschrieben. Das Eingabebild wird auf 384 × 384 verkleinert, und die Ausgabe ist ein Satz von 24 × 24 × 96 Feature-Maps, eine für jeden Schlüsselpunkt und eine für der Hintergrund.

Tabelle 2. Die in diesem Dokument verwendete MobileNetV3-Netzwerkarchitektur. HS **= Hard-Swish,** RE **= ReLU,** s **= Schrittweite.**

Operator	Exp-Größe	#aus	ES	[illegible]	s
Eingabe 3842 × 3 conv2d, 3 ×	-	16	-	HS	2
3 1922 × 16 bneck, 3 × 3 962 ×	16	16	x	RE	2
16 bneck, 3 × 3 482 × 24	72	24	-	RE	2
bneck, 3 × 3 482 × 24 bneck,	88	24	-	RE	1
5 × 5 242 × 40 bneck, 5 × 5	96	40	x	HS	2
242 × 40 bneck, 5 × 5 242 × 40	240	40	x	HS	1
bneck, 5 × 5 242 × 48 bneck,	240	40	x	HS	1
5 × 5 242 × 48 bneck, 5 × 5	120	48	x	HS	1
242 × 96 bneck, 5 × 5 242 × 96	144	48	x	HS	1
bneck, 5 × 5	288	96	x	HS	1
	576	96	x	HS	1
	576	96	x	HS	1

Machine Translated by Google

TS-Schätzung. Die Feature-Maps von MobileNetV3 sind die Eingabe, um einen Satz von Kandidaten-Schlüsselpunkten für jeden Objektteil im Bild zu generieren. Tatsächlich erfassen die Feature-Maps die räumlichen Informationen im Eingabebild und bieten eine umfassende Darstellung des Bildes, die zum Erkennen von Schlüsselpunkten verwendet werden kann. MobileNetV3 fügt einen benutzerdefinierten Kopf hinzu, um Schlüsselpunktpositionen vorherzusagen, der aus einer Reihe von Faltungsschichten besteht, die Heatmaps generieren, die vom Backbone extrahierten Features verfeinern und Heatmaps für jeden Schlüsselpunkt generieren. Abbildung 4 zeigt ein Beispiel für die Ausgabe des TS-Schätzers für den Ölabscheiderdeckel des Kurbelgehäuses aus Gusseisen , der aus fünf Heat Jede Heatmap hat die gleiche räumliche Auflösung wie die Feature-Maps und ist auf Werte zwischen 0 und 1 normalisiert. Jeder Pixel in der Heatmap gibt die Wahrscheinlichkeit an, dass der entsprechende Körperteil an dieser Stelle im Bild vorhanden ist.

PAF-Netzwerk. Es verwendet die vom Feature-Extraktor generierten Feature-Maps als Eingabe und gibt einen Satz von PAF-Feature-Maps aus, eine für jedes Paar der erkannten Schlüsselpunkte. Die PAF-Feature-Maps kodieren die Richtung und Stärke der Verbindungen zwischen Schlüsselpunkten mithilfe einer Zweikanaldarstellung, wobei jeder Kanal einen anderen Aspekt der Verbindung kodiert. Insbesondere kodiert ein Kanal den Einheitsvektor, der die Richtung der Verbindung darstellt, während der andere Kanal den Konfidenzwert kodiert, der die Stärke der Verbindung darstellt.

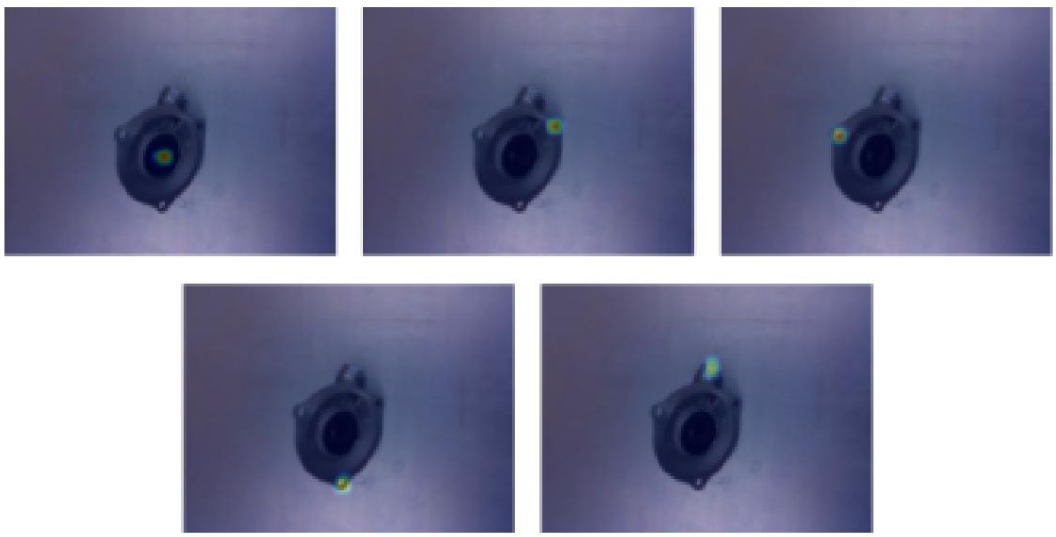


Abbildung 4. Heatmap-Beispiele für Kurbelgehäuse-Ölabscheiderdeckel aus Gusseisen. Es gibt fünf Heatmaps, die den betrachteten Schlüsselpunkten entsprechen.

Endgültige TS-Berechnung. Sobald PAF und Heatmaps generiert sind, werden sie zusammen verwendet, um die einzelnen Schlüsselpunkte in das endgültige TS zu gruppieren. Das endgültige TS wird erhalten, indem zunächst die Kandidatenverbindungen mithilfe der PAFs identifiziert und dann die Verbindungen anhand der Wahrscheinlichkeit bewertet werden, mit der sie eine gültige Verbindung bilden. Die Verbindungen werden dann verwendet, um das endgültige TS zu erstellen, indem die einzelnen Schlüsselpunkte zu einem vollständige

Sobald die Schlüsselpunkte berechnet sind, verwenden wir die Tiefeninformationen zum Erstellen des endgültigen 3D-TS anhand des Satzes von Schlüsselpunkten aus Lightweight OpenPose. Abbildung 5 zeigt einige Beispiele endgültiger TSs für die drei betrachteten Objekte, wobei die **Robustheit des vorgeschlagenen TS-Extraktionsansatzes in Bezug auf verschiedene Ansichten des** Objekts, photochrome Änderungen und teilweise Okklusionen **erkennbar ist . Der Ansatz funktioniert auch mit mehreren Instanzen des Objekts.**

Machine Translated by Google

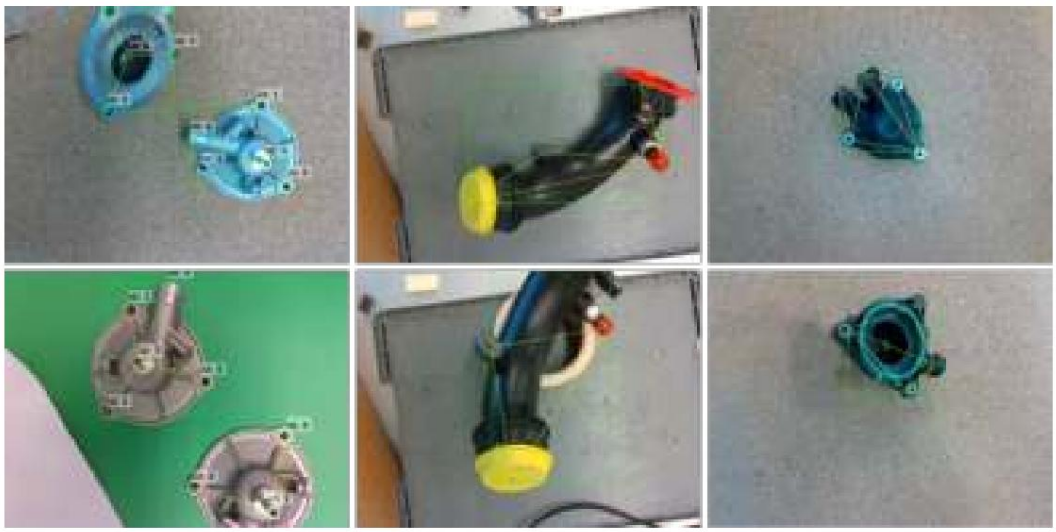


Abbildung 5. Beispiele für TS-Extraktion an verschiedenen Objekten: der gusseiserne Kurbelgehäuse-Ölabscheiderdeckel links, das Luftrohr in der Mitte und der Kunststoff-Kurbelgehäuse-Ölabscheiderdeckel rechts. Unser Ansatz ist robust gegenüber unterschiedlichen Ansichten desselben Objekts, gegenüber photochromen Änderungen un

3.3. Auswahl der Greifposition **Nach**

der Auswahl der Nk **Schlüsselpunkte für die TS-Extraktion werden diese Schlüsselpunkte auch im CAD-Modell des Objekts mithilfe einer 3D-Modellierungssoftware identifiziert. Dies führt zur Generierung einer nominalen dreidimensionalen Darstellung des TS,** TSN, **im CAD- Koordinatensystem** Ff . **Darüber hinaus werden die Posen in** Ff **aller möglichen** Ng **Greifreferenzrahmen (siehe Abbildung** 6) über die (4 **× 4) homogene Transformationsmatrix [22],** f **T , ausgedrückt.**

gj , j = **1, . . . ,** Ng, **können auf dem Modell lokalisiert werden. Der Übersichtlichkeit halber nehmen wir an, dass jeder Greifpunkt mit einem Schlüsselpunkt zusammenfällt.**



Abbildung 6. Die Greifrahmen für die betrachteten Objekte: Kurbelgehäuse-Ölabscheider-Deckel aus Gusseisen (obere Reihe), Kurbelgehäuse-Ölabscheider-Deckel aus Kunststoff (mittlere Reihe) und Luftrohr (untere Reihe).

Machine Translated by Google

Dann, gegeben alle möglichen Kombinationen von drei Schlüsselpunkten

$$St = ti , i = \mathbf{1, \ldots ,} Nt = \binom{Nk}{\mathbf{3}} \mathbf{:} ti = (Pj , Pl , Pm), \mathbf{j, l,} m \ddot{y} \mathbf{\{1, \ldots , Nk\},} j \ddot{y}= l \ddot{y}= m ,$$

für jedes Tripel ti **wird eine Ebene über ein daran angehängtes Koordinatensystem identifiziert, deren Pose** f **durch die homogene Transformationsmatrix T** ti bezeichnet wird . **Für jedes Greifreferenzsystem kann die relative Pose in Bezug auf die** i-te Ebene bestimmt werden als

$$\mathbf{T}^{ti} = \mathbf{T}_{gj\,ti}^{\ddot{y}1} \, f \, \mathbf{T}_{gj} . \quad \mathbf{(2)}$$

Somit kann für jeden Greifpunkt eine Liste von Nt **Transformationsmatrizen T** $_{gj}^{ti}$ **berechnet werden, die die Greifrahmenpositionen im ebenen Rahmen darstellen. Dieser Satz von Operationen, zusammengefasst in Algorithmus 1, wird nur einmal ausgeführt.**

Algorithmus 1: Vorverarbeitungsalgorithmus

Eingabe : TSN, Ng, Nt , $\mathbf{T}_{gj}^{F}$ (j = **1, . . . ,** Ng)

Ausgabe: T $_{gj}^{ti}$ (i = **1, . . . ,** Nt ; j = **1, . . . ,** Ng) **1 für jedes Tripel**

ti **von Schlüsselpunkten in** TSN **do** f **2 Berechne T** ti **3 end für 4 für** i =

1, . . . , Nt **do für** j = **1, . . . ,** Ng

do ÿ1 ff ti **T = TT**

gj ti gj

5

6

7 Ende für

8 Ende für

9 Rückkehr T gj ^{ti}

Zur Laufzeit werden folgende Schritte ausgeführt:

1. Zur Unterscheidung der Objekte wird **ein YOLO-Detektor [23] eingesetzt. YOLO wurde gewählt, da es schneller als klassifikatorbasierte Systeme ist, aber eine ähnliche Genauigkeit aufweist und Vorhersagen mit einer einzigen Netzwerkauswertung trifft, indem die Objekterkennung als einzelnes Regressionsproblem betrachtet wird, was zu einer hohen Genauigkeit führt. Darüber hinaus kann YOLO mehrere Objekte gleichzeitig in einem Bild erkennen und klassifizieren.**

2. Der aktuelle 3D-TS, TSC, **wird extrahiert.**

3. Der Greifpunkt, der der Kamera am nächsten ist, p 4a. Wenn $_{c)}^{C}$ **wird als das Beste ausgewählt. mindestens 3 Schlüsselpunkte sichtbar sind, ein Satz von drei Schlüsselpunkten,** tk , **die dient zur Berechnung entsprechende Ebene im Kamerarahmen, T** tk $_{danke.}^{C}$ **und Auswahl der homogenen Transformationsmatrix, T, wird** $_{gcc}$, **das die Greifposition im ebenen Rahmen identifiziert. Dann, das Verfahren mit Schritt 5 fortgesetzt. 4b. Wenn nur 2 oder weniger Schlüsselpunkte sichtbar sind, beginnt der Roboter, sich kreisförmig um die Mitte des Objektbegrenzungsrahmens zu bewegen, um ein neues Bild aus einem anderen Blickwinkel aufzunehmen . Dann kehrt das Verfahren zu Schritt 1 zurück.**

5. Die Greifpose im Kamerabild wird berechnet als

$$\mathbf{T}_{gcc}^{C} \, tk = \mathbf{TT}_{gcc}^{C} \, tk .$$

Dieses Verfahren ist in Algorithmus 2 zusammengefasst.

Definieren wir die homogene Transformationsmatrix T $^{t}_{C}$, **d. h. die konstante homogene Matrix, die die Transformation zwischen dem Kamerabild und dem Endeffektorbild durchführt ,** die über die in Abschnitt **3.1 beschriebene Kalibrierungsmethode erhalten wurde .**

Machine Translated by Google

Algorithmus 2: Laufzeitalgorithmus objT

Eingabe: für jedes Objekt gj ^{tj}

Ausgabe: T $^{c}_{gcc}$

1 Bild von der Kamera abrufen 2

Aktuelles Objekt im Bild erkennen 3 TSC ÿ **TS**

ti extrahieren **4 T** = **T ÿ Liste**

der Transformationen relativ zum aktuellen Objekt gj **i=1,...,Nt ; j=1,...,Ng abrufen ÿ Greifpunkt extrahieren, der der Kamera am nächsten ist**

5 p $^{c}_{cc}$

6 wenn sichtbare Schlüsselpunkte **ÿ 3**

7 **dann** tk **ÿ extrahiere ein Tripel sichtbarer Schlüsselpunkte in** TSC

8 Berechnen Sie T tk ^{c}

9 T gcc **extrahieren** ^{tk} **aus T**

10 sonst

11 Bewegen Sie den Roboter zu einem anderen Standpunkt

12 GEHE ZU 1

13 Ende, wenn

14 T $^{c}_{gcc}$ tk **ÿ T** $^{c}_{tk\,gcc}$ **T**

15 Rückkehr T $^{c}_{gcc}$

Um die Greifposition im Inertialsystem zu erfassen, T ^{c} **wird wie folgt transformiert** gcc

$$Tgcc = Te\ \mathbf{T}\ {}^{t}_{c}\mathbf{T}\ {}^{c}_{gcc}\ , \tag{3}$$

wobei Te **die homogene Matrix ist, die die Stellung des Endeffektors im Inertialsystem darstellt.**

Anmerkung 1. Es ist zu beachten, dass das obige Verfahren auch dann anwendbar ist, wenn der Greifpunkt nicht mit einem Schlüsselpunkt zusammenfällt. Allerdings muss eine weitere konstante Transformation angewendet werden, um den Greifpunkt mit einem der Schlüsselpunkte zu verknüpfen, die zur Ebene gehören.

3.4. Roboter greifen **Um**

den Greifvorgang auszuführen, muss der Endeffektor angewiesen werden, seinen Referenzrahmen am Greifreferenzrahmen auszurichten. Die Flugbahn wird durch die Zuweisung einer Folge von drei Punkten geplant: Der erste ist die Sichtposition des Roboters, der mittlere ist der Annäherungspunkt , d . h. ein Punkt, der entlang der z- **Achse des Greifreferenzrahmens in einem Abstand von 10 cm zum Ursprung positioniert ist, und der letzte ist die geschätzte Greifposition, pˆ die gewünschte Position des Endeffektors** $_{G}$**. Genauer gesagt, pe,d** (t), **ist definiert als**

$$\mathbf{pe,d}\ (t) = \begin{cases} \ddot{y}\ \mathbf{(pa}\ \ddot{y}\ \mathbf{p0}\ \mathbf{)\ f\ddot{u}r\ 0}\ \ddot{y}\ t\ \ddot{y}\ ta\ \mathbf{p0} + \ddot{y}\mathbf{pa}\ \ddot{y}\ \mathbf{p0}\ \ddot{y}\ \dfrac{s1(t)}{\ddot{y}\ pa\ \ddot{y}\ \mathbf{p0}} & \\ + \ddot{y}\mathbf{p\hat{}}\ \dfrac{s2(t)}{\ddot{y}\ pa\ \ddot{y}_{G}}\ \mathbf{(p\hat{}}\ pa_{G}\ \ddot{y}\ pa\ )\ \mathbf{f\ddot{u}r}\ ta < t\ \ddot{y}\ tf & \end{cases} . \tag{4}$$

wobei p0 die Betrachtungsposition und pa **die Position des Annäherungspunkts ist,** s1(t) (s2(t)) **ist die** Bogenlänge **von p0 nach** pa **(von** pa **zu pˆ Pfadpunkten, sowohl für** s1(t) als **auch** für s2(t), **das** $_{G}$**). Um eine kontinuierliche Beschleunigung und Geschwindigkeit bei der Zeitgesetz kann als Polynom fünften Grades entworfen werden.**

Bei den Zeitpunkten ist tf **die Dauer der Bewegung und** ta **der Zwischenzeitpunkt am Annäherungspunkt, der so gewählt wird, dass bis zum Annäherungspunkt eine schnelle Bewegung erfolgt und in der Nähe des Objekts eine langsame Bewegung.**

Machine Translated by Google

Bezüglich der Endeffektorausrichtung ist geplant, am Annäherungspunkt die gleiche Ausrichtung wie die geschätzte Greifpose Rg zu erreichen und diese Ausrichtung **während des letzten Teils des Pfads konstant zu halten.**

Die geplante Trajektorie in Bezug auf Position und Orientierung ist die Eingabe des geschlossenen inversen Kinematikalgorithmus [22], der darauf abzielt, die Referenzwerte der Gelenkpositionen und -geschwindigkeiten zu berechnen. Bezeichnen wir mit pe (t) **und** Re(t) **die Position bzw. Orientierung des Endeffektors und mit Re,d** (t) **die gewünschte Orientierung des Endeffektors.** Die **Referenzgeschwindigkeiten der Robotergelenke, ÿqr** (t), **werden wie folgt berechnet:**

$$\mathbf{Q}_R(t) = \mathbf{J}^{\dagger}(q(t))(v\ddot{y}\,\mathbf{e,d}(t) + Ke(t)), \quad \textbf{(5)}$$

† wobei J (q(t)) **bezeichnet die rechte Pseudoinverse der Roboter-Jacobi-Matrix, K ÿ IR6×6 ist eine positive definite Matrixverstärkung, ve,d** $^{T}_{e,d}$ **ÿT** $_{e,d}$ **ist der gewünschte Endeffektor linear und** = **pÿ Winkelgeschwindigkeit und e ist der Trackingfehler, definiert als**

$$\mathbf{e} = \begin{matrix} \mathbf{pe,d}\ \ddot{y}\ pe \\ \ddot{y}e\ddot{y}e,d\ \ddot{y}\ \ddot{y}e,d\ddot{y}e\ \ddot{y}\ S(\ddot{y}e,d\ )\ddot{y}e\ , \end{matrix} \quad \textbf{(6)}$$

wobei Qe = *{ÿe ,* ÿe} **und Qe,d** = {ÿe,d , **ÿe,d}** die aus Re **bzw. Re,d extrahierten Einheitsquaternionen sind und S(·) der schiefsymmetrische Matrixoperator ist, der das Kreuzprodu**

Abbildung 7 zeigt eine Flussdiagrammdarstellung des gesamten Prozesses .

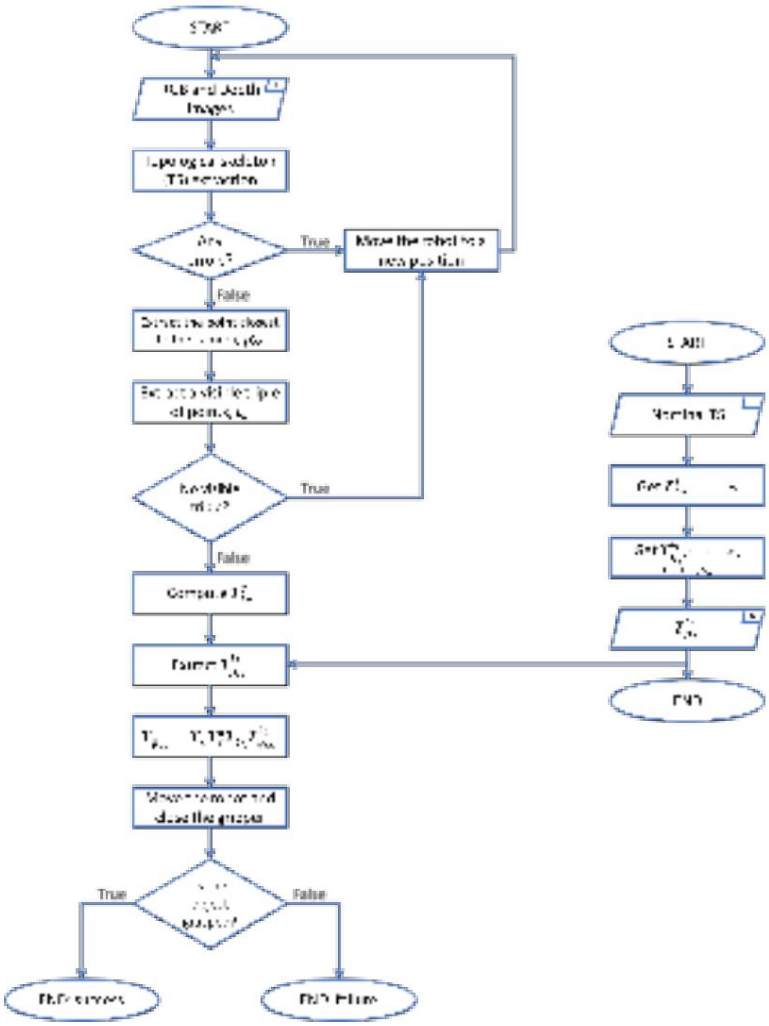


Abbildung 7. Flussdiagrammdarstellung des gesamten Prozesses.

Machine Translated by Google

4. Experimentelle Ergebnisse

Der Versuchsaufbau besteht aus einer Intel RealSense D435 Kamera montiert auf einem Franka Emika Panda Robotermanipulator, gekennzeichnet durch 7 Drehgelenke. Der Roboter kann über das Franka Control Interface (FCI) und die libfranka **C++** gesteuert werden **Open-Source-Bibliothek, die den Roboter direkt mit einer externen Workstation steuert durch eine Ethernet-Verbindung. In dieser Arbeit wird das Metapaket** franka_ros verwendet, das libfranka **in ROS, wurde verwendet. Die Workstation läuft mit Ubuntu 18.04 LTS und einem Echtzeit Kernel auf einer Intel Xeon 3,7 GHz CPU mit 32 GB RAM. Wir haben Experimente durchgeführt mit den drei betrachteten Objekten in Abbildung 6, und die quantitativen Ergebnisse sind weiter unten berichtet.**

4.1. TS-Extraktionsergebnisse

Mit Coco Annotator [24] wurden 5992 Bilder annotiert. Die gelabelten Daten haben wurde in Trainings-, Validierungs- und Test-Sets aufgeteilt, die aus 4618, 229 und 1145 Bildern bestehen, Tabelle 3 zeigt die Anzahl der Bilder in den Trainings-, Validierungs- und Testsätzen für jedes betrachtete Objekt.

Tabelle 3. Anzahl der in Trainings-, Validierungs- und Test-Sets für die betrachteten Objekte verwendeten Beispielbilder.

Objekt	Ausbildung	Validierung	Prüfen
Kurbelgehäuse-Ölabscheiderdeckel aus Gusseisen	1440	60	300
Luftleitung	1406	46	228
Kunststoff-Kurbelgehäuse-Ölabscheiderdeckel	1772	123	617

Die Metrik, die wir zur Bewertung der TS-Erkennung verwendet haben, ist die Object Keypoint Similarity (OKS) **[25], wie folgt definiert:**

$$OKS = \frac{\sum_i \left[\delta(v_i > 0) \exp\left(\frac{-d_i^2}{2 s^2 k_i^2}\right)\right] \delta(v_i > 0)}{\sum_i \delta(v_i > 0)} \quad (7)$$

Wo:

ÿ s **ist der Objektmaßstab;**
ÿ di **ist die Entfernung des vorhergesagten Schlüsselpunkts** i **von der Grundwahrheit;**
ÿ ki **ist eine Konstante pro Schlüsselpunkt, die den Abfall steuert;**
ÿ vi **ist das Sichtbarkeitsflag.**

OKS **wird für jede Probe berechnet, die ein Objekt darstellt. Die Sichtbarkeitsflagge wird Berücksichtigt wird, ob ein Punkt sichtbar ist oder nicht: Wenn der Schlüsselpunkt beschriftet ist,** ist ÿ(vi > 0) **gleich 1, andernfalls ist es gleich 0 ohne verdeckte Schlüsselpunkte zu berücksichtigen.**

In unserem Szenario haben wir OKS verwendet , **um True Positive (TP), False Negative (FN) zu berechnen. Falsch positive (FP) und wahre negative (TN) Erkennungen. Wenn eine Erkennung** OKS > Schwellenwert hat, **es wird als TP betrachtet, andernfalls als FP. Insbesondere haben wir zwei Schwellenwerte berücksichtigt, nämlich 0,5 und 0,75, und berechnete die folgenden Metriken: Präzision, Rückruf, F1-Score, und Average Precision (AP). Tabelle 4 zeigt die Ergebnisse unseres Algorithmus für einen Testsatz von 1145 Bilder.**

Tabelle 4. Ergebnisse des TS-Detektors bei verschiedenen Schwellenwerten für einen Testsatz von 1145 Bildern.

Schwelle	Präzision	Abrufen	F1-Ergebnisse	AP
0,5	0,92	0,90	0,91	0,82
0,75	0,86	0,89	0,87	0,72

Um die Laufzeitleistung unseres TS-Extraktors zu berechnen, testeten wir ihn an einer Teilmenge von 60 Bilder mit einer NVIDIA RTX A5500, mit einer durchschnittlichen Ausführungszeit von 0,012 s und

Machine Translated by Google

eine Standardabweichung von 0,0018 s. Bei einer Teilmenge von 40 Bildern beträgt die durchschnittliche Ausführungszeit bei Verwendung einer NVIDIA QUADRO T2000 0,019 s und die Standardabweichung 0,0025 s.

4.2. Ergebnisse des

Objektdetektors **Zum Trainieren des Objektdetektors haben wir 750 Bilder der Größe 640 × 480 mit dem Annotationstool LabelImg [26] annotiert. Wir haben den Datensatz in Trainings-, Validierungs- und Testsätze mit jeweils 450, 150 und 150 Bildern aufgeteilt. Nach der Trainingsphase beträgt die durchschnittliche Präzision des Testsatzes 97,32 % und die Erfolgsrate 96,70 %. Die Inferenz der Bilder wurde auf einer NVIDIA QUADRO T2000 ausgeführt. Bei einer Teilmenge von 40 Bildern beträgt die durchschnittliche Ausführungszeit 0,323 s, während die**

4.3. Ergebnisse des Robotergreifens

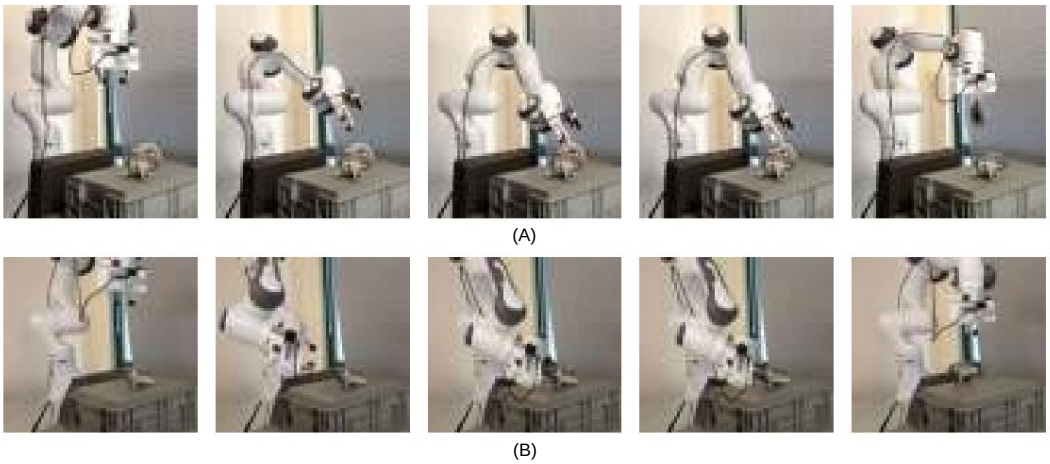
Definieren wir die Schätzfehler bei Greifposition und Orientierung als

$$e_{ep} = p_g^t - \hat{p}_g^t, \quad (8)$$

$$e_{\dot{y}}^e = \dot{y}_g^t - \hat{\dot{y}}_G^t, \quad (9)$$

wobei p_G^t ist die tatsächliche Greifposition, während $\hat{p}_G^t$ ist die Schätzung, die durch die visuelle Algorithmus. Bezüglich der Orientierung $_G(\dot{y}_G^t)$ sind die Eulerwinkel, die aus dem tatsächlichen ist ẏe **(geschätzte) Greifhaltung. Die Verwendung von Eulerwinkeln anstelle von Quaternionen wie in (6) liefert eine klarere physikalische Interpretation der Orientierungsfehler. Das hochgestellte** e **bedeutet , dass die Variablen im Endeffektorrahmen ausgedrückt werden (siehe Abbildu**

Um statistisch signifikante Ergebnisse zu erhalten, wurden 54 Greiftests (20 für den Kurbelgehäuse- Ölabscheiderdeckel aus Gusseisen, 19 für das Luftrohr und 15 für den Kurbelgehäuse-Ölabscheiderdeckel aus Kunststoff) durchgeführt, indem die Objekte in verschiedenen Konfigurationen, unter verschiedenen Lichtbedingungen und mit verschiedenen Hintergründen platziert wurden, sodass der Roboter alle möglichen Greifpositionen erkunden konnte. Ein Greiftest gilt als erfolgreich, wenn der Greifer das Objekt 10 s lang stabil festhält. Eine Reihe von Schnappschüssen des Greifvorgangs ist in Abbildung 8 dargestellt, wobei die obere Reihe einen erfolgreichen Test und die untere Reil



(A)

(B)

Abbildung 8. Momentaufnahmen von zwei Greiffällen. (a) Erfolgreiches Greifen. (b) Fehlschlag.

Nur bei 7 Versuchen (2 mit dem Kurbelgehäuse-Ölabscheiderdeckel aus Gusseisen, 3 mit dem Luftrohr und 2 mit dem Kurbelgehäuse-Ölabscheiderdeckel aus Kunststoff) kam es zu einem Fehlschlag. Somit wurde eine Erfolgsquote von 87,04 % erreicht. Die Tabellen 5–7 zeigen die mittleren Positions- und Orientierungsfehler und die entsprechende Standardabweichung für die erfolgreichen Versuche

Machine Translated by Google

Tabelle 5. Durchschnittliche Fehler für den Ölabscheiderdeckel des Kurbelgehäuses aus Gusseisen.

Erfolgreiche Tests	t_{px} [mm] e	t_{py} [mm] e	$t_{\ddot{y}x}$ [Grad]	$t_{\ddot{y}y}$ [Grad]
1	0,734	9.471	3.128	3.515
2	6.867	0,717	3.284	3.091
3	1.081	5.857	6.092	12.433
4	13.032	10.752	3.132	6.256
5	1.796	0,775	6.801	0,77
6	3.832	0,759	3.913	3.546
7	3.629	1.531	2,981	3.416
8	1.324	6.747	0,674	0,92
9	3.073	7.437	0,62	35.308
10	1.293	1.021	0,952	1.605
11	3.916	1.631	0,391	6.217
12	4,62	2.697	3.044	4.508
13	2.368	4.09	9.12	3.953
14	0,463	6.314	2.456	10,075
15	1.816	2,214	2,217	5.806
16	2,958	4,751	1,842	7.763
17	2,86	8,736	1,354	4.138
18	3,018	0,747	13,855	2.16
Bedeuten	3,26	4.236	3.659	6.415
Standardabweichung	2,820	3.278	3.329	7.607

Tabelle 6. Mittlere Fehler für die Luftleitung.

Erfolgreiche Tests	$t_{t_{px}}$ [mm]	$t_{t_{py}}$ [mm]	$t_{t_{\ddot{y}x}}$ [Grad]	$t_{t_{\ddot{y}y}}$ [Grad]
	0,043	0,989	2.439	8.109
1	9,586	17,223	4.597	0,555
2	1,098	2,612	16.723	13.204
3 4	3,777	6,188	19.545	10.702
5	7,213	0,227	3.663	1.042
6	3,882	1,516	1.684	0,315
7	2,897	2,176	11.183	1.733
8	0,144	1.932	10.255	7.891
9	2.41	1.412	12.057	7.392
10	5.042	0,797	12.674	20,755
11	2,989	2.809	11.516	26.216
12	14.5	4.903	0,988	2.294
13	11.149	1.207	5.192	0,53
14	3.736	5.841	9.211	25.563
15	0,818	8.494	12.022	20.307
16	0,636	14.538	15.122	7.086
Bedeuten	4.37	4.554	9.304	9.606
Standardabweichung	4.085	4.843	5.447	8.801

Für alle Objekte haben die Positionsfehler entlang der Z-Achse des Endeffektorrahmens wurden nicht berichtet, da sie aufgrund der Objektgeometrie vernachlässigbar sind. Aus dem gleichen Grund Die Orientierungsfehler um die Z-Achse des Endeffektorrahmens können vernachlässigbar sein für die Kurbelgehäuse-Ölabscheiderdeckel aus Gusseisen und das Luftrohr.

Bei einigen Tests wurden große Fehler festgestellt, vor allem entlang der y-Achse des Endeffektorrahmens , aber das Objekt wurde erfolgreich gegriffen, da der Greifer gekennzeich durch parallele Finger, und Fehler entlang der Schließrichtung werden eher toleriert.

Die Systemausfälle können in zwei Hauptkategorien unterteilt werden:

1. Fehler im Zusammenhang mit fehlenden (siehe Abbildung 9a,e) oder ungenauen (siehe Abbildung 9b,d,f) Schlüsselpunkten Erkennung oder Vorhersage und falsche Tiefenschätzung.

2. Fehler bei der Poseneinschätzung, die zum Verrutschen des Objekts führen können.

Machine Translated by Google

Tabelle 7. Durchschnittliche Fehler für die Kunststoff-Abdeckung des Kurbelgehäuse-Ölabscheiders.

Erfolgreiche Tests	t_{px} [mm] e	t_{py} [mm] e	$t_{\ddot{y}x}$ [Grad]	$t_{\ddot{y}y}$ [Grad] e	$t_{\ddot{y}z}$ [Grad]
1	13.453	1.036	7.726	4.594	10.505
2	1.795	8.396	11.419	16.797	1,977
3	1.387	5.306	2.214	3.451	11.761
4	7.493	5.909	3.41	5.945	7.285
5	0,799	6.753	3.439	12.744	2,62
6	2,997	1.521	2.03	4.885	9.134
7	5.897	9.452	10,56	1.247	4,62
8	0,973	1.319	7.221	0,007	4.146
9	2.246	5.577	2.104	3.025	2.231
10	1.268	13.631	3.325	4.165	26.108
11	4.547	13.786	2.265	1.418	7,61
12	0,575	1.472	8.638	2.531	10.629
13	3.606	2.318	1.721	6.289	23.799
Bedeuten	3.618	5.883	5.082	5.161	9.417
Standardabweichung	3.487	4.281	3.381	4.527	7.367

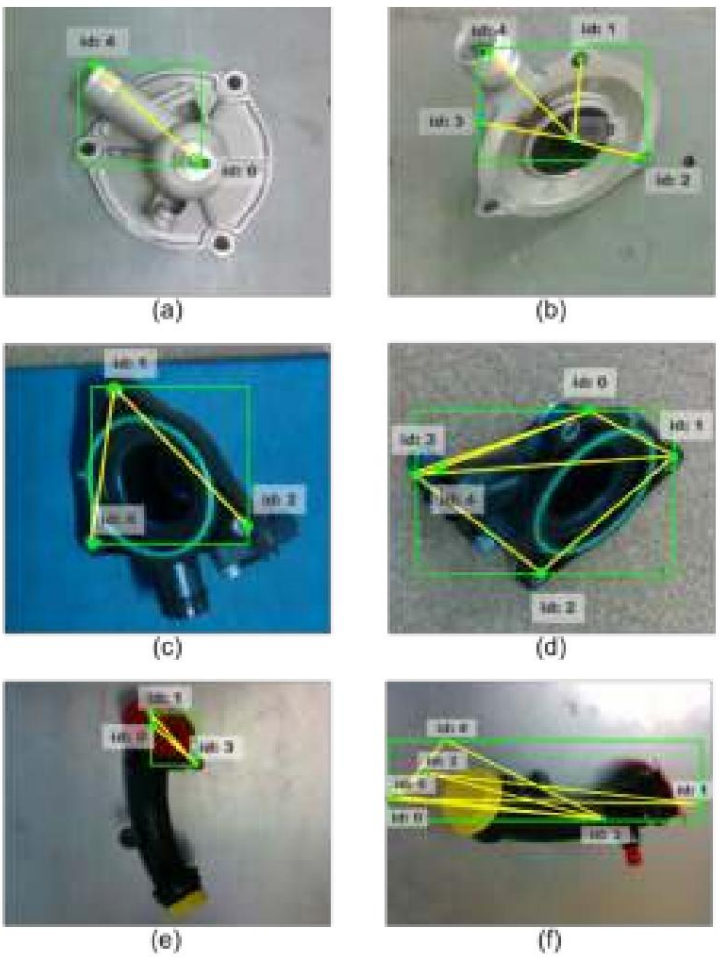


Abbildung 9. Beispiele für fehlende (a,e) und ungenaue (b,d,f) Schlüsselpunkterkennung in TS. Beispiel für fehlende Keypoint-Erkennung, die zu einem erfolgreichen Objektgreifen führen kann (c).

Im Falle einer fehlenden Keypoint-Erkennung, z.B. aufgrund der relativen Objekt-Kamera Position kann der Fehler behoben werden, indem der Blickwinkel der Kamera verändert wird und ein neue Vorhersage (siehe Abschnitt 3.3). In den anderen Fällen ist der Greifvorgang abgeschlossen mit einem Fehler. Da der Roboter den Greiffehler erkennen kann, wird der gesamte Vorgang wiederholt.

Machine Translated by Google

Es ist bemerkenswert, dass gemäß dem im Algorithmus 2 beschriebenen Verfahren das Greifen eines Objekts sogar dann möglich ist, wenn nur drei sichtbare Schlüsselpunkte (siehe Abbildung 9c) korrekt erkannt wurden, vorausgesetzt, dass einer dieser Schlüsselpunkte ein Greifpunkt ist, der sich an einer Stelle befindet, die mit dem verfügbaren Endeffektor gegriffen werden kann.

5. Schlussfolgerungen

In dieser Arbeit wird eine robuste Methode zum Greifen von Teilen mit komplexer Geometrie in einem industriellen Szenario vorgeschlagen. In einer solchen Umgebung sind die Herausforderungen beim Greifen auf das Vorhandensein von Unsicherheiten in der Position des zu greifenden Objekts und auf die Wahrnehmung von Rauschen aufgrund seines Materials zurückzuführen. Insbesondere haben wir uns auf reale Autoteile mit komplexen Geometrien und reflektierenden Oberflächen konzentriert, die Rauschen in der Tiefenkarte verursachen. Die vorgeschlagene Lösung basiert auf einem TS-Extraktionsnetzwerk, das in Echtzeit eine graphenbasierte Darstellung des Objekts erstellen kann. In einem Argumentationsschritt wird entschieden, ob die aktuelle Ansicht des Objekts für das tatsächliche Greifen gut genug ist oder ob sich der Manipulator bewegen muss, um das Objekt besser greifen zu können . Es wurden quantitative Experimente mit einem 7-DoF-Roboter und drei verschiedenen komplex geformten Autoteilen durchgeführt, die zeigen, dass der vorgeschlagene Ansatz schnell und robust ist. Die hohe Genauigkeit und Echtzeitfähigkeit dieses vorgeschlagenen Ansatzes machen ihn zu einer geeigneten Lösung für industrielle Anwendungen, bei denen eine schnelle und genaue Leistung erforde

Aufgrund der Komplexität der betrachteten Objekte ist ein vollständiger quantitativer Leistungsvergleich mit anderen in der Literatur vorhandenen Ansätzen kaum durchführbar.
Der Testdatensatz ist jedoch öffentlich verfügbar, um zukünftige Vergleiche anzustellen.

In Zukunft möchten wir die Integration der Tiefendaten in den TS- Extraktionsprozess untersuchen, um die 3D-Positionen der Schlüsselpunkte direkt zu erhalten. Darüber hinaus könnte auch die Objekterkennungsphase in das TS-Extraktionsverfahren integriert werden.

Beiträge der Autoren: Konzeptualisierung, AP, MS, DDB und FP; Methodik, AP, MS, DDB und FP; Software, AP, MS und DDB; Untersuchung, AP und MS; Validierung, AP und MS; Formale Analyse, AP, MS und DDB; Schreiben – Originalentwurf, AP und MS; Schreiben – Überprüfung und Bearbeitung, DDB und FP; Überwachung, DDB und FP; Projektverwaltung, DDB und FP; Mittelbeschaffung, FP Alle Autoren haben gleichermaßen zu dieser Arbeit beigetragen. Alle Autoren haben **die veröffentlichte Version des Manuskripts gelesen und ihr zugestimmt .**

Finanzierung: Diese Forschung wurde vom italienischen Ministerium für Universität und Forschung im Rahmen des Stipendiums PRIN 2022 PNRR MELODY (Multi-robot collaborativE manipuLation suppOrting DisassemblY tasks) Nr. P2022XALNS unterstützt.

Datenverfügbarkeitserklärung: Der Quellcode des TS-Detektoransatzes ist öffentlich verfügbar unter https://github.com/apennisi/CoGP-TS. Der ROS-basierte Quellcode unseres Ansatzes ist öffentlich verfügbar unter https://github.com/sileom/graspingWithSkeleton.git. Unter **https://tinyurl.com/bdhyf493** sind **mehrere Videos der Experimente verfügbar . Die Testset-Bilder sind unter https://tinyurl verfügbar. com/bdxs8n7z. Alle in der beschriebenen Strategie verwendeten Modelle sind öffentlich verfügbar und können von https://tinyurl.com/3a4nnc88** heruntergeladen werden. **(Alle Links wurden am 23. Juli 2024 aufgerufen).**

Danksagungen: Die Autoren möchten Alessandro Lorenzo, Simona D'Amato und Antonio Giardiello für ihre Hilfe beim Bildannotationsprozess danken.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Abkürzungen

In diesem Manuskript werden die folgenden Abkürzungen verwendet:

CNN	**Faltungsneuronales Netzwerk**
Schärfentiefe	**Freiheitsgrad**
FCI	**Franka-Steuerungsschnittstelle**
FCN	**Vollständig gefaltetes Netzwerk**
FN	**Falsch negativ**

Machine Translated by Google

FP	**Falsch-Positiv**
GG-CNN	**Generatives Greif-Faltungs-Neuralnetzwerk**
MBConv	**Mobile umgekehrte Engpassfaltung**
OKS	**Ähnlichkeit der Objekt-Schlüsselpunkte**
PAF	**Teileaffinitätsfelder**
ROS	**Roboter-Betriebssystem**
TN	**Wahres Negativ**
TP	**Wahres Positiv**
TS	**Topologisches Skelett**

Verweise

1. **Sileo, M.; Bloisi, DD; Pierri, F. Echtzeit-Objekterkennung und -erfassung durch Hintergrundsubtraktion in einem industriellen Szenario In Proceedings des 2021 IEEE 6th International Forum on Research and Technology for Society and Industry (RTSI), virtuell, 6.–9. September 2021; S. 283–288.**
2. **Sileo, M.; Bloisi, DD; Pierri, F. Greifen fester Industrieobjekte mittels 3D-Registrierung.** Machines **2023,** 11, **396. [CrossRef]**
3. **Costanzo, M.; De Maria, G.; Lettera, G.; Natale, C. Können Roboter ein Supermarktregal auffüllen?: Bewegungsplanung und Greifkontrolle.** IEEE-Roboter. Autom. Mag. **2021,** 28, **61–73. [CrossRef]**
4. **Asif, U.; Tang, J.; Harrer, S. GraspNet: Ein effizientes Convolutional Neural Network zur Echtzeit-Grifferkennung für Geräte mit geringem Stromverbrauch. In Proceedings of the IJCAI, Stockholm, Schweden, 13.–19. Juli 2018; S. 4875–4882.**
5. **Zhang, H.; Tan, J.; Zhao, C.; Liang, Z.; Liu, L.; Zhong, H.; Fan, S. Eine schnelle Erkennungs- und Greifmethode für mobile Manipulatoren basierend auf verbessertem schnellerem R-CNN.** Ind. Robot. Int. J. Robot. Res. Appl. **2020,** 47, **167-175. [CrossRef]**
6. **Mahler, J.; Liang, J.; Niyaz, S.; Laskey, M.; Doan, R.; Liu, X.; Ojea, JA; Goldberg, K. Dex-net 2.0: Deep Learning zur Planung robuster Greifvorgänge mit synthetischen Punktwolken und analytischen Greifmetriken. In Proceedings of the Robotics: Science and Systems (RSS), Cambridge, MA, USA, 12.–16. Juli 2017.**
7. **Pinto, L.; Gupta, A. Supersizing self-supervision: Lernen, mit 50.000 Versuchen und 700 Roboterstunden zu begreifen. In Proceedings der 2016 IEEE international conference on robotics and automation (ICRA), Stockholm, Schweden, 16.–21. Mai 2016; S. 3406–3413.**
8. **Schmidt, P.; Vahrenkamp, N.; Wächter, M.; Asfour, T. Greifen unbekannter Objekte mithilfe tiefer Convolutional Neural Networks auf Basis von Tiefenbildern. In Proceedings der 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australien, 21.–25. Mai 2018; S. 6831–6838.**
9. **Morrison, D.; Corke, P.; Leitner, J. Closing the Loop for Robotic Grasping: Ein generativer Grasp-Synthese-Ansatz in Echtzeit. In Proceedings of the Robotics: Science and Systems (RSS), Pittsburgh, PA, USA, 26.–30. Juni 2018.**
10. **Morrison, D.; Corke, P.; Leitner, J. Erlernen robusten, reaktiven Robotergreifens in Echtzeit.** Int. J. Robot. Res. **2020,** 39, **027836491985906. [Querverweis]**
11. **Yu, F.; Koltun, V. Multiskalige Kontextaggregation durch erweiterte Faltungen.** arXiv **2015, arXiv:1511.07122.**
12. **Dune, C.; Marchand, E.; Collowet, C.; Leroux, C. Aktive grobe Formschätzung unbekannter Objekte. In Proceedings der 2008 IEEE/RSJ International Conference on Intelligent Robots and System, Nizza, Frankreich, 22.–26. September 2008; Seiten 3622–3627.**
13. **Kraft, D.; Pugeault, N.; Ba¸seski, E.; POPOVIC, M.; Kragi´c, D.; Kalkan, S.; Wörgötter, F.; Krüger, N. Geburt des Objekts: Erkennung der Objekthaftigkeit und Extraktion der Objektform durch Objekt-Aktionskomplexe.** Int. J. Humanoid Robot. **2008, 5, 247–265. [CrossRef]**
14. **Detry, R.; Ek, CH; Madry, M.; Piater, J.; Kragic, D. Verallgemeinerung von Griffen über teilweise ähnliche Objekte. In Proceedings der 2012 IEEE International Conference on Robotics and Automation, St. Paul, MI, USA, 14.–18. Mai 2012; S. 3791–3797.**
15. **Bloisi, DD; Pennisi, A.; Iocchi, L. Hintergrundmodellierung im maritimen Bereich,** Mach. Vis. Appl. **2014,** 25, **1257–1269. [CrossRef]**
16. **Marchand, E.; Spindler, F.; Chaumette, F. ViSP für visuelle Servosteuerung: Eine generische Softwareplattform mit einer breiten Palette an Robotersteuerungsfähigkeiten .** IEEE Robot. Autom. Mag. **2005,** 12, **40–52. [CrossRef]**
17. **Kannala, J.; Brandt, SS Ein generisches Kameramodell und Kalibrierungsverfahren für konventionelle, Weitwinkel- und Fischaugenobjektive.** IEEE Trans. Pattern Anal. Mach. Intell. **2006,** 28, **1335–1340. [CrossRef] [PubMed]**
18. **Tsai, RY; Lenz, RK Eine neue Technik zur vollständig autonomen und effizienten 3D-Robotik-Hand-/Augenkalibrierung.** IEEE Trans. Robot. Autom. **1989, 5, 345–358. [CrossRef]**
19. **Osokin, D. Echtzeit-2D-Mehrpersonen-Pose-Schätzung auf der CPU: Lightweight OpenPose.** arXiv **2018, arXiv:1811.12004.**
20. **Cao, Z.; Hidalgo, G.; Simon, T.; Wei, SE; Sheikh, Y. OpenPose: Echtzeit-2D-Pose-Schätzung für mehrere Personen mithilfe von Teileaffinität Felder.** IEEE Trans. Pattern Anal. Mach. Intell. **2021,** 43, **172–186. [CrossRef] [PubMed]**
21. **Howard, A.; Sandler, M.; Chu, G.; Chen, L.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Suche nach MobileNetV3.** arXiv **2019, arXiv:1905.02244.**
22. **Siciliano, B.; Sciavicco, L.; Villani, L.; Oriolo, G.** Robotik – Modellierung, Planung und Steuerung; **Springer: London, UK, 2009.**
23. **Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. Man schaut nur einmal hin: Einheitliche Objekterkennung in Echtzeit. In Proceedings der IEEE-Konferenz zu Computer Vision und Mustererkennung, Las Vegas, NV, USA, 27.–30. Juni 2016; S. 779–788.**
24. **Brooks, J. COCO Annotator. 2019. Online verfügbar: https://github.com/jsbroks/coco-annotator/ (abgerufen am 23. Juli 2024).**

Machine Translated by Google

25. Ronchi, MR; Perona, P. Benchmarking und Fehlerdiagnose bei der Multi-Instance-Pose-Estimation. arXiv **2017, arXiv:1707.05388.**
26. Heartexlabs; Lin, T. LabelImg. 2015. Online verfügbar: https://github.com/heartexlabs/labelImg (abgerufen am 23. Juli 2024).

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Verbesserte aktive Störungsunterdrückung für permanente Magnet-Synchronmotor

Zhiwei Huang [1], Yuanhao Cheng 1,2,*, Si Chen 1,2, Xuhui Zhang [3], Jiawei Xiang 1,2 und Sun'an Wang [4]

[1] Hochschule für Maschinenbau und Elektrotechnik, Wenzhou-Universität, Wenzhou 325060, China; 22461440036@stu.wzu.edu.cn (ZH); 20200316@wzu.edu.cn (SC); jwxiang@wzu.edu.cn (JX)
[2] Pingyang-Institut und intelligente Fertigung, Wenzhou-Universität, Wenzhou 325035, China 3
Shaanxi-Schlüssellabor für intelligente Überwachung elektromechanischer Minenausrüstung, Xi'an-Universität für Wissenschaft und Technologie, Xi'an 710064, China; zhangxh@xust.edu.cn
[4] Fakultät für Maschinenbau, Xi'an Jiaotong-Universität, Xi'an 710049, China; sawang@xjtu.edu.cn
* Korrespondenz: yhcheng@wzu.edu.cn

Zusammenfassung: Um die Regelleistung eines Permanentmagnet-Synchronmotors (PMSM) bei externen Störungen zu verbessern, wird ein verbesserter Algorithmus zur aktiven Störungsunterdrückung (IADRC) vorgeschlagen. Da die nichtlineare Funktion im herkömmlichen ADRC-Algorithmus an den Knickpunkten nicht glatt genug ist, was sich direkt auf die Regelleistung auswirkt, wird eine innovative nichtlineare Funktion vorgeschlagen, um die Konvergenz und Stabilität effektiv zu verbessern. Auf dieser Grundlage wird die vorgeschlagene IADRC erstellt, und vergleichende Simulationsergebnisse mit ADRC und anderen IADRC zeigen, dass eine schnellere Reaktionsgeschwindigkeit, höhere Genauigkeit und stärkere Robusth

Schlüsselwörter: verbesserte ADRC; nichtlineare Funktion; PMSM; Geschwindigkeitsregelung

1. Einleitung

Aufgrund der Vorteile einer hohen Leistungsdichte, eines einfachen Aufbaus, einer leichten Wartung und einer bequemen Drehzahlregelung werden Permanentmagnet-Synchronmotoren (PMSMs) häufig in Fahrzeugen mit alternativer Antriebskraft, computergesteuerten Werkzeugmaschinen (CNC) und anderen hochwertigen Geräten eingesetzt [1,2]. Die Proportional-Integral-Differential-Regelungstechnik (PID) wird in zahlreichen Branchen eingesetzt und steuert das traditionelle Drehzahlregelungssystem des PMSM. Zwar hat der PID-Regler die Vorteile einer einfachen Struktur und einer leichten Implementierung [3,4], doch ist es für eine genaue Regelung notwendig, ein genaues mathematisches Modell des geregelten Objekts zu erstellen. Bei der Modellierung von PMSMs wird die Motorstruktur jedoch häufig vereinfacht. Diese Vereinfachung hat sich als unzureichend erwiesen, wenn der PMSM mit niedriger Drehzahl arbeitet, was zu einer deutlichen Verschlechterung der Regelungsleistung führt, wie in Referenz [5] hervorgehoben wird. Darüber hinaus lassen sich Rauschen und Störungen durch PID nicht gut unterdrücken, sodass Wissenschaftler verschiedene verbesserte PID- Systeme vorgeschlagen haben [6,7]. Durch die Optimierung der Anpassungsparameter kann die Regelleistung deutlich verbessert werden, die Entstörungsfähigkeit, Reaktionsgeschwindigkeit und Genauigkeit müssen jedoch noch verbessert werden.

Die moderne Regelungstheorie bietet eine neue Lösung zur Leistungssteigerung von Regelungssystemen. Parallel zur Weiterentwicklung der PMSM-Regelungstechnologie wurden verschiedene effektive Regelungsmethoden vorgeschlagen. Laut Ref. [8] verbessert der nichtlineare PID- Regler die Regelungsleistung durch adaptive Fehlersignaltransformation und dynamische Anpassung der PID-Parameter, wodurch der Regler das Rauschen im Eingangssignal effizient dämpfen und einen Kompromiss zwischen Reaktionsgeschwindigkeit und Überschwingen, der herkömmlichen PID-Reglern innewohnt, angehen kann. Darüber hinaus wird Fuzzy-PID eingesetzt, um die Robustheit des Systems zu verbessern [9]. Darüber hinaus hat sich die Gleitmodusregelung (SMC) in den letzten Jahren zweifellos als entscheidender Fortschritt in der Motorregelungstechnik herausgestellt [10,11]. Sie kann die Regelungswirksamkeit von PMSM-Systemen verbessern und gleichzeitig eine schnelle Konvergenz aufweisen

check for updates

Zitat: Huang, Z.; Cheng, Y.; Chen, S.; Zhang, X.; Xiang, J.; Wang, S. Verbesserte aktive Störungsunterdrückungssteuerung für Permanentmagnet-Synchronmotoren. Electronics 2024, 13, 3023. https:// doi.org/10.3390/electronics13153023

Wissenschaftlicher Redakteur: Fabio Corti

Erhalten: 25. Juni 2024
Überarbeitet: 25. Juli 2024
Akzeptiert: 30. Juli 2024
Veröffentlicht: 31. Juli 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.
Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

Machine Translated by Google

und hohe Störfestigkeit. Aufgrund der Rattereigenschaften der Gleitmodus-Variablenstruktur ist die Frage, wie dieses Rattern wirksam unterdrückt werden kann, jedoch zu einem wichtigen Thema geworden. Robuste Steuerung ist auch eine Klasse von Methoden, die auf unsichere Systeme abzielen, und es wurden robuste Steuerungsgesetze auf der Grundlage von Hÿ- Paradigmen und µ-Synthese untersucht, die eine gute Störungsunterdrückung und Robustheit gegenüber Parametervariationen erreichen [12]. Obwohl diese Algorithmen die Regelungswirkung des PMSM-Systems effektiv verbessern, haben sie immer noch Probleme, eine bessere Regelungsleistung zu erreichen, wenn es um die Berücksichtigung von Unsicherheiten in den Modellen und den Umgang mit externen Störungen geht.

Active Disruption Rejection Control (ADRC) ist eine nichtlineare Steuerungsstrategie, die von Jingqing Han vorgeschlagen wurde und die Unsicherheiten und externe Störungen effektiv bewältigen kann [13]. Die Nützlichkeit dieses Ansatzes zeigt sich in seiner überlegenen Entstörungsfähigkeit , seiner starken Anpassungsfähigkeit an Änderungen und Unsicherheiten der Systemparameter, seiner hohen Steuerungspräzision , seiner schnellen Reaktionsgeschwindigkeit, seiner einfachen Struktur und seiner einfachen Implementierung [14] und hat in vielen technischen Bereichen große Aufmerksamkeit erregt [15–17]. ADRC besteht hauptsächlich aus drei Komponenten: Tracking Differentiator (TD), Extended State Observer (ESO) sowie Nonlinear State Error Feedback Control Law (NLSEF). Der TD kann das Eingangssignal glätten, den Rauscheinfluss reduzieren und differenzielle Informationen des Signals bereitstellen, damit der Regler besser auf Änderungen im System reagieren kann. Der ESO ist die Kernkomponente von ADRC und wird zur Schätzung der Zustandsvariablen und Gesamtstörungen des Systems verwendet. Er ist in der Lage, Unsicherheiten und externe Störungen in Echtzeit zu beobachten und zu mildern und so die Robustheit und Widerstandsfähigkeit des Systems gegen Störungen zu verbessern. Darüber hinaus generiert das NLSEF Steuerbefehle basierend auf beobachteten Zustandsfehlern sowie Störungsinformationen, um eine genaue Steuerung des gesteuerten Objekts zu erreichen. Die nichtlineare Rückkopplungssteuerungsstrategie des NLSEF kann die dynamische Reaktion des Systems drastisch verbessern. ADRC übernimmt die Essenz des traditionellen PID und kann Systeme mit Nichtlinearität und Unsicherheit effektiv steuern, ohne sich auf genaue Modelle zu verlassen [18,19]. Beispielsweise schlugen Fang et al. [20] basierend auf ADRC und Gleitmodussteuerung eine neue integrierte Designmethode für Geschwindigkeits- und Positionsschleifen vor und realisierten die Geschwindigkeits- und Positionssteuerung von PMSMs. Während ein minimales Systemüberschwingen erreicht wird, bleibt die Komplexität der Parameterkonfiguration eine Herausforderung. Um dieses Problem zu lösen, führten Li et al. [21] einen Gleitmodus- ADRC ein, bei dem der ESO durch einen nichtlinearen Störungsbeobachter ersetzt wurde, um die Parameteranpassungen zu rationalisieren . Aufbauend auf ihrer Forschung verwendeten Ge et al. [22] einen Partikelschwarmoptimierungsalgorithmus , um ADRC für die Lageregelung von Raumfahrzeugen zu verbessern. Anstatt sich auf eine empirische Auswahl zu verlassen, werden die Parameter nun durch mathematische Optimierung bestimmt . Ref. [23] schlug eine Fuzzy-ADRC mit gebrochener Ordnung vor, die die Beschränkungen herkömmlicher Steuerungsmethoden (wie geringe Präzision und langsame Reaktion) mildert und zur Steuerung der Bewegung der zahlreichen Gelenke des Manipulators verwendet wird. Ramlavi und Chidan [24] untersuchten den linearen Tracking-Steuerungsansatz mit aktiver Störungsunterdrückung, um die Tracking-Leistung und Robustheit des Controllers zu verbessern und lösten die Modellunsicherheit und Umgebungsstörungen eines einrädrigen Roboters. Aufbauend auf der Forschung von Partovibakhsh und Liu [25] verbesserten Guo und Zhao [26] die Tracking-Fähigkeiten mobiler Roboter durch den Kompensationsmechanismus der ADRC. Lv [27] schlägt eine Methode zur Fuzzy-Steuerung mit automatischer Störungsunterdrückung (Fuzzy-ADRC) für einen dreiphasigen Vierarm-Wechselrichter vor, um Motordrehmomentpulsationen unter komplexen Betriebsbedingungen zu unterdrücken. Als Reaktion auf die schlechte Regelleistung, die durch feste Parameter in ADRC für ein lagerloses PMSM verursacht wird, schlägt Wang [28] eine dynamische Parameteranpassungsmethode für ADRC vor, die auf einem genetischen Algorithmus und einem Backpropagation-Neuralnetz basiert. Fang [29] schlug eine ADRC-Methode vor, die auf einem verbesserten ESO bas

Obwohl die ADRC-Technologie die Steuerwirkung nichtlinearer und unsicherer Systeme effektiv verbessert, gibt es immer noch Einschränkungen, wie z. B. die Unzulänglichkeit effektiver Parametereinstellungsmethoden , das Fehlen einer Schätzfähigkeit für schnell zeitabhängige Störungen und die von der herkömmlichen ADRC übernommene Fal- Funktion, die an den Umschaltpunkten zwischen dem nichtlinearen Abschnitt und dem linearen Abschnitt nicht glatt genug ist [30,31]. In Anbetracht der Herausforderungen, denen sich PMSM-Steuersysteme mit ADRC-Reglern

Machine Translated by Google

leisten den folgenden Beitrag: **Zur Verbesserung der** Fal- Funktion wird **eine polynomische nichtlineare Funktion in Kombination mit einer Kosinusfunktion** vorgeschlagen , **die die Glätte der nichtlinearen Funktion effektiv verbessert und die Robustheit von ADRC erhöht. Die untersuchten Ergebnisse zeigen, dass der hier vorgeschlagene IADRC im Vergleich zum herkömmlichen ADRC und anderen verbesserten ADRC (IADRC) die Reaktionsgeschwindigkeit, Präzision und Robustheit des Systems erheblich steigert.**

Die Organisationsstruktur dieses Dokuments wird unten umrissen. Abschnitt 2 stellt das mathematische Modell des PMSM vor. Abschnitt 3 präsentiert die Verbesserungsstrategie des ADRC und konstruiert einen IADRC-Controller. In Abschnitt 4 wird ein Entwurf für das PMSM-Geschwindigkeitsregelungssystem unter Verwendung des vorgeschlagenen IADRC vorgestellt. In Abschnitt 5 werden das Simulationsexperiment und die Analyse durchgeführt und die gute Regelungsleistung des vorgeschlagenen IADRC für PMSM-Systeme überprüft.

2. Mathematische Modellierung von PMSM

Die Rotorstruktur und die Permanentmagnetverteilung des PMSM [32] sind **in Abbildung** 1 **dargestellt . Entsprechend ihrer strukturellen Eigenschaften können PMSMs in zwei Typen unterteilt werden: oberflächenmontierte und eingebaute. Der oberflächenmontierte PMSM wird auch als PMSM mit nicht ausgeprägten Polen bezeichnet. Da seine Permeabilität sehr nahe an der Vakuumpermeabilität liegt, ist die Änderung zwischen Reluktanz und Induktivität sehr gering, wodurch der Rotor gute nicht ausgeprägte Poleigenschaften aufweist. Dies führt auch dazu, dass das Magnetfeld des Permanentmagneten eine annähernd sinusförmige Verteilung aufweist und der Motor eine bessere Leistung erbringt. Beim eingebauten PMSM ist die Reluktanz der Längsachse viel höher als die der Querachse, sodass die Induktivität der Längsachse viel niedriger ist als die der Querachse und der Rotor ausgeprägte Poleigenschaften aufweist, weshalb er auch als PMSM mit ausgeprägten Polen bezeichnet wird. Der eingebaute PMSM realisiert eine sensorlose Steuerung durch den Effekt ausgeprägter Pole und erhöht die Leistungsdichte durch das elektromagnetisch Allerdings ist die Rotorstruktur komplex, der magnetische Streuflusskoeffizient groß und die Herstellungskosten hoch.**

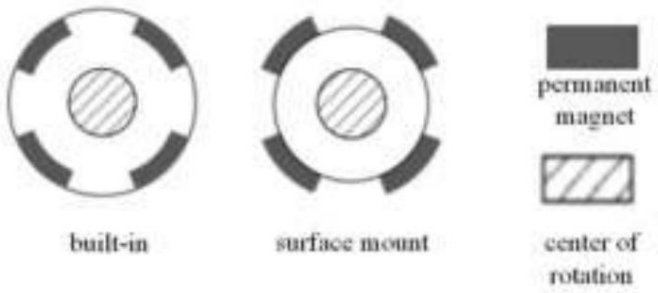


Abbildung 1. Schematische Darstellung der PMSM-Rotorstrukturen.

Aufgrund der unterschiedlichen Strukturen der PMSMs **unterscheiden sich auch ihre Betriebsarten und die Form** ihrer elektromagnetischen Drehmomente. Um eine bessere Steuerwirkung zu erzielen , **ist es daher notwendig, eine Vielzahl unterschiedlicher Steuerschemata entsprechend der internen Strukturen der PMSMs zu formulieren. Angesichts der geringen Kosten der oberflächenmontierten PMSM und der gleichen Induktivität der dq-Achse ist es einfacher, die Gleichung für das elektromagnetische Drehmoment aufzustellen. Daher wurde in dieser Studie die oberflächenmontierte PMS**

Zur Analyse des mathematischen Modells des dreiphasigen PMSM werden folgende Annahmen getroffen dienen der Vereinfachung der Analyse:

(1) Ignorieren Sie die Reluktanz der Stator- und Rotorkerne, ohne Berücksichtigung der Verluste durch Wirbelströme oder Hysterese.

(2) Das Permanentmagnetmaterial ist nichtleitend und seine Permeabilität ist gleich zu dem der Luft.

(3) Der Rotor besitzt keine Dämpfungswicklungen.

Machine Translated by Google

(4) Das Erregermagnetfeld des Permanentmagneten und das Ankerreaktionsmagnetfeld der Dreiphasenwicklung sind sinusförmig über den Luftspalt verteilt.

(5) Die induzierte elektromotorische Kraft in der Phasenwicklung folgt einer sinusförmigen Muster.

Unter der Annahme, dass die oben genannten Bedingungen erfüllt sind, kann **die Spannungsgleichung eines dreiphasigen PMSM im natürlichen Koordinatensystem gemäß der elektromagnetischen** Induktion und dem Kirchhoffschen Spannungsgesetz wie folgt ausgedrückt werden:

$$\begin{aligned} \ddot{y} &= dt\, d\ddot{y}b \frac{d\ddot{y}a}{} + Rsia\ \mathbf{ua} \\ \ddot{y}\ddot{y}\ \mathbf{ub} &= \frac{}{dt} + Rsis \\ \ddot{y}\ddot{y}\ \mathbf{uc} &= \frac{d\ddot{y}c}{dt} + Rsic \end{aligned} \quad (1)$$

Die Statorspannungsgleichung des PMSM in einem zweiphasigen rotierenden Koordinatensystem kann sein erhält man durch die Clarke-Transformation und die Park-Transformation [33]:

$$\begin{aligned} d\ddot{y}q \quad & \frac{d\ddot{y}d}{} + Rsid\ \ddot{y}\ \ddot{y}e\ddot{y}q\ \mathbf{ud} = dt \\ & \frac{}{} + Rsiq + \ddot{y}e\ddot{y}d\ \mathbf{uq} = dt \end{aligned} \quad (2)$$

Die Statormagnetismus-Kettengleichung lautet:

$$\mathbf{\ddot{y}d} = Ldid + \mathbf{\ddot{y}f}\ \mathbf{\ddot{y}q} = Lqiq \quad (3)$$

Bringt man Gleichung (3) in (2) ein, erhält man:

$$\begin{aligned} d\ddot{y}q \quad & \frac{d\ddot{y}d}{} + Rsid\ \ddot{y}\ \mathbf{\ddot{y}eLqiq}\ \mathbf{ud} = dt \\ & \frac{}{} + Rsiq + \ddot{y}e(Ldid + \mathbf{\ddot{y}f}\ )\ \mathbf{uq} = dt \end{aligned} \quad (4)$$

wobei ua, ub, uc die Dreiphasenspannungen der Statorwicklung sind, mit V als Einheit; Rs der Statorwiderstand ist, mit ÿ als Einheit; ia, ib, ic der Dreiphasenstrom der Statorwicklung sind, mit A als Einheit; ÿa, ÿb, ÿc der Fluss der Statorwicklung sind, mit Wb als Einheit; ÿf **die Permanentmagnet-Flussverkettung ist, mit Wb als Einheit; Ld, Lq die Induktivitätskomponenten der Statorinduktivität jeweils auf der dq-Achse sind, mit H als Einheit;** ÿe die **elektrische Winkelgeschwindigkeit des Motors darstellt ,** mit rad/s als Einheit.

Durch die Analyse der Komponenten Widerstandsdrehmoment und dynamisches Drehmoment wird die mechanische Bewegungsgleichung des Motors abgeleitet:

$$J\frac{d\ddot{y}m}{} = \mathbf{Te}\ \ddot{y}\ B\ddot{y}m\ \ddot{y}\ T\ dt \quad (5)$$

wobei J **das Trägheitsmoment,** T **das Lastdrehmoment und** B **der viskose Reibungskoeffizient ist.**

Durch Clarke-Transformation und Park-Transformation ergibt sich die folgende Gleichung des elektromagnetischen Drehmoments:

$$\mathbf{Te} = \frac{\mathbf{3}}{\mathbf{2}}\ pn\ \mathbf{\ddot{y}f}\ \mathbf{iq} + (Ld\ \ddot{y}\ Lq)idiq \quad (6)$$

Dabei ist Te **das elektromagnetische Drehmoment. 3/2 ist der Koeffizient bei Anwendung des Transformationsprinzips mit gleicher Amplitude.** Pn **bezieht sich auf die Gesamtzahl der Polpaare des Motors.**

Die transformierte mechanische Bewegungsgleichung lautet:

$$J\frac{d\ddot{y}m}{dt} = \frac{\mathbf{3}}{\mathbf{2}}\ \mathbf{pn}\ \mathbf{\ddot{y}f}\ \mathbf{iq} + (Ld\ \ddot{y}\ Lq)idiq\ \ddot{y}\ B\ddot{y}m\ \ddot{y}\ T \quad (7)$$

3. Entwurf eines verbesserten ADRC-Controllers

Die traditionelle ADRC behandelt Unsicherheit, interne Systemstörungen und externe Störungen zusammen als Gesamtstörung des Systems. Und sie kann ech

Machine Translated by Google

Zeitgenaue Schätzung und Kompensation dieser Störung mithilfe von ESO und NLSEF. Dieser Ansatz basiert nicht auf einem exakten Modell der Anlage und verbessert die Entstörungsfähigkeiten erheblich. Allerdings ist die Fal- Funktion des herkömmlichen ADRC an den Umschaltpunkten zwischen dem nichtlinearen und dem linearen Abschnitt nicht glatt genug, was dazu führt, dass der Steuereffekt nicht ideal ist. Daher **wird in diesem Artikel ein verbesserter ADRC-Regler** vorgeschlagen , **der auf einer verbesserten** Fal- Funktion basiert.

3.1. Verbessertes Fal-Funktionsdesign

Die Fal- **Funktion ist eine nichtlineare Funktion, die bei ADRC eine wichtige Rolle spielt. Sie ermöglicht ADRC nicht nur, Systemstörungen besser abzuschätzen und auszugleichen und so die Störungsbeständigkeit des Systems zu verbessern, sondern ermöglicht auch, die Verstärkung bei großen Fehlern zu verringern und bei kleinen Fehlern zu erhöhen. Sie ist vorteilhaft, um die Reaktionsgeschwindigkeit des Systems auf große Fehler und die Regelgenauigkeit auf kleine Fehler zu verbessern und gleichzeitig die Robustheit und Anpassungsfähigkeit des Systems zu verbessern.**

Bei herkömmlichem ADRC kann fal(e, **ÿ,** ÿ) **als entscheidende nichtlineare Funktion sowohl interne als auch externe Störungen des Systems effektiv abschätzen und so das Ausgangssignal des ADRC erzeugen. Die Form der** fal- **Funktion ist in Gleichung (8) dargestellt:**

$$f\,al(e, \ddot{y}, \ddot{y}) = \begin{cases} |e|^{\ddot{y}}\,\text{Vorzeichen}(e), & |e| \\ \dfrac{t}{\ddot{y}^{\,1\ddot{y}\ddot{y}}} & \ddot{y}\,\ddot{y}\,, \; |e| < \ddot{y} \end{cases} \quad \textbf{(8)}$$

Wenn *ÿ* **= 0,001 ist** und *ÿ* **verschiedene Werte annimmt, ist die Kurve der** Fal- Funktion **in Abbildung 2 dargestellt .**

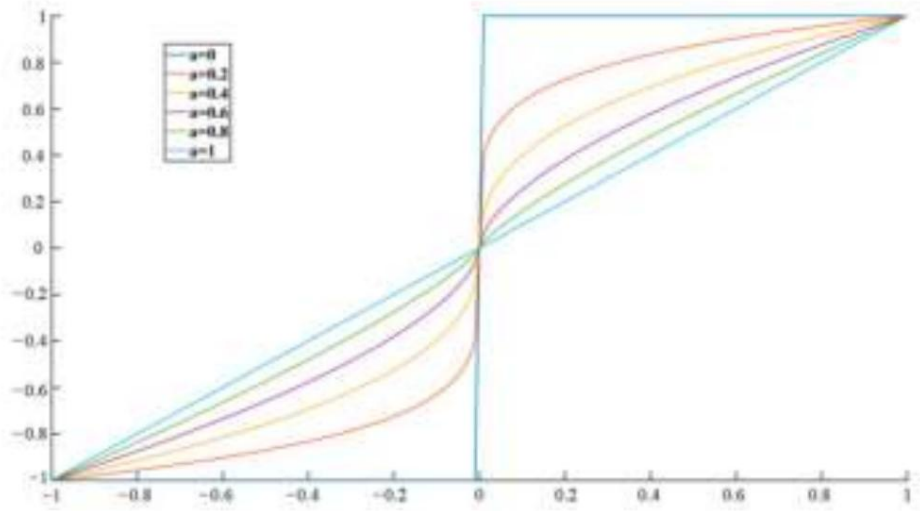


Abbildung 2. Kurven der Funktion fal(e, **ÿ,** ÿ) **unter verschiedenen** ÿ **-Werten.**

Aus der Perspektive der Konvergenz ist der terminale Attraktor |e| $^{\ddot{y}}$ sign(e) konvergiert **in endlicher Zeit , was besonders für die Regelung in der Nähe des Ursprungs** (|e| **< 1) geeignet ist, da dies eine Verstärkungswirkung auf den Fehler hat, aber nicht für die Regelung weiter vom Ursprung entfernt , da die Konvergenz langsamer ist als bei der klassischen linearen Regelung. Der lineare Term** konvergiert exponentiell, was für die Regelung weit vom Ursprung entfernt (|e| **> 1) geeignet ist . In der** Fal- **Funktion wird jedoch die Verwendung des terminalen Attraktorterms und des linearen Terms vertauscht, was dazu führt, dass die** Fal- **Funktion als Ganzes in nicht-endlicher Zeit konvergiert, unabhängig davon, wie klein der anfängliche Fehler e(0) ist. Daher kann der lineare Term verbessert werden, um die Konvergenzfähigkeit zu steigern.**

Aus Gleichung (8) ist ersichtlich, dass sich die Ableitungswerte an den Knickpunkten nach Anwendung der Ableitung der fal- Funktion unterscheiden. Außerdem ist aus Abbildung **2 ersichtlich, dass die** Kurven der fal(e, **ÿ,** ÿ) **-Funktion stetig, aber nicht glatt sind, wenn** e = |ÿ|, **und dass es zu einer plötzlichen Änderung kommt, die zu einer schlechten Regelleistung des Systems**

Machine Translated by Google

Basierend auf den oben genannten Problemen wird in diesem Artikel eine polynomische nichtlineare Funktion Ifal(e, **ÿ,** ÿ) vorgeschlagen, die lineare und trigonometrische Funktionen **kombiniert .** Die Gründe für die Verwendung einer **Kombination aus trigonometrischen**

Funktionen und Polynomen sind folgende: (1) Trigonometrische Funktionen und Polynome haben explizite mat

sie sind analytisch und theoretisch leicht handhabbar.

(2) Durch die geeignete Gestaltung trigonometrischer Funktionen und Polynome kann das Überschwingen des Systems verringert und die Reaktionsgeschwindigkeit sowie Stabilität des Systems verbessert werden.

(3) Obwohl die hinzugefügte trigonometrische Funktion Ifal **komplexer macht als die ursprüngliche** Fal- **Funktion, kann die Berechnungseffizienz durch die Nachschlagetabelle und Interpolationsmethoden verbessert und unter der Voraussetzung der Gewährleistung einer Echtzeitleistung und geringerer Hardwareanforderungen ein besserer Kontrolleffekt erzielt w**

Der Funktionsausdruck wird unten angezeigt:

$$\text{Wenn f al}(e, \ddot{y}, \ddot{y}) = \begin{cases} \dfrac{|e|^{\ddot{y}}}{\ddot{y}\,\ddot{y}\ddot{y}3_{\text{Vorzeichen(e)}}} & , |e| \ddot{y}\ \ddot{y} \\ \mathbf{k1}\cos(e\ddot{y}\ddot{y})\ \mathbf{k1} & \end{cases} \quad \begin{cases} e^{\ddot{y}\ddot{y}2\,\ddot{y}}\ \mathbf{k2}\cos(e\ddot{y}\ddot{y}) \\ + \mathbf{k3\ddot{y}} + \ddot{y}\ddot{y}\ddot{y}\ddot{y}1\ \mathbf{e}, |e| < \ddot{y}\ \mathbf{k2} \end{cases} \tag{9}$$

Die Designidee der Funktion ist wie folgt: Da das Hauptproblem der ursprünglichen Fal- **Funktion darin besteht, dass die Segmente nicht glatt genug sind, wird die ursprüngliche lineare Funktion in diesem Dokument in eine Polynomfunktion geändert, die mit einer trigonometrischen Funktion und einer linearen Funktion kombiniert wird, wenn** |e| ÿ *ÿ* . **Unter der Bedingung, dass die** Fal- **Funktion an den Segmenten kontinuierlich ist und die Werte der Ableitungsfunktion gleich sind, wird dann die Beziehung zwischen den drei konstanten Koeffizienten ve**

Basierend auf den obigen Überlegungen kann man erkennen, dass, wenn e = |ÿ|, **die Funktionswerte auf beiden Seiten der** Ifal- **Funktion gleich sind, und dass nach der Ableitung der** Ifal- **Funktion die Ableitungswerte auf beiden Seiten ebenfalls gleich sind.**

Im Lichte der vorgenannten Konzepte ergibt sich folgende Gleichung:

$$\begin{aligned} &\text{I f alÿ}(e, \ddot{y}, \ddot{y}) = \text{I f al+}(e, \ddot{y}, \ddot{y}) \\ &\text{I f alÿ ÿ}(e, \ddot{y}, \ddot{y}) = \text{I f alÿ +}(e, \ddot{y}, \ddot{y}) \end{aligned} \tag{10}$$

Der Ableitungsausdruck der Ifal- **Funktion lautet nach der Berechnung wie folgt:**

$$\text{Wenn f alÿ }(e, \ddot{y}, \ddot{y}) = \begin{cases} \dfrac{\ddot{y}\ddot{y}(\ddot{y}\ddot{y}1)}{\ddot{y}\ \ddot{y}\ddot{y}3\ \text{t}\ (\text{k1}\ddot{y}1)\ \ddot{y}\ \ddot{y}} & \ddot{y}\ddot{y}2\ e \\ (\text{k2}\ddot{y}1) + \ddot{y}\ddot{y}(\ddot{y}\ddot{y}1)\ , |e| \ddot{y}\ \ddot{y} & \end{cases} \quad , |e| > \ddot{y} \tag{11}$$

Durch Kombinieren der Gleichungen (10) und (11) können wir Gleichung (12) ableiten.

$$\frac{\ddot{y}\ \ddot{y}\ddot{y}3\text{t}\ \mathbf{k1}\cos(e\ddot{y}\ddot{y})\ \mathbf{k1}}{\ddot{y}\ \ddot{y}\ddot{y}3\ \text{t}\ (\text{k1}\ddot{y}1)\ \ddot{y}\ \ddot{y}} \quad \begin{matrix} e^{\ddot{y}\ddot{y}2\,\ddot{y}}\ \mathbf{k2}\cos(e\ddot{y}\ddot{y}) + \\ \mathbf{k3\ddot{y}} + \ddot{y}\ddot{y}\ddot{y}\ddot{y}1\ e = |e|\ \mathbf{k2}\ (\text{k2}\ddot{y}1) + \ddot{y}\ddot{y}(\ddot{y}\ddot{y}1) \end{matrix}^{\ddot{y}} \text{Zeichen(e)} \tag{12}$$

Wenn |e| = **ÿ, vereinfacht sich Gleichung (3) zu folgender Form:**

$$\frac{\ddot{y}\ k1 + \ddot{y}\ \ddot{y}}{3\ \mathbf{k1}} \quad \frac{\ddot{y}\ k2+\ddot{y}\ddot{y}2}{\ddot{y}} + \mathbf{k3\ddot{y}} + \ddot{y}\ddot{y} = \ddot{y}\ \mathbf{k2}\ \ddot{y}\ k1+\ddot{y}\ddot{y}3\ \ddot{y}^{\ddot{y}} \tag{13}$$

$$k2+\ddot{y}\ddot{y}2 + \ddot{y}\ddot{y}\ddot{y}\ddot{y}1 = \ddot{y}\ddot{y}\ddot{y}\ddot{y}1$$

Dann wird die Beziehung für k1, k2 und k3 hergeleitet, k2 = **k1** ÿ **1, k3** = **k1** (k1ÿ1) , die in die Ifal- Funktion **eingebracht** $\frac{\mathbf{k1}\ 2\ddot{y}\mathbf{k1}+1}{\mathbf{wird,\ um\ die}}$ ÿ **ÿ, neue nichtlineare Funktion zu erhalten:**

$$\text{Wenn f al}(e, \ddot{y}, \ddot{y}) = \begin{cases} \dfrac{|e|^{\ddot{y}}\ \text{Zeichen(e)}}{\ddot{y}\ \ddot{y}\ddot{y}3\text{t}\ \mathbf{k1}\cos(e\ddot{y}\ddot{y})\ \mathbf{k1}} & , |e| > \ddot{y} \\ e^{\ddot{y}\ddot{y}2\,\ddot{y}}\ \text{k1}\ddot{y}1\cos(e\ddot{y}\ddot{y})\ \ddot{y}\ k1+\ddot{y}\ddot{y}3 + \ddot{y}\ddot{y}\ddot{y}\ddot{y}1\ \mathbf{e}, |e|\ \ddot{y}\ \ddot{y} + (\mathbf{1}\ \ddot{y}\ \ddot{y})\ddot{y}\ \ddot{y} + \text{k1}\ddot{y}1\ \mathbf{k1}\ (\text{k1}\ddot{y}1) & \end{cases} \tag{14}$$

Die Fal- **Funktionskurven vor und nach der Verbesserung werden mit MATLAB 2020a gezeichnet von MathWorks und sind in Abbildung 3 dargestellt .**

Machine Translated by Google

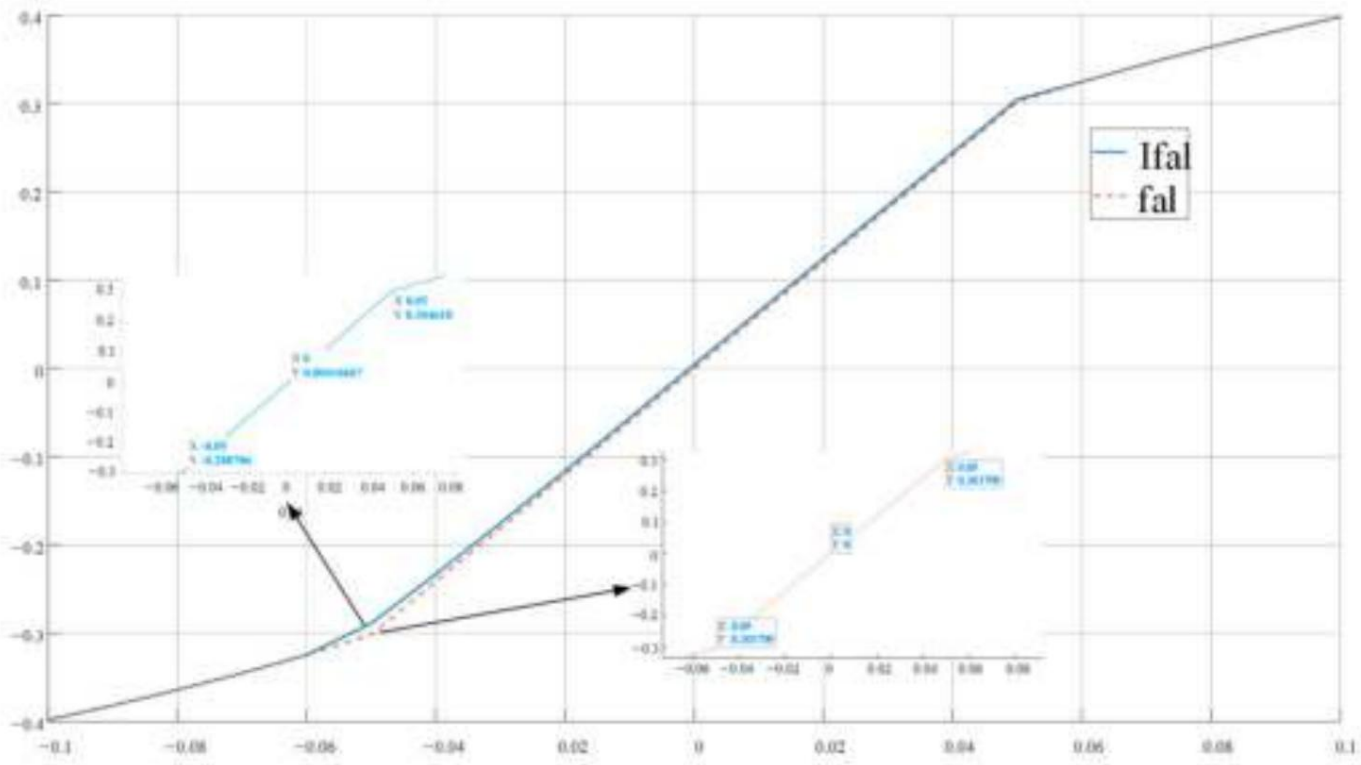


Abbildung 3. Kurven der Funktion fal(e, **ÿ,** ÿ) **vor und nach der Verbesserung.**

Aus der Grafik ist ersichtlich, dass die Kurve der Ifal(e, **ÿ,** ÿ) **-Funktion glatter ist als die der** Fal- **Funktion. Die** Ifal- **Funktion ist an den Punkten, an denen** |e| = ÿ ist, stetig und ableitba **Im Bereich (ÿ0,06; 0,06) ist die Ausgabe der** Ifal- **Funktion höher als die der** Fal- **Funktion, d.ÿh. die** Ifal- **Funktion konvergiert schneller und reibungsloser.**

3.2. Verbessertes ESO-Design (IESO)

Darüber hinaus dient das IESO als Kern für IADRC. Es bietet nicht nur die Zustandsschätzung des PMSM-Systems, sondern klassifiziert auch alle Arten interner und externer Störungen sowie Unsicherheiten, die die Gesamtheit der Störungen des PMSM-Systems umfassen, und schätzt und kompensiert die Gesamtstörung in Echtzeit. Ähnlich wie das ESO stützt sich das IESO ausschließlich auf die Eingabe- und Ausgabedaten des Systems und erfordert kein genaues mathematisches Störungsmodell. Dieser Ansatz verbessert effektiv die Leistung und Robustheit des Steuerungssystems. Die Formulierung des IESO lautet wie folgt:

$$\begin{array}{ll} \ddot{y} & \mathbf{e2} = \mathbf{z21}\ \ddot{y}\ n = \\ \ddot{y} & \dot{z}_{21}\ \ bu(t) + \mathbf{z22}\ \ddot{y}\ \mathbf{k31}\ l\ f\ al(e2, \mathbf{\ddot{y}2, \ddot{y}2}) = \mathbf{\ddot{y}k32}\ l\ f \\ \ddot{y} & \dot{z}_{22}\ \ al(e2, \mathbf{\ddot{y}2, \ddot{y}2}) \end{array} \quad \textbf{(15)}$$

3.3. Verbessertes NLSEF-Design (INLSEF)

Das NLSEF ermöglicht die effektive Kontrolle von Nichtlinearität, Unsicherheit und externen Störungen in ADRC und verbessert die Steuerungsleistung und Robustheit des Systems. Die Fal- **Funktion ist ein wichtiges nichtlineares Element im NLSEF zur Verarbeitung von Fehlersignalen. Sie kann die Verstärkung basierend auf der Fehlergröße adaptiv manipulieren, indem sie die Verstärkung bei erheblichen Fehlern verringert und die Verstärkung erhöht, wenn der Fehler minimal ist, wodurch die Genauigkeit und Robustheit des Steuerungsmechanismus verbessert wird. Der Ausdruck des INLSEF lautet wie folgt:**

$$\begin{array}{l} \mathbf{e1} = \mathbf{z11}\ \ddot{y}\ \mathbf{z21}\ \mathbf{z22} + \mathbf{k2} \\ u(t) = \ddot{y} \quad \overline{W}\text{enn}\ f\ al(e1, \mathbf{\ddot{y}1, \ddot{y}1})\ b \end{array} \quad \textbf{(16)}$$

Machine Translated by Google

wobei z21 eine Beobachtung der dynamischen Reaktion des Reglers darstellt, während z21 eine Beobachtung der Störung bezeichnet , b **der Kompensationsfaktor ist und k2 die Reglerverstärku**

3.4. Entwurf eines PMSM-Drehzahlregelungssystems auf Grundlage des von der IADRC vorgeschlagenen

Abbildung 4 zeigt das schematische Diagramm des von der IADRC vorgeschlagenen PMSM-Drehzahlregelungssystems . Der Controller steuert die Drehzahl und das Drehmoment des PMSM indirekt, indem er den Ausgang des Wechselrichters steuert. Das System verwendet eine doppelte Regelung, wobei die Stromschleife unter PI arbeitet, während die Drehzahlschleife durch die von der IADRC vorgeschlagene Regelung

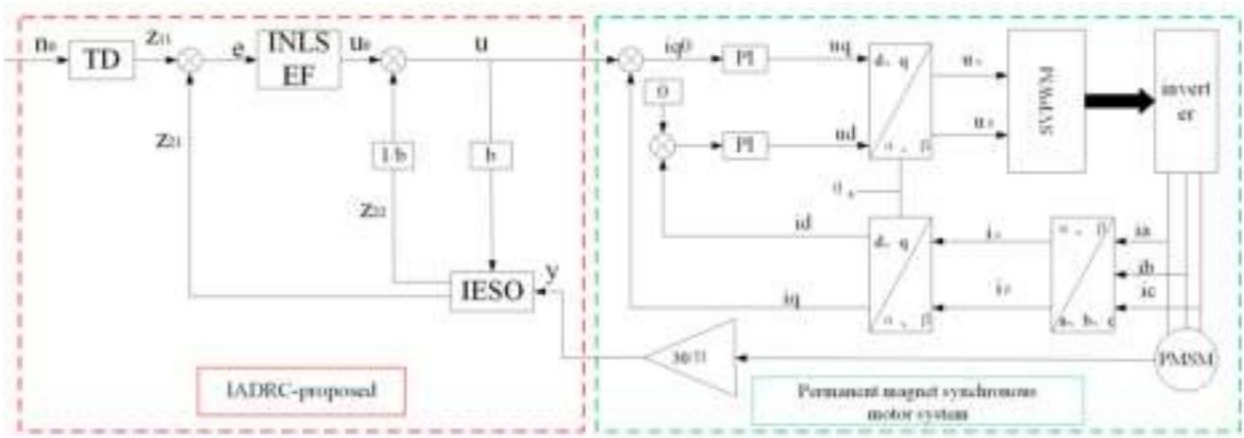


Abbildung 4. Diagramm, das das vom IADRC vorgeschlagene Geschwindigkeitskontrollsystem von PMSM darstellt.

4. Simulation und experimentelle Auswertung

Um die Wirkung des von IADRC vorgeschlagenen Algorithmus auf die PMSM-Drehzahlregelung zu überprüfen, werden die **traditionellen ADRC-, SMC-, Fuzzy-PID- und IADRC-Algorithmen aus Ref. [34]** als Vergleichsalgorithmen verwendet . **Die auf diesen fünf Regelalgorithmen basierenden Modelle des PMSM-Drehzahlregelungssystems werden in MATLAB/Simulink erstellt und anschließend die Simulationsvergleichsexperimente durchgeführt. Die Spezifikationen für das PMSM in den Simulationsmodel sind in Tabelle 1 aufgeführt.**

Tabelle 1. Spezifikationen für PMSM.

Spezifikationen	Wert
Statorphasenwiderstand R (Ohm)	**2,785**
Statorlängsinduktivität LD (H)	**8,5** × 10ÿ3
Statorquerinduktivität LQ (H)	**8,5** × 10ÿ3
Permanentmagnet-Flussverkettung ÿf **(Wb)**	**0,175**
Trägheit J (kg·m2/rad)	**1** × 10ÿ3
Polpaare, P	
Reibungskoeffizient B	**4 1** × 10ÿ4

Die in diesem Manuskript verwendete Methodik wird wie folgt beschrieben: Die erwartete Drehzahl des PMSM beträgt 1500 U/min. Um die Robustheit des von IADRC vorgeschlagenen Algorithmus zu überprüfen, wird zunächst ein Lastdrehmoment hinzugefügt. Wenn die tatsächliche Drehzahl die erwartete Drehzahl erreicht und stabil ist, wird das Lastdrehmoment von 2 Nm **plötzlich nach 0,1 s auf die Ausgangswelle des PMSM angewendet, um die Leerlaufreaktionskurve und die Entstörungsfähigkeit des PMSM unter Einwirkung der oben genannten fünf Steueralgorithmen zu vergleiche**

Die Parametereinstellungen im von IADRC vorgeschlagenen Algorithmus lauten wie folgt. Gemäß den spezifischen Beschreibungen der drei Komponenten von ADRC in Abschnitt 3 werden die beiden Parameter ÿ, *ÿ* **in TD und NLSEF erfahrungsgemäß auf ÿ1 = 0,4, ÿ1 = 0,09, ÿ2 = 0,4, ÿ2 = Die beiden Parameter im ESO werden durch kontinuierliche Simulation und Abstimmung bestimmt: ÿ3 = 0,25, ÿ3 = 0,04. Geschwindigkeitskoeffizient k1 im TD, Reglerverstärkung k2 und Kompensationsfaktor b** des **NLSEF, abgeleitet aus dem mathematischen Modell und den Paramet**

Machine Translated by Google

System, werden als k1 = 5355, k2 = 5000, b **= 1030,66 angenommen. Ebenso die Kalibrierung**
Die Gewinne der Ausgabe im ESO werden als k31 = 8500, k32 = 500.000 angenommen. Die Schlüsselparameter in
Die Parameter für den Controller sind in Tabelle 2 aufgeführt.

Tabelle 2. Die Werte der Schlüsselparameter des Controllers.

Komponente	Spezifikationen	Wert
TD	Tracking-Faktor ÿ1	0,4
	Filterfaktor ÿ1	0,09
	Geschwindigkeitsfaktor k1	5355
NLSEF	Tracking-Faktor ÿ2	0,4
	Filterfaktor ÿ2	0,09
	Reglerverstärkung k2	5000
	Kompensationsfaktor b	1030,66
ESO	Tracking-Faktor ÿ3	0,25
	Filterfaktor ÿ3	0,04
	Die Kalibriergewinne des Ausgangs k31	8500
	Die Kalibrierungsgewinne des Ausgangs k32	500.000

Entsprechend den obigen Parametereinstellungen werden die drei Algorithmen angewendet auf die
Geschwindigkeitsregelung von PMSM. Die Ergebnisse des Experiments sind unten dargestellt:

Aus den Abbildungen **5** **und** **6** **ist ersichtlich, dass in der ersten Startphase des PMSM das IADRC**
Der in diesem Artikel vorgeschlagene Algorithmus hat die beste Wirkung unter den drei Algorithmen. Er hat nicht nur keine
Überschwingen, sondern erreicht und stabilisiert auch die erwartete Geschwindigkeit bei 0,027 s. Die maximale
Der Fehlerschwankungswert beträgt 31,28 U/min, was etwa 2,09 % der erwarteten Drehzahl entspricht.
IADRC in Ref. [34] benötigt 0,045 s, um die erwartete Geschwindigkeit zu erreichen und sich dort zu stabilisieren, und die
Überschwingen beträgt 61,64 U/min, was etwa 4,11 % der erwarteten Drehzahl entspricht. Und die Zeit für die
Die traditionelle ADRC zum Erreichen und Stabilisieren der erwarteten Geschwindigkeit beträgt 0,085 s, und ihr Überschwingen
ist auch die größte, nämlich 119,04 U/min, also etwa 7,89 % der erwarteten Drehzahl. Der SMC
benötigt 0,075 s, um die erwartete Geschwindigkeit zu erreichen und sich dort zu stabilisieren, mit einem maximalen Überschwingen von
423,71 U/min, was etwa 28,25 % der erwarteten Drehzahl entspricht. Der Fuzzy-PID nimmt
0,043 s, um die erwartete Geschwindigkeit zu erreichen und sich dort zu stabilisieren, mit einer maximalen Fehlerschwankung von
30,27 U/min, was etwa 2,02 % der erwarteten Geschwindigkeit entspricht.

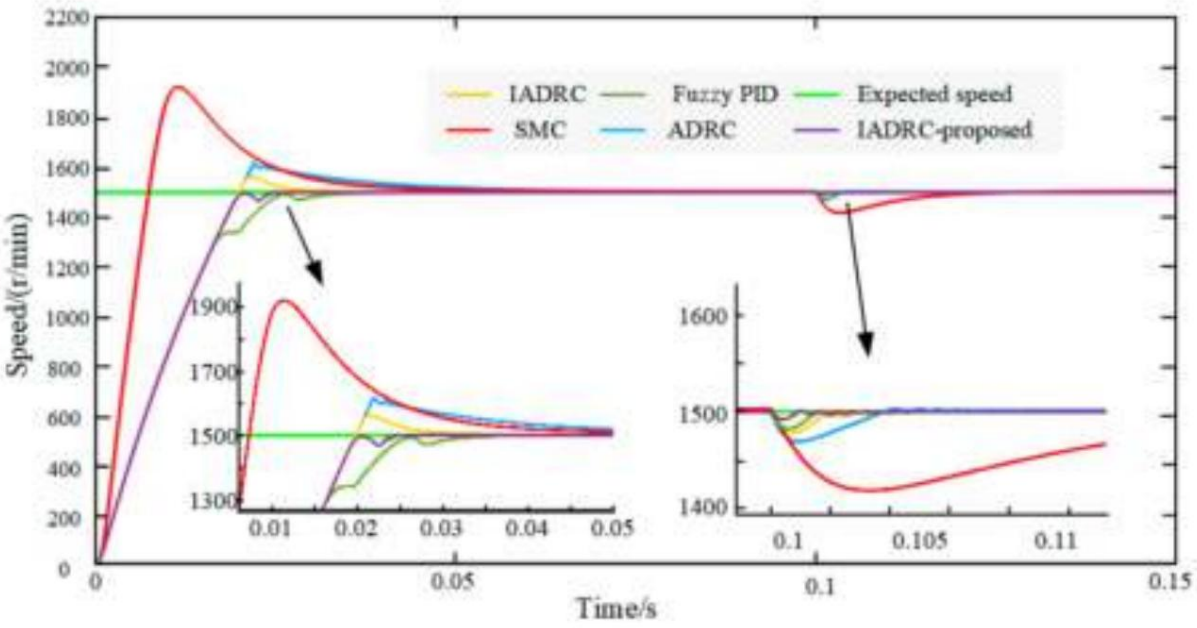


Abbildung 5. Geschwindigkeitsreaktionskurven von 5 Algorithmen ohne Gaußsches Rauschen.

Machine Translated by Google

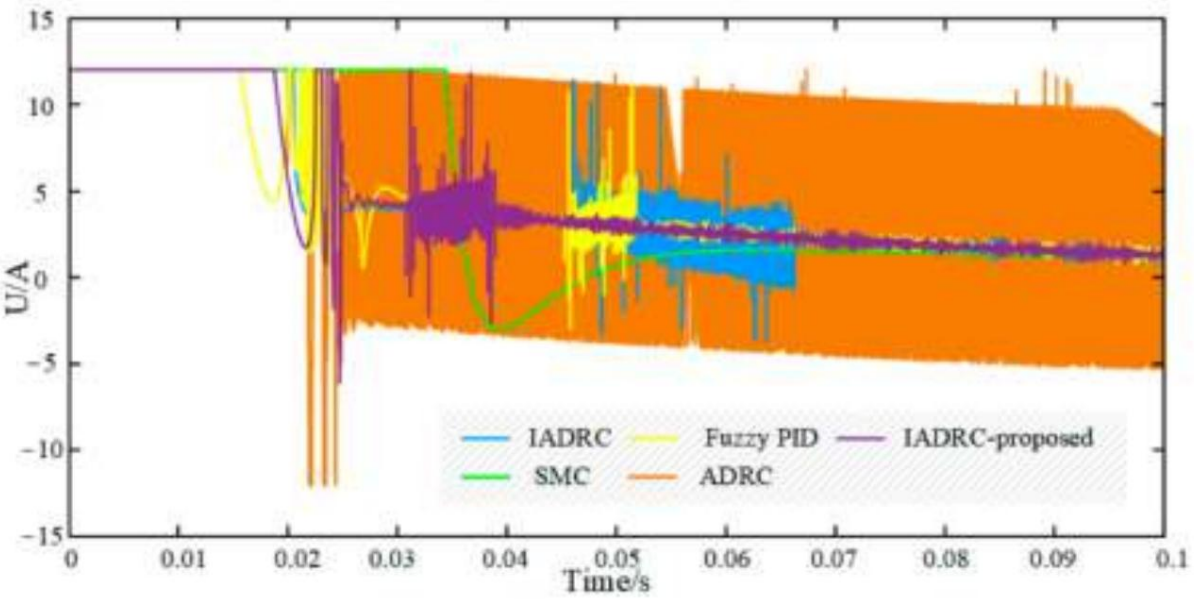


Abbildung 6. Die Anlagensteuerungseingaben von 5 Algorithmen ohne Gaußsches Rauschen.

Wenn sich die Drehzahl des PMSM auf der erwarteten Drehzahl stabilisiert hat, wird das stufenweise Lastdrehmoment nach 0,1 s angelegt. Der in diesem Dokument vorgeschlagene IADRC-Algorithmus bleibt davon grundsätzlich unberührt und die Drehzahlkurve schwankt nur geringfügig. Der maximale Drehzahlfehler beträgt 7,39 U/min, was etwa 0,49 % der erwarteten Drehzahl entspricht, und kehrt dann schnell (0,004 s) zum erwarteten Wert zurück. Die Drehzahlkurve des IADRC-Algorithmus in Ref. [34] erzeugte einen maximalen Drehzahlfehler von 20,73 U/min, was etwa 1,38 % der erwarteten Drehzahl entspricht, und kehrte nach 0,006 s zum erwarteten Wert zurück. Und die Drehzahlkurve des herkömmlichen ADRC-Algorithmus erzeugt einen maximalen Drehzahlfehler von 29,85 U/min, was etwa 1,99 % der erwarteten Drehzahl entspricht, und kehrt nach 0,034 s zum erwarteten Wert zurück. Die Geschwindigkeitskurve des SMC- Algorithmus erzeugt einen maximalen Geschwindigkeitsfehler von 81,98 U/min, was etwa 5,47 % der erwarteten Geschwindigkeit entspricht, und kehrt nach 0,028 s zum erwarteten Wert zurück. Die Geschwindigkeitskurve des Fuzzy-PID-Algorithmus erzeugt einen maximalen Geschwindigkeitsfehler von 18,11 U/min, was etwa 1,21 % der erwartete

Im praktischen Einsatz wird das PMSM-System von verschiedenen Unsicherheitsfaktoren beeinflusst, wie z. B. Umgebungsgeräusche und Widerstandsänderungen, die durch Dauerbetrieb des Motors oder Änderungen der Umgebungstemperatur verursacht werden. Um die Robustheit des vom IADRC vorgeschlagenen Algorithmus bei verschiedenen Parametervariationen und Unsicherheiten zu überprüfen, werden dem System Simulationsexperimente für Rauschstörungen und Motorwiderstandsänderunge

Um die Robustheit des vom IADRC vorgeschlagenen Algorithmus zu überprüfen, enthält dieses Dokument auch Vergleichsexperimente des Anlagensteuereingangs U **unter Gaußschem Rauschen und vergleicht ihn mit anderen Algorithmen. Das Gaußsche Rauschen wird am Anfang hinzugefügt und das Signal wird auf einen Mittelwert von 1 und eine Varianz von 200 eingestellt. Andere Motorparameter sind in den Tabellen 1 und 2 aufgeführt. Die Simulationsergebnisse sind in den Abbildungen 7 u**

Vergleicht man die Abbildungen 5–8, ist es offensichtlich, dass sich die Geschwindigkeitsreaktionskurven und die Systemleistung U vor **und nach dem Hinzufügen des Gaußschen Rauschens für alle fünf Algorithmen nicht wesentlich ändern. Darüber hinaus haben die Geschwindigkeitsreaktionskurven und die Systemleistung** U , die durch den vom IADRC vorgeschlagenen Algorithmus **gesteuert werden, die kürzeste Schwingungszeit und die kleinste Schwingungsamplitude, was die Stab**

Um die Auswirkungen von Änderungen des Motorwiderstands auf den IADRC-Vorschlag weiter zu validieren, werden die Motorwiderstände auf 2,5875, 2,875 bzw. 3,1625 festgelegt. Andere Parameter stimmen mit Tabelle 1 überein. Die experimentellen Ergebnisse sind in Abbildung 9 dargestellt.

Machine Translated by Google

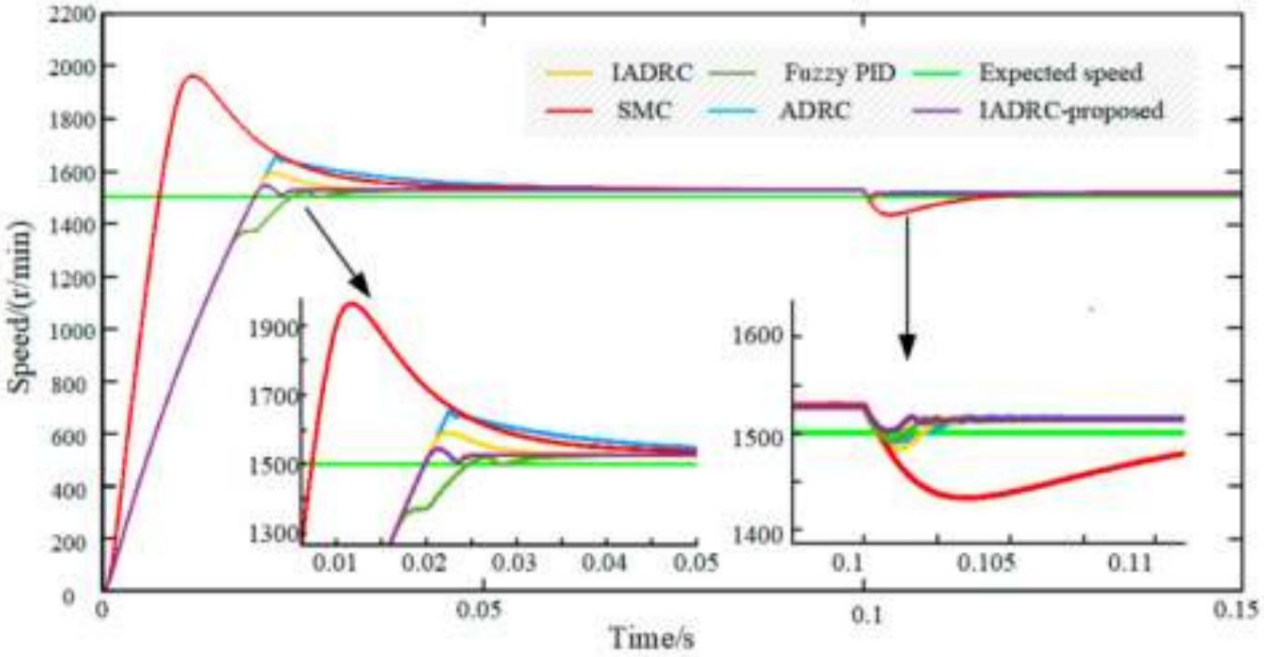


Abbildung 7. Geschwindigkeitsreaktionskurven von 5 Algorithmen unter Gaußschem Rauschen.

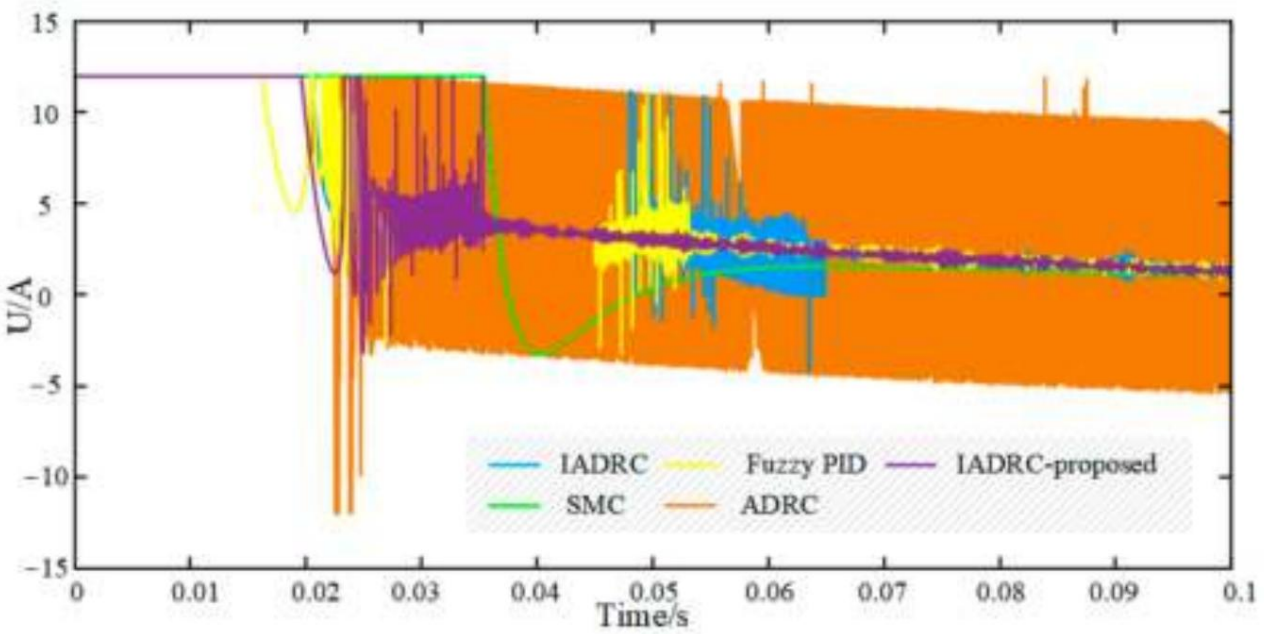


Abbildung 8. Die Anlagensteuerungseingaben von 5 Algorithmen unter Gaußschem Rauschen.

Aus Abbildung 9 ist ersichtlich, dass sich Überschwingen und Zeit bis zum Erreichen der Stabilität zwar alle ändern, die Änderung aber gering ist. Dies deutet darauf hin, dass der Kontrolleffekt nach Änderungen der äußeren Bedingungen immer noch sehr gut ist, was die Robustheit des vom IADRC vorgeschlagenen Algorithmus bestätigt .

Zusammenfassend ist damit erwiesen, dass unser vorgestellter IADRC eine verbesserte Reaktionsfähigkeit, höhere Regelgenauigkeit, stärkere Robustheit und Entstörungsfäh Die spezifische Leistung ist wie folgt: Unser proprietärer IADRC-Algorithmus ist 40,00 % besser als der IADRC in Ref. [34], 68,24 % besser als der traditionelle ADRC, 64,00 % besser als SMC und 37,21 % besser als Fuzzy-PID in Bezug auf die Schnelligkeit. In Bezug auf die Genauigkeit ist er 49,25 % besser als der IADRC in Ref. [34], 73,72 % besser als der traditionelle ADRC, 92,62 % besser als SMC und 3,35 % schlechter als Fuzzy-PID. Darüber hinaus ist

Machine Translated by Google

Angewandt im stabilen Betriebszustand ist der in diesem Dokument vorgeschlagene IADRC-Algorithmus in Bezug auf die Schnelligkeit 33,33 % besser als der IADRC in Ref. [34], 88,24 % besser als der traditionelle ADRC, 85,71 % besser als SMC und 94,29 % besser als Fuzzy-PID. In Bezug auf die Genauigkeit ist er 64,35 % besser als der IADRC in Ref. [34], 75,24 % besser als der traditionelle ADRC, 90,99 % besser als SMC und 59,19 % besser als Fuzzy-PID. Darüber hinaus belegen experimentelle Tests mit zusätzlichem gaußschen Rauschen und sich änderndem Motorwiderstand, dass der vorgeschlagene IADRC-Algorithmus eine gute Robustheit aufweist.

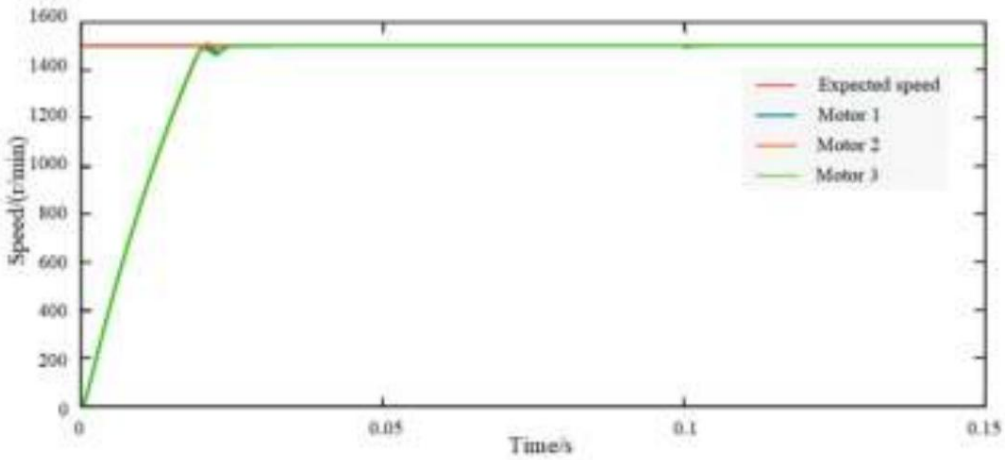


Abbildung 9. Motordrehzahlkurven mit unterschiedlichen Widerständen.

5. Schlussfolgerungen

In diesem Artikel wird **eine verfeinerte,** auf der Fal- **Funktion basierende IADRC-Strategie erläutert und im Vergleich zum herkömmlichen ADRC, IADRC in Ref. [34], SMC und Fuzzy-PID wird eine bessere Regelungsleistung für den PMSM erzielt. Simulationsergebnisse zeigen, dass der in diesem Artikel vorgeschlagene IADRC im Vergleich zu ADRC, IADRC in Ref. [34], SMC und Fuzzy-PID eine schnellere Reaktionsgeschwindigkeit aufweist. Die Zeit zum Erreichen und Stabilisieren der erwarteten Geschwindigkeit ist 68,24 % kürzer als bei ADRC, 40,00 % kürzer als bei IADRC in Ref. [34], 64,00 % kürzer als bei SMC und 37,21 % kürzer als bei Fuzzy-PID. Darüber hinaus weist der in diesem Artikel vorgeschlagene IADRC kein Überschwingen auf und die maximale Fehlerreduzierung beträgt 73,72 % im Vergleich zu ADRC und 49,25 % im Vergleich zu IADRC in Ref. [34], 92,62 % im Vergleich zu SMC und ÿ3,35 % im Vergleich zu Fuzzy-PID. Im Fall einer plötzlichen Störung, wenn die Geschwindigkeit des PMSM auf die erwartete Geschwindigkeit stabilisiert wird, ist die Zeit zum Erreichen und Stabilisieren der erwarteten Geschwindigkeit 88,24 % kürzer als bei ADRC, 33,33 % kürzer als bei IADRC in Ref. [34], 85,71 % kürzer als bei SMC und 94,29 % kürzer als bei Fuzzy-PID. Darüber hinaus ist der maximale Fehler des in diesem Dokument vorgeschlagenen IADRC 75,24 % niedriger als bei ADRC, 64,35 % niedriger als bei IADRC in Ref. [34], 90,99 % niedriger als bei SMC und 59,19 %** niedriger als bei Fuzzy-PID. Abgesehen davon wird die gute Robustheit auch durch **Hinzufügen von gaußschem Rauschen und Ändern des Motorwiderstands überprüft .**

In unserer zukünftigen Arbeit werden wir uns auf die Anwendung des vom IADRC vorgeschlagenen Algorithmus zur physischen Steuerung von PMSMs konzentrieren und den Algorithmus entsprechend der tatsächlichen Steuerungssituation weiter verbessern.

Beiträge der Autoren: Schreiben – Vorbereitung des Originalentwurfs, ZH; Schreiben – Überprüfung und Bearbeitung, YC; Ressourcen, SC; Betreuung, XZ, JX und SW. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Forschung wurde von der National Natural Science Foundation of China unter der Fördernummer 52205316 und der Zhejiang Provincial Natural Science Foundation of China unter der Fördernummer LTGN24E050002 finanziert. Und das APC wurde vom Wenzhou Major Science and Technology Innovation Project of China unter der Fördernummer ZN2022002 finanziert .

Machine Translated by Google

Datenverfügbarkeitserklärung: Die in der Studie vorgestellten Originalbeiträge sind im Artikel enthalten, weitere Anfragen können an den entsprechenden Autor gerichtet werden.

Danksagung: Die Autoren danken den Herausgebern und den anonymen Gutachtern für ihre aufschlussreichen Kommentare und Vorschläge.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte vorliegen.

Verweise

1. **Bolognani, S.; Tubiana, L.; Zigliotto, M. Erweiterte Kalman-Filterabstimmung in sensorlosen PMSM-Antrieben.** IEEE Trans. Ind. Appl. **2003,** 39, **1741–1747.** [CrossRef]
2. **Wallmark, O.; Harnefors, L.; Carlson, O. Steuerungsalgorithmen für einen fehlertoleranten PMSM-Antrieb.** IEEE Trans. Ind. Electron. **2007,** 54, **1973–1980.** [CrossRef]
3. **Gashti, A.; Akbarimajd, A. Entwicklung eines Anti-Windup-PI-Reglers für die LFC eines nichtlinearen Stromsystems in Kombination mit DSTS eines Kernkraftwerks und einer HGÜ-Verbindung.** Electr. Eng. **2020,** 102, **793–809.** [CrossRef]
4. **Errouissi, R.; Al-Durra, A.; Muyeen, SM Experimentelle Validierung eines neuartigen PI-Drehzahlreglers für Wechselstrommotorantriebe mit verbesserten Übergangseigenschaften.** IEEE Trans. Control Syst. Technol. **2018,** 26, **1414–1421.** [CrossRef]
5. **Zaky, M.; Khater, M.; Yasin, H.; Shokralla, S. Schätzungen von sehr niedriger Geschwindigkeit und Nullgeschwindigkeit von sensorlosen Induktionsmotorantrieben.** Electr. Power Syst. Res. **2010,** 80, **143–151.** [CrossRef]
6. **Shi, R. Verbesserung des prädiktiven Regelungsalgorithmus basierend auf Fuzzy-PID mit fraktionaler Ordnung.** J. Intell. Syst. **2023,** 32, **876.** [CrossRef]
7. **Tan, W.; Han, W.; Xu, J. State-Space PID: Ein fehlendes Bindeglied zwischen klassischer und moderner Regelung.** IEEE Access **2022,** 10, **116540–116553.** [CrossRef]
8. **Atalik, G.; Senturk, S. Intuitionistische Fuzzy-Kontrollkarten basierend auf der intuitionistischen Fuzzy-Ranking-Methode für TIFNs.** Soft Comput. **2022,** 26, **11403–11407.** [CrossRef]
9. **Tu, WW; Dong, JX Robuste Gleitmodussteuerung für eine Klasse nichtlinearer Systeme durch Dual-Layer-Gleitmodusschema.** J. Frankl. Inst. Eng. Appl. Math. **2023,** 360, **10227–10250.** [CrossRef]
10. **Zhang, GC; Li, XF; Xia, YQ; He, S. Adaptive Gleitmodussteuerung für ein 2D-nichtlineares Fornasini-Marchesini-Modell, das Quantisierung und Paketausfällen unterliegt.** Int. J. Syst. Sci. **2021,** 52, **3001–3012.** [CrossRef]
11. **Zhao, J.; Zhao, T.; Liu, N. Aktive Störungsunterdrückungssteuerung fraktionaler Ordnung mit Fuzzy-Selbstoptimierung für präzise stabilisierte Plattformen.** Entropy **2022,** 24, **1681.** [CrossRef] [PubMed]
12. **Xiang, CD; Petersen, IR; Dong, DY Kohärente robuste Hÿ- Kontrolle unsicherer linearer Quantensysteme mit direkter und indirekter Kupplungen.** J. Frankl. Inst. Eng. Appl. Math. **2023,** 360, **13845–13869.** [CrossRef]
13. **Han, J. Von PID zur aktiven Störungsunterdrückungsregelung.** Trans. Ind. Electron. **2009,** 56, **900–906.** [CrossRef]
14. **Liu, JJ; Chen, H.; Wang, L.; Yu, MY Aktive Störungsunterdrückungskontrolle für ein verbessertes Tiefenmodell von AUV.** Appl. Mech. Mater. **2014,** 687, **157–162.** [CrossRef]
15. **Meng, Y.; Liu, B.; Wang, L. Geschwindigkeitsregelung von PMSM basierend auf einem optimierten ADRC-Controller.** Math. Probl. Eng. **2019,** 2019, **18.** [Querverweis]
16. **Xiong, S.; Xie, H.; Song, K.; Zhang, G. Eine Geschwindigkeitsverfolgungsmethode für autonomes Fahren über ADRC mit erweitertem Zustandsbeobachter.** Appl. Sci. **2019, 9, 3339.** [CrossRef]
17. **Wan, J.; Liu, H.; Yuan, J.; Shen, Y.; Zhang, H.; Wang, H.; Zheng, Y. Bewegungssteuerung eines autonomen Unterwasserfahrzeugs basierend auf Fraktionale Analysis: Aktive Störungsunterdrückung.** J. Mar. Sci. Eng. **2021, 9, 1306.** [CrossRef]
18. **Yang, Z.; Qian, C.; Sun, X.; Wang, K. Verbesserte lineare ADRC-Strategie für die sensorlose Steuerung von IPMSM unter Berücksichtigung von Cross-Kopplungsfaktoren.** IEEE Trans. Transp. Electrif. **2023, 9, 4437–4446.** [CrossRef]
19. **Zuo, YF; Zhang, J.; Liu, C.; Zhang, T. Integriertes Design für Permanentmagnet-Synchronmotor-Servosysteme basierend auf aktiv Störungsunterdrückungskontrolle.** Trans. China Electrotech. Soc. **2016,** 31, **51–58.** [CrossRef]
20. **Fang, S.; Meng, J.; Meng, Y.; Wang, Y.; Huang, D. Zeitdiskrete aktive Störunterdrückungsstromregelung des PM-Motors bei Niedrige Geschwindigkeit im resonanten Gleitmodus.** IEEE Trans. Transp. Electrif. **2023, 9, 4783–4794.** [CrossRef]
21. **Li, Z.; Zhang, Z.; Wang, J.; Wang, S.; Chen, X.; Sun, H. ADRC-Steuerungssystem von PMLSM basierend auf einem neuartigen nicht-singulären Terminal- Sliding-Mode-Beobachter.** Energies **2022,** 15, **3720.** [CrossRef]
22. **Ge, X.; Sun, K. Optimale Steuerung eines Raumfahrzeugs mit ausfahrbaren Solarzellen unter Verwendung eines Partikelschwarmoptimierungsalgorithmus.** Sci. China Technol. Sci. **2011,** 54, **1107–1112.** [CrossRef]
23. **Lamraoui, HC; Qidan, Z. Geschwindigkeitsverfolgungssteuerung eines mobilen Einradroboters basierend auf LADRC. In Proceedings der 2017 3. IEEE International Conference on Control Science and Systems Engineering (ICCSSE), Peking, China, 17.–19. August 2017; S. 200–204.** [CrossRef]
24. **Partovibakhsh, M.; Liu, G. Schlupfverhältnisschätzung und Steuerung von mobilen Robotern auf Rädern auf unterschiedlichem Gelände. In Proceedings der 2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), Shenyang, China, 8.–12. Juni 2015; S. 566–571.** [CrossRef]
25. **Ding, L.; Gao, H.-B.; Deng, Z.-Q.; Li, Z.; Xia, K.-R.; Duan, G.-R. Pfadfolgesteuerung von Planetenerkundungsrobotern auf Rädern Fortbewegung auf verformbarem unebenem Gelände.** Sci. World J. **2014,** 2014, **793526.** [CrossRef] [PubMed]
26. **Feng, H.; Guo, B.-Z. Aktive Störungsunterdrückungskontrolle: Alte und neue Ergebnisse.** Annu. Rev. Control **2017,** 44, **238–248.** [CrossRef]

Machine Translated by Google

27. Lv, C.; Wang, B.; Chen, J.; Zhang, R.; Dong, H.; Wan, S. Forschung zu einer Methode zur Unterdrückung von Drehmomentwelligkeiten durch Fuzzy- Steuerung mit aktiver Störungsunterdrückung für einen Permanentmagnet-Synchronmotor. Electronics **2024,** 13, **1280. [CrossRef]**

28. Wang, X.; Zhu, H. Aktive Störungsunterdrückungssteuerung eines lagerlosen Permanentmagnet-Synchronmotors basierend auf einem genetischen Algorithmus und einer Methode zur dynamischen Anpassung der Parameter neuronaler Netze. Electronics **2023,** 12, **1455. [CrossRef]**

29. Fang, Q.; Zhou, Y.; Ma, S.; Zhang, C.; Wang, Y.; Huangfu, H. Elektromechanische Aktuator-Servosteuerungstechnologie basierend auf Aktive Störungsunterdrückungsregelung. Electronics **2023,** 12, **1934. [CrossRef]**

30. Sun, C.; Liu, M.; Liu, C.; Feng, X.; Wu, H. Eine industrielle Quadrotor-UAV-Steuerungsmethode basierend auf Fuzzy Adaptive Linear Active Störungsunterdrückungsregelung. Electronics **2021,** 10, **376. [CrossRef]**

31. Štumberger, B.; Štumberger, G.; Hadžiselimoviˊc, M.; Hamler, A.; Goriÿcan, V.; Jesenik, M.; Trlep, M. Vergleich der Drehmomentfähigkeit von dreiphasigen Permanentmagnet-Synchronmotoren mit unterschiedlicher Permanentmagnetanordnung. J. Magn. Magn. Mater. **2007, 316 , e261–e264.**

32. Zhang, XG; Fang, SJ; Zhang, H. Prädiktive Stromfehlerkompensation basierend auf einem starken, robusten Modell zur prädiktiven Steuerung für PMSM-Antriebssysteme. IEEE Trans. Ind. Electron. **2024, 1–10. [CrossRef]**

33. Vitor, ALO; Scalassara, PR; Goedtel, A.; Endo, W. Muster basierend auf Clarke- und Park-Transformationen von Wavelet-Koeffizienten für Klassifizierung von Fehlern an elektrischen Maschinen. J. Control Autom. Electr. Syst. **2023,** 34, **230–245. [CrossRef]**

34. Xu, P. Forschung zum PMSM-Geschwindigkeitskontrollsystem für Bergbaufahrzeuge auf Basis von Fuzzy-Selbstimmunität. Masterarbeit, Anhui Technische Universität, Anhui, China, 2022.

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Untersuchung und Abmilderung des Kompromisses zwischen Leistung und Fairness durch Protected-Category-Sampling

Gideon Popoola und John Sheppard *

Gianforte School of Computing, Montana State University, Bozeman, MT 59717, USA;
gideon.popoola@student.montana.edu * Korrespondenz: john.sheppard@montana.edu

Zusammenfassung: Algorithmen für maschinelles Lernen sind bei alltäglichen Entscheidungen mittlerweile weit verbreitet und Entscheidungshilfesysteme sind in unserem Alltag allgegenwärtig. Daher hat die Forschung zur Vermeidung und Eindämmung potenzieller Verzerrungen und Ungerechtigkeiten bei den von diesen Algorithmen getroffenen Vorhersagen in den letzten Jahren zugenommen. Die meisten Forschungsarbeiten zu Fairness und Verzerrungsminderung beim maschinellen Lernen behandeln oft jede geschützte Variable separat, aber in Wirklichkeit ist es möglich, dass eine Person mehreren geschützten Kategorien angehört . Daher wurde in dieser Arbeit die Kombination einer Reihe geschützter Variablen und die Generierung neuer Spalten untersucht, die diese geschützten Variablen in Diese neuen Unterkategorien sind häufig extrem unausgewogen, sodass die Risikominderung als Problem der unausgewogenen Klassifizierung angegangen wurde. Konkret wurden vier neue benutzerdefinierte Stichprobenverfahren entwickelt und untersucht, um diese neuen Unterkategorien zu beproben. Diese neuen Stichprobenverfahren werden als „ Protected-category Oversampling“, „Protected-category Proportional Sampling“, „Protected-category Synthetic Minority Oversampling Technique“ (PC-SMOTE) und „Protected-category Adaptive Synthetic Sampling“ (PC-ADASYN) bezeichnet. Diese Stichprobenverfahren modifizieren das bestehende Stichprobenverfahren, indem sie sich bei der Stichprobennahme auf die neuen Unterkategorien statt auf die Klassenbezeichnung konzentrieren. Die Auswirkung dieser Stichprobenstrategien wurde dann anhand der klassischen Leistung und Fairness in Klassifizierungseinstellungen bewert Die Klassifizierungsleistung wurde anhand von Genauigkeit und F1 gemessen, basierend auf dem Training univariater Entscheidungsbäume, und die Fairness wurde anhand von ausgeglichenen ungeraden Differenzen und statistischer Parität gemessen. Um die Auswirkung von Fairness gegenüber Leistung zu bewerten, wurden diese Maßnahmen anhand der Tiefe des Entscheidungsbaums bewertet. Die Ergebnisse zeigen, dass die vorgeschlagenen Methoden in der Lage waren, optimale Punkte zu bestimmen, wodurch die Fairness erhöht wurde, ohne die Leistung zu verringern, wodurch jeder potenzielle Kompromiss zwischen Leistung und Fairness gemildert wurde.

Schlüsselwörter: Fairness, geschützte Kategorien, maschinelles Lernen, Sampling

check for updates

Zitat: Popoola, G.; Sheppard, J. Untersuchung und Eindämmung der Kompromiss zwischen Leistung und Fairness durch Stichprobennahme geschützter Kategorien. Electronics **2024,** 13, **3024. https:// doi.org/ 10.3390/electronics13153024**

Akademische Herausgeber: Niusha Shafiabady und Jianlong Zhou

Empfangen: 11. Juni 2024
Überarbeitet: 23. Juli 2024
Akzeptiert: 24. Juli 2024
Veröffentlicht: 31. Juli 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.
Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

Da Algorithmen des maschinellen Lernens (ML) Entscheidungsfindungs- und Entscheidungsunterstützungssysteme zunehmend dominieren, wirft ihr weitverbreiteter Einsatz in verschiedenen Sektoren dringende Fragen hinsichtlich der Fairness und Transparenz ihrer Vorhersagen auf **[1]. Das Potenzial dieser Algorithmen, bestehende gesellschaftliche Vorurteile aufrechtzuerhalten oder zu verschärfen, hat eine bedeutende Forschungsarbeit ausgelöst, die sich mit der Untersuchung und Eindämmung algorithmischer Ungerechtigkeit befasst. Dies ist von entscheidender Bedeutung, da die von diesen Algorithmen beeinflussten Entscheidungen tiefgreifende Auswirkungen auf Einzelpersonen haben und sich auf Ergebnisse in Bereichen von Finanzen und Beschäftigung bis hin zu Strafjustiz und Gesundheitswesen** a

Die Quelle von Ungerechtigkeit und Verzerrung im ML ist vielschichtig [3]. Insbesondere ist es möglich, dass Ungerechtigkeit direkt aus den ML-Algorithmen selbst entsteht, aufgrund einer möglichen Fehlausrichtung der zugrunde liegenden induktiven Verzerrung **der Algorithmen gegenüber dem Zielkonzept** und der Datenverteilung. Dies wird als algorithmische Verzerrung **bezeichnet . Ein alternatives Problem liegt in einer potenziellen Verzerrung in den Daten, die zum Trainieren der Modelle verwendet werden, wo als direkte Folge der typischen Annahme „unabhängig und identisch verteilt“ (I**

Machine Translated by Google

Methoden besteht das Ergebnis des Lernens darin, die Verzerrung in den Vorhersagen so zu propagieren, dass sie der Verzerrung in den zugrunde liegenden Daten selbst entsprechen. Diese letztere Situation bildet den Schwerpunkt unserer Arbeit hier.

1.1. Voreingenommenheit und Ungerechtigkeit beim maschinellen Lernen

Unter Voreingenommenheit versteht man die nachteilige, unfaire oder ungleiche Behandlung einer Einzelperson oder Gruppe aufgrund bestimmter Merkmale, die oft als sensible oder geschützte Merkmale bezeichnet werden [4]. Beispiele für diese geschützten Merkmale sind Alter, Rasse, Behinderung, Geschlecht und Gender [5]. Voreingenommenheit im ML kann grob in ungleiche Behandlung (direkte Ungerechtigkeit) und wirksame Behandlung (indirekte Ungerechtigkeit) unterteilt werden **[6]. Direkte Ungerechtigkeit liegt vor, wenn geschützte Merkmale explizit bei der Entscheidungsfindung verwendet werden. Indirekte Ungerechtigkeit kommt heute immer häufiger vor. Bei dieser Art der Ungerechtigkeit werden geschützte Attribute nicht explizit verwendet. Sie liegt stattdessen vor, wenn das Vertrauen auf Variablen, die mit diesen Attributen verbunden sind, zu erheblich unterschiedlichen Ergebnissen für die geschützten Gruppen führt. Diese anderen Variablen werden als Proxy-Merkmale bezeichnet. Beispiele für Voreingenommenheit in der realen Welt sind die historische Praxis des „Redlining“ in den USA, bei der Bewohnern von Postleitzahlengebieten mit überwiegend Minderheitenanteil Hypotheken verweigert wurden, geschlechtsspezifische Voreingenommenheit im Einstellungsverfahre**

Obwohl diese Entscheidungshilfe-Tools dabei helfen, den Entscheidungsprozess zu automatisieren, können sie zu einer unfairen Behandlung von Einzelpersonen oder Gruppen führen, sowohl direkt als auch indirekt [8]. Ungerechtigkeiten können in mehreren Bereichen der Modellierung auftreten, beispielsweise im Trainingsdatensatz . Dies kann passieren, wenn der Trainingsdatensatz keine faire Darstellung der geschützten Kategorien bietet, sodass die „Grundwahrheit“ schwer zu bestimmen ist. Betrachten Sie beispielsweise einen Datensatz eines Unternehmens, in dem eine bestimmte Gruppe in der Vergangenheit diskriminiert wurde. Nehmen wir beispielsweise an, weibliche Mitarbeiter in diesem Unternehmen wurden nicht so befördert wie ihre männlichen Kollegen, die im Gegensatz dazu Karriere machen konnten, obwohl beide Gruppen die gleiche Leistung erbrachten. In dieser Situation ist der wahre Wert der Beiträge weiblicher Mitarbeiter – die Grundwahrheit – nicht sichtbar. Infolgedessen erkennt und berücksichtigt ein ML-Algorithmus, der mit diesen Daten trainiert wird, diese Verzerrung wahrscheinlich und verewigt so bestehende Vorurteile. Dies könnte dazu führen, dass der Algorithmus diskriminierende Entscheidungen trifft, beispielsweise männliche Kandidaten häufiger für eine Einstellung oder Beförderung empfiehlt als gleich oder besser qualifizierte weibliche Kandidaten.

Ein weiterer Bereich, in dem Ungerechtigkeit auftreten kann, ist der ML-Algorithmus selbst [9]. ML-Algorithmen können immer noch diskriminierende Entscheidungen treffen, selbst wenn sie mit einem unvoreingenommenen Datensatz trainiert werden, in dem die „Grundwahrheit“ genau dargestellt ist. Diese Situation entsteht, wenn die Fehler des Systems Einzelpersonen einer bestimmten Gruppe oder Minderheit überproportio Betrachten wir beispielsweise einen Algorithmus zur Brustkrebserkennung, der bei schwarzen Personen deutlich mehr Falsch-Negativ-Ergebnisse liefert als bei weißen. Das heißt, er erkennt Brustkrebs bei schwarzen Patienten häufiger nicht als bei weißen. Würde dieser Algorithmus zur Abgabe von Behandlungsempfehlungen verwendet, würde er fälschlicherweise mehr schwarzen als weißen Personen von einer Behandlung abraten, was zu rassistischen Ungleichheiten bei den Behandlungsergebnissen führen würde. Dies unterstreicht die dringende Notwendigkeit, sicherzustellen, dass die Algorithmen hinsichtlich ihrer Trainingsdaten und der Auswirkungen ihrer Fehler auf unterschiedliche Populationen für alle Gruppen gleich gut funktionieren. Ergebnisse aus der bisherigen Literatur berichten von mehreren Fällen, in denen Algorithmen zu unfairer Behandlung führten, z. B. Redlining und Racial Profiling [10], Diskriminierung bei Hypotheken [11], Einstellung und Personalauswahl [12].

Während erhebliche Anstrengungen unternommen wurden, um die Verzerrung bei ML-Vorhersagen zu beheben [13,14], konzentrierte sich ein Großteil der bestehenden Forschung auf die Abschwächung der Verzerrung für einzelne geschützte Attribute isoliert [15]. Beispielsweise können die meisten bestehenden Ansätze in einem Datensatz mit zwei geschützten Attributen, Rasse und Geschlecht, entweder ein faires Modell mit Rasse oder ein faires Modell mit Geschlecht lernen, aber kein faires Modell mit sowohl Rasse als auch Geschlecht [7]. Identitäten in der realen Welt sind jedoch nicht singulär; sie sind komplex und vielschichtig, wobei Einzelpersonen oft mehreren geschützten Gruppen gleichzeitig angehören [16]. Beispielsweise kann eine Person aufgrund mehrerer geschützter Attribute wie Alter, Rasse und Geschlecht gleichzeitig diskriminiert werden. Diese Intersektionalität kann zu zusammengesetzten Formen der Verzerrung führen und

Machine Translated by Google

Diskriminierung, die durch Fairnessmaßnahmen mit einer einzigen Variable nicht ausreichend angegangen werden. Daher ist es wichtig, Methoden zu entwickeln, die die mehrdimensionale Natur persönlicher Identitäten ganzheitlich berücksichtigen. Dieses Projekt versucht, diese Lücke zu schließen, indem es Kombinationen geschützter Kategorien in Betracht zieht und diese geschützten Kategorien so zu umfassenden Mehrkategoriengruppen zusammenfasst. Es zielt darauf ab, die vielschichtigen Komplexitäten der Voreingenommenheit mithilfe neuartiger Stichprobenmethoden für geschützte Kategorien effektiver anzugehen und so die Vielschichtigkeit persönlicher Identitäten und potenzielle

Die in diesem Artikel vorgestellte Arbeit ist motiviert durch das Problem der Verwendung von ML-Algorithmen für die Entscheidungsfindung in sozial sensiblen Bereichen wie Kreditbewertung, Einstellung oder Hypothekenbewertung , wobei mit dieser Situation gearbeitet wird, in der eine Person mehreren geschützten Kategorien angehören kann . Bei einem gekennzeichneten Trainingsdatensatz mit zwei oder mehr geschützten Merkmalen kombiniert die vorgeschlagene Methode diese geschützten Attribute und teilt sie dann in neue Multikategorien auf. Diese neuen Kategorien sind wahrscheinlich extrem unausgewogen und müssen ausgeglichen werden, um die Fairness der Vorhersage unserer ML-Algorithmen zu verbessern. Beliebte Stichprobenmethoden wie Oversampling [17], Synthetic Minority Oversampling Technique (SMOTE) [18], Adaptive Synthetic Sampling (ADASYN) [19] usw. ziehen Stichproben über Klassenbezeichnungen hinweg , was nicht mit dem Ziel unserer Forschung der Stichprobennahme über die neuen Multikategorien hinweg übereinstimmt. Daher wird eine neue Klasse von Modifikationen dieser Stichprobenmethoden vorgeschlagen, die Stichproben über die neue Kategorie statt über Klassenbezeichnungen hinweg ziehen können. Diese neue Klasse modifizierter Stichprobennahme wird als geschützte Kategoriestichprobennahme **bezeichnet .** Die daraus resultierenden vorgeschlagenen **Stichprobenverfahren für geschützte Kategorien** werden verwendet, um die neuen Kategorien zu prüfen und auszugleichen, bevor **die Klassifizierung durchgeführt wird. Das Neue an dieser Arbeit ist zweifacher Natur. Erstens kombiniert der vorgeschlagene Ansatz die geschützten Kategorien, um neue Multikategorien zu bilden, die nachahmen, wie die Identität des Menschen in der realen Welt aussieht. Zweitens werden vorhandene Stichprobenverfahren so modifiziert, dass sie mit der Stichprobenziehung**

Nur zu Demonstrationszwecken wurde ein univariater Entscheidungsbaum als Klassifizierungsalgorithmus gewählt. Ziel ist es, die Auswirkungen der verschiedenen Stichprobenverfahren auf die Leistung zu demonstrieren , wobei davon ausgegangen wird, dass sich unabhängig von der **zugrunde liegenden Lernmethode ähnliche Trends zeigen .** Das vorgeschlagene Stichprobenverfahren wurde mit der **Baseline (nicht abgetastete Daten) verglichen, wobei Genauigkeit und F1 als Leistungskennzahlen für die Klassifizierung sowie ausgeglichene Quotenunterschiede und statistische Parität als Fairnesskennzahlen verwendet wurden. Außerdem wurden mehrere Analysen durchgeführt, um zu zeigen, wie sich die maximale Tiefe des Entscheidungsbaums auf Genauigkeit und Fairness auswirkt.**

1.2. Forschungsfrage

Ausgehend von der empirischen Beobachtung, dass manchmal ein Kompromiss zwischen Fairness und ML-Leistung besteht [20], versucht diese Forschung, mehrere Fragen zu beantworten, beispielsweise wie dieser Kompromiss gemildert werden könnte. Insbesondere wollen wir beantworten, ob die Methode der geschützten Kategorie-Stichprobennahme zur Bewältigung der Fairness diesen Kompromiss abmildern kann. Darüber hinaus können wir einen methodischen Rahmen entwickeln, der Verzerrungen über diese kombinierten geschützten Variablen hinweg effektiv abmildert, ohne die Vorhersagegenauigkeit von ML-Modellen zu beeinträchtigen? Schließlich planen wir, die Frage zu beantworten, wie sich die Tiefe eines Entscheidungsbaums sowohl auf die Genauigkeits- als auch auf die Fairnessmetriken auswirkt, und so die Beziehung zwischen dem Anpassungsgrad (Unteranpassung durch Überanpassung) und der

1.3. Hypothese

Wir gehen von der Hypothese aus, dass wir durch den Einsatz ausgefeilter geschützter Kategorienstichprobentechniken, die für diese neu formulierten Mehrkategoriengruppen entwickelt wurden, die Modellfairness in Bezug auf ausgeglichene Quotenunterschiede deutlich steigern können, ohne die Klassifizierungsleistung in Bezug auf Genauigkeit und F1 zu verringern. Darüber hinaus untersuchen wir das empfindliche Gleichgewicht zwischen Fairness und Genauigkeit und gehen davon aus, dass es möglich ist, strategische Punkte zu identifizieren, an denen die Fairness ohne nachteilige Auswirkungen auf die Leistung maxir Diese Forschung stellt bestehende Behauptungen über die Existenz von Kompromissen in Fairness und ML in Frage

Machine Translated by Google

Vorhersage. Es bereitet den Boden für zukünftige Untersuchungen der mehrdimensionalen Natur von Identität und Diskriminierung in automatisierten Entscheidungssystemen.

1.4. Beiträge

Das allgemeine Problem der Fairness beim maschinellen Lernen ist insofern bedeutsam, als die heutige Verbreitung von KI- und ML-Systemen erhebliche Auswirkungen auf das Leben und den Lebensunterhalt der Menschen Obwohl die Aufmerksamkeit für faires ML deutlich zugenommen hat, besteht weiterhin Bedarf an Methoden, um faires ML zu fördern, ohne die ML-Leistung negativ zu beeinflussen. Basierend auf einer gründlichen Literaturrecherche und dem oben genannten Bedarf an dieser Art von Arbeit leisten die hier vorgestellten Methoden die folgenden Beiträge: 1. Die weit verbreitete

Annahme, dass es beim maschinellen Lernen einen inhärenten Kompromiss zwischen Fairness und Leistung (d. h. Genauigkeit) gibt, wird durch Beweise in Frage gestellt, die diese Infragestellung untermauern. Insbesondere deuten die Ergebnisse in diesem Dokument darauf hin, dass ein solcher Kompromiss gemildert werden kann, was darauf hindeutet, dass jeder Kompromiss höchstwahrscheinlich mit der Art und Weise zusammenhängt, wie die Daten verwaltet werden.

2. Es werden vier neuartige Vorverarbeitungsmethoden für die Stichprobenentnahme von Daten vorgestellt, die auf der Anwendung einer Multikategorie-Stichprobenstrategie unter Verwendung von in geschützten Kategorien erfassten Daten basieren. Die Methoden gehen von der Annahme und der entsprechenden Hypothese aus, dass das Ausbalancieren der Daten auf der Grundlage dieser Multikategorieeigenschaften die Fairness erhöhen kann, ohne die Leistung des maschinellen Lernmodells zu be

3. Die experimentellen Ergebnisse werden anhand von drei Datensätzen präsentiert, die in der Fair-ML-Community ausführlich untersucht wurden. Die Experimente umfassen Vergleiche mit herkömmlichen Trainingsmethoden ohne Resampling, um die relativen Auswirkungen der vorgeschlagenen Methoden zu demonstrieren. Die Ergebnisse zeigen, dass zwei der vorgeschlagenen Methoden, Protected- Category Synthetic Minority Oversampling Technique (PC-SMOTE) und Protected- Category Adaptive Synthetic Sampling (PC-ADASYN), besonders effektiv bei der Verbesserung von Fairness und Leistung sind.

4. Eine detaillierte Analyse der möglichen Auswirkungen von Unter- und Überanpassung auf die Fairness wird durch die Untersuchung verschiedener Ebenen in einem Entscheidungsbaummodell mit und ohne Verwendung der vorgeschlagenen Stichprobenmethoden präsentiert. Die Ergebnisse zeigen, dass die vorgeschlagenen Methoden in der Lage sind, eine ideale Ebene des Baums zu identifizieren, auf der sowohl Fairness als auch Genauigkeit maximiert werden.

Als Ergebnis der oben genannten Beiträge stellt diese Arbeit einen bedeutenden Fortschritt bei der Lösung von Fairnessbedenken im maschinellen Lernen dar. Eine wichtige Erkenntnis aus den hier vorgestellten Methoden und Ergebnissen ist, dass Fairness ohne Beeinträchtigung der Modellleistung erreicht werden kann.

1.5. Organisation

Dieses Dokument ist wie folgt organisiert. Abschnitt 2 enthält eine detaillierte Erklärung von Fairness und eine Diskussion mehrerer technischer Fairnessmetriken. Anschließend wird in Abschnitt 3 frühere Literatur zu Strategien zur Minderung von Verzerrungen beschrieben. Abschnitt 4 beschreibt unsere vorgeschlagenen Sampling-Techniken, Datensätze und unseren Ansatz zur Hyperparameter-Optimierung. Abschnitt 5 präsentiert die Ergebnisse mehrerer Experimente zusammen mit statistischen Hypothesentests als Mittel zur Validierung dieser Ergebnisse. Abschnitt 6 diskutiert die experimentellen Ergebnisse und analysiert die Leistung jedes Algorithmus bei jedem Datensatz und jeder Metrik. Außerdem werden weitere Ergebnisse zum Einfluss der Baumtiefe auf Fairness und Genauigkeit präsentiert . Abschnitt 7 präsentiert die Grenzen dieser Arbeit und entsprechende Richtungen für künftige Arbeiten und Abschnitt 8 präsentiert eine Reihe von Schlussfolgerungen.

2. Hintergrund

Diese Studie untersucht die Fairness bei der Vorhersage eines Ergebnisses y ÿ **Y aus einer Reihe turen x ÿ X ÿ** Rd **und einige zusätzliche geschützte Attribute s ÿ S ÿ** Rp **Geschlecht und , von Merkmalen wie Rass Geschlecht. Bei der Kreditvorhersage beispielsweise** steht **x für die finanzielle Vergangenheit eines Antragstellers, s für die von ihm selbst angegebene Rasse und sein Geschlecht und** y **für die Frage, ob sein Kredit ge**

Machine Translated by Google

oder verweigert. Ein Vorhersagemodell gilt als fair, wenn seine Fehler gleichmäßig auf geschützte Gruppen wie unterschiedliche Rassen oder Geschlechter verteilt sind. Die Klassenvorhersagen aus den Trainingsdaten D werden als Yˆ D := h(x, s) **für ein** h : **X × S ÿ Y aus einer Klasse H bezeichnet. Die geschützten Attribute s ÿ S in unserer Studie werden als binär angenommen, wobei ein spezieller Wert** n **die nicht privilegierte Gruppe bezeichnet . Zum Beispiel** könnte **S die Rasse und** n **„nicht weiß“ sein; daher ist** die binäre **Natur von S** {w, **n}, wobei** w **weiße Bewerber darstellt, die die privilegierte Gruppe sind, und** n **nicht-weiße Bewerber darstellt, die die nicht privilegierte Gruppe sind. Die Definition lässt sich weiter auf nicht-binä**

X , Datensatz D, einen Merkmalssatz und einen geschützten Attributsatz S mit dem Domänenwert {w, **n}. Die Unterscheidung in beschrifteten Datensätzen kann wie folgt definiert werden: Es gibt einen Die Unterscheidung in D in Bezug auf die Gruppe S** = n , **bezeichnet als** diss=n(D), **ist definiert als**

$$\text{diss=n(D)} = \frac{|\{x \ddot{y} D : x(s) = w, h(x) = +\}|}{|\{x \ddot{y} D : x(s) = w\}|} \quad \frac{|\{x \ddot{y} D : x(s) = n, h(x) = +\}|}{\{x \ddot{y} D : x(s) = n\}|}$$

Die obige Definition kann als die Differenz der Wahrscheinlichkeit übersetzt werden, dass ein Bewerber für jeden geschützten Attributbereich {w, **n} zur positiven Klasse gehört. Unsere Studie erweitert die obige Definition, indem sie den Datensatz D betrachtet, der zwei oder mehr geschützte Attribute enthält.**

Es werden zwei gängige Fairnessmaße verwendet. Das erste ist die Equalized Odd Difference **(EOD), die misst, wie diskriminativ oder fair unsere Vorhersage ist. EOD besagt, dass ein binärer Klassifikator yˆ fair ist, wenn seine Falsch-Negativ-Rate (FNR) und seine Wahr-Positiv-Rate (TPR) über den gesamten Bereich von S** gleich sind **[21]. FPR und TPR in Bezug auf das geschützte Attribut s ÿ S mit dem Wert** n **können definiert werden als**

$$TPRn(\hat{y}) = P(\hat{y} = \mathbf{1}|\mathbf{y} = \mathbf{1}, S = n)$$

$$FPRn(\hat{y}) = P(\hat{y} = \mathbf{1}|\mathbf{y} = \mathbf{0}, S = n)$$

EOD wird dann mathematisch als die Differenz zwischen TPR und FPR über verschiedene Gruppen in einem geschützten Attribut definiert. Das heißt,

$$EOD = TPRn(\hat{y}) \ddot{y} FPRn(\hat{y}).$$

Ein fairer Klassifikator hat einen EOD von 0, während ein unfairer Klassifikator einen EOD von 1 hat. Obwohl es in der Praxis fast unmöglich ist, einen vollständig fairen Klassifikator zu erreichen, zielt diese Forschung darauf ab, den EOD zu verbessern, ohne die Genauigkeit zu verringern. Für den EOD gilt:

$$FPRn(\hat{y}) = P(\hat{y} = \mathbf{1}|\mathbf{y} = \mathbf{0}, S = n) = TPRn(\hat{y}) = P(\hat{y} = \mathbf{1}|\mathbf{y} = \mathbf{1}, S = n)$$

Und

$$FPRw(\hat{y}) = P(\hat{y} = \mathbf{1}|\mathbf{y} = \mathbf{0}, S = w) = TPRw(\hat{y}) = P(\hat{y} = \mathbf{1}|\mathbf{y} = \mathbf{0}, S = w)$$

Um die obige EOD-Definition auf unsere Multikategorie auszudehnen, wird der EOD für jede Spalte berechnet und dann der Makrodurchschnitt des EOD als endgültiger EOD dargestellt. Die zweite Metrik, die zur Messung der Fairness bei der ML-Vorhersage verwendet wird, ist die statistische Parität **(SP). SP definiert Fairness als gleiche Wahrscheinlichkeit, als positiv klassifiziert zu werden [22]. Dies kann so interpretiert werden, dass** jede Gruppe in einem geschützten Attribut die gleiche Wahrscheinlichkeit hat, mit einem **positiven Ergebnis klassifiziert zu werden .**

$$P(\hat{y} = \mathbf{1}|\mathbf{S} = w) = P(\hat{y} = \mathbf{1}|\mathbf{S} = n)$$

3. Literaturübersicht

ML-Algorithmen, die zunehmend für die Entscheidungsfindung in kritischen Anwendungen wie Rückfall, Kreditscoring, Kreditentscheidungen usw. eingesetzt werden, könnten zunächst als fair und frei von inhärenter Voreingenommenheit angesehen werden. In Wirklichkeit können sie jedoch jede Voreingenommenheit oder Diskriminierung übernehmen, die in den Daten vorhanden ist, mit denen sie trainiert werden, wie Burt [23] feststellt. Darüber hinaus reicht das bloße Entfernen geschützter Variablen aus dem Datensatz nicht aus, um indirekte D

Machine Translated by Google

könnte es sogar verbergen. Diese Erkenntnis hat den Bedarf an fortschrittlicheren Werkzeugen erhöht, wodurch die Entdeckung und Prävention von Diskriminierung zu einem wichtigen Forschungsgebiet geworden ist, wie [24–27] hervorhebt .

Voreingenommenheit in ML ist ein schnell wachsendes Thema in der Forschungsgemeinschaft für maschinelles Lernen. Voreingenommenheit in einem ML-Modell kann zu einer unfairen Behandlung von Menschen führen, die bestimmten geschützten Gruppen angehören. In letzter Zeit legen Branchenführer immer mehr Wert auf Voreingenommenheit in ML-Modellen und -Software. Das Institute for Electrical and Electronics Engineers (IEEE) [28], Microsoft [29] und die Europäische Union [30] haben kürzlich Grundsätze für faires KI-Verhalten veröffentlicht. Diese Organisationen haben erklärt, dass ML-Modelle in realen Anwendungen fair sein müssen. Strategien zur Minderung von Voreingenommenheit beinhalten die Änderung eines oder mehrerer der folgenden Elemente, um sicherzustellen, dass die vom ML-Algorithmus getroffenen Vorhersagen weniger voreingenommen sind: (a) die Trainingsdaten, (b) der ML-Algorithmus und (c) die daraus resultierenden Vorhersagen selbst. Diese werden jeweils in Vorverarbeitungsansätze [31], Inverarbe

Erstens können die Trainingsdaten vor dem Training des Modells vorverarbeitet werden, um Ungerechtigkeit oder Verzerrung zu verringern. Kamiran und Calders [6] schlagen vor, die Daten zu stichprobenartig zu erfassen oder neu zu gewichten, um Diskriminierung zu neutralisieren. Mit diesem Ansatz kann die Darstellung oder Bedeutung bestimmter Datenpunkte angepasst werden, um eine Klasse gegenüber einer anderen zu bevorzugen (oder zu verringern). Eine andere Methode besteht darin, einzelne Datensätze direkt zu ändern, um Diskriminierung zu verringern, wie in [33] untersucht wird. Bei diesem Ansatz werden beispielsweise Werte in einem Datensatz geändert, um erkennbare Verzerrungen gegenüber bestimmten Gruppen zu verringern. Darüber hinaus wird das von Sondeck et al. [34] eingeführte Konzept der t-Nähe in der Arbeit von [35] auf die Diskriminierungskontrolle angewendet . Die Verwendung der t-Nähe stellt sicher, dass die Verteilung sensibler Attribute in einer bestimmten Gruppe nahe an der Verteilung des Attributs im gesamten Datensatz liegt, wodurch die Privatsphäre gewahrt und Diskriminierung aufgrund sensibler Attribute verhindert wird. Ein gemeinsamer Nenner dieser Ansätze ist die Abwägung der Diskriminierungskontrolle mit dem Nutzen der verarbeiteten Daten, d. h. die Minimierung der Verzerrung, ohne die Genauigkeit, Repräsentativität und allgemeine Nützlichkeit der Daten für prädiktive Modellierung oder Analyse erheblich zu beeinträchtigen . Dieses Gleichgewicht ist von entscheidender Bedeutung, um sicherzustellen,

Insgesamt kann die Vorverarbeitungsmethode weiter in drei Kategorien unterteilt werden: (1) Datenänderung, (2) Datenentfernung und (3) Datenneuabtastung. Methoden der ersten Kategorie zielen darauf ab, die Werte der Trainingsdatenpunkte (einschließlich geschützter Attributwerte , Klassenwerte und Merkmalswerte) zu ändern, um die Verzerrung im Datensatz zu verringern. Ein Beispiel für diese Methode ist die von [15] vorgeschlagene Datenmassage . Ihr Ansatz ordnet die Trainingsdaten und Daten nahe der Entscheidungsgrenze in sowohl privilegierten als auch nicht privilegierten Gruppen werden umgedreht. Alternativ wurde von Fahse et al. [23] eine optimierte Vorverarbeitungsmethode vorgeschlagen, die eine probabilistische Transformation lernt, die die Klassen und Merkmale mit individueller Verzerrung und Gruppengerechtigkeit bearbeitet. In [36] werden die ursprünglichen Attributwerte durch Werte ersetzt, die unabhängig vom Klassenlabel gewählt werden, um ein Modell zu trainieren, das ungefähr ausgeglichene Chancen erreicht. In ähnlicher Weise ersetzen Peng et al. [37] die geschützten Attributwerte durch Werte, die basierend auf anderen Attributen vorhergesagt werden, ähnlich der Datenimputation.

Methoden der zweiten Kategorie zielen darauf ab, ein faires Modell zu trainieren, indem bestimmte Merkmale aus dem Trainingsdatensatz entfernt werden. Ein Beispiel für diese Methode ist die von Dhar et al. [38] vorgeschlagene Datenunterdrückung. In ihrem Artikel werden die geschützten Attribute und Merkmale, die stark mit geschützten Attributen korreliert sind (auch als Proxy-Attribute bekannt), aus dem Datensatz entfernt, um ein faires Modell zu trainieren.

Methoden der dritten Kategorie zielen darauf ab, ein faires Modell zu trainieren, indem entweder die Stichprobengewichte angepasst oder der Datensatz überabgetastet wird. Krasanakis et al. [39] schlugen beispielsweise eine Neugewichtungsmethode vor, bei der die Gewichte der Trainingsstichproben iterativ mit einem theoretisch fundierten Modell angepasst werden, um den Kompromiss zwischen Bias und Genauigkeit zu mildern. In [40] schlugen Chakraborty et al. FairSMOTE als Methode vor, um Trainingspunkte von Minderheitengruppen mit künstlichen Datenpunkten auf der Grundlage der Synthetic Minority Oversampling Technique (SMOTE) [18] zu überabtasten, um ausgewogene Klassenverteilungen zu erreichen. Auch Yan et al. [41] schlugen vor, die Trainingsdaten der Minderheitengruppen mit künstlichen Datenpunk

Machine Translated by Google

Im Gegensatz zu FairSMOTE konzentrierten sich die Autoren auf Szenarien, in denen geschützte Attribute unbekannt sind, und verwendeten eine Clustermethode, um verschiedene demografische Gruppen zu identifizieren.

Bei der Inprocessing-Methode werden Methoden eingesetzt, die die Art und Weise ändern, wie ein ML-Modell trainiert wird, um Verzerrungen zu reduzieren. In [42] wurde ein Ansatz zur kontradiktorischen Debiasing-Methode vorgeschlagen. Dieser Ansatz lernt einen Klassifikator, um die Genauigkeit und Fairness der Vorhersage zu erhöhen, indem eine Variable für die interessierte Gruppe einbezogen wird, indem gleichzeitig ein Prädiktor und ein Gegner gelernt werden. Dies führt zur Generierung eines unvoreingenommenen Klassifikators, da die Vorhersagen keine Informationen zur Gruppenunterscheidung enthalten , die der Gegner nutzen könnte. Alternativ wurde in **[43] ein Algorithmus vorgeschlagen, der eine Fairnessmetrik als Teil der Verlustfunktion verwendet und ein für diese Fairnessmetrik** trainiertes Modell zurückgibt. **Kamishima et al. [22] schlugen eine Regularisierungsmethode vor , die einen Strafterm in die Verlustfunktion eines Klassifikators einbezog, um eine unvoreingenommene Vorhersage zu erzielen. Zafar et al. [44] entwickelten eine neue Gewichtungsmethode, bei der sie das Stichprobengewicht für jedes Trainingsdatum anpassen, um ein bestimmtes Fairnessziel zu erreichen, wie z. B. ausgeglichene Chancen für die Validierungsdaten. In jüngster Zeit wurde die Bias-Minderung als ein Optimierungsproblem mit Nebenbedingungen angegangen, indem eine Fairness-Nebenbedingung hinzugefügt und der Verlust so optimiert wurde, dass er mit dieser Nebenbedingung übereinstimmt [45,46]. Einige Arbeiten modifizieren außerdem n**

Nachbearbeitungsmethoden verringern den Fehler nach der Anpassung eines ML-Modells und umfassen Ansätze wie Kalibrierung, Constraint-Optimierung und Transformationsschwellenwertbildu Solche Methoden schlagen einen Algorithmus vor, der benachteiligten Gruppen günstige Ergebnisse und begünstigten Gruppen ungünstige Ergebnisse innerhalb eines gegebenen Konfidenzintervalls um die Entscheidungsgrenze mit der höchsten Unsicherheit zuweist. Ein Ansatz beispielsweise modifiziert die Spitzenschwellen des Klassifikators, um ein bestimmtes Ziel gleicher Chancen oder gleicher Quoten zu erreichen. Ein weiterer Ansatz beinhaltet die zufällige Mutation der Klassen bestimmter Vorhersagen in verschiedene Klassen [48].

Mehrere neuere Studien [49,50] kombinierten entweder Vorverarbeitung, Zwischenverarbeitung oder Nachverarbeitung zu einer Ensemble-Methode. Beispielsweise kombinieren Bhaskaruni et al. **[50] die Überabtastung der ungleichgewichtsgeschützten Klasse mit einer Entscheidungsgrenzenverschiebung als Nachverarbeitungsmethode, um das Ungerechtigkeitsproblem anzugehen.**

Forscher haben sich mit verschiedenen Konzepten der Diskriminierung und Fairness bei **algorithmischen Entscheidungsprozessen befasst .** Ungleiche Auswirkungen (früher als indirekte Fairness bezeichnet) **werden beispielsweise durch statistische Parität und Gruppenfairness gemessen, wie von Bhaskaruni et al. [50] diskutiert. Andererseits betont das Konzept der individuellen Fairness, das ebenfalls von Bhaskaruni et al. eingeführt wurde, dass ähnliche Personen unabhängig von ihrer Gruppenzugehörigkeit ähnlich behandelt werden sollten. Dieser Ansatz konzentriert sich auf Fairness auf individueller Ebene und stellt sicher, dass Entscheidungen auf der Grundlage relevanter Attribute und nicht auf der Grundlage gruppenbasierter Stereotypen oder Vorurteile getroffen werden.**

Bei Klassifikatoren und anderen Vorhersagemodellen ist das Erreichen gleicher Fehlerraten über verschiedene Gruppen hinweg ein wichtiges Ziel, wie Zhang und Neill [16] hervorheben. Ebenso ist es von entscheidender Bedeutung, eine Kalibrierung **oder das Fehlen von Vorhersageverzerrungen in den Vorhersagen sicherzustellen, wie von Hardt et al. [21]** erörtert . **Die Spannung zwischen diesen Konzepten – Kalibrierung und gleiche Fehlerraten – wird jedoch** von Dwork et al. **[51] und Pleiss et al.** [52] **untersucht , was darauf hinweist, dass es eine Herausforderung sein kann, beide gleichzeitig zu erfüllen. Karimi-Haghighi und Castillo [53] präsentieren verwandte Arbeiten, in denen die Komplexität untersucht wird, die mit der Erreichung algorithmischer Fairness verbunden ist. Friedler et al. [54]** untersuchen **weiter die Kompromisse bei der Erfüllung verschiedener Definitionen algorithmischer Fairness, insbesondere aus der Perspektive der öffentlichen Sicherheit. Da sich unsere Arbeit eher auf die Vorverarbeitung als auf die Modellierung konzentriert, sind Überlegungen wie ausgeglichene Fehlerraten und Vorhersageverzerrungen weniger direkt anwend**

Aus unserer Überprüfung verschiedener Vorverarbeitungsmethoden geht hervor, dass keine Arbeiten durchgeführt wurden, die versuchten, die Fairness für zwei oder mehr geschützte Attribute gleichzeitig zu Außerdem konzentrierte sich die in früheren Arbeiten verwendete Sampling-Methode nur auf Sampling basierend auf Klassenbezeichnungen und nicht auf geschützte Kategorien. Daher wird in diesem Artikel die Vorverarbeitung betont, da sie den anpassungsfähigsten Aspekt der Data Science-Pipeline darstellt [55]. Die Vorverarbeitung ist insofern einzigartig, als sie nicht von der Wahl des Modellierungsalgorithm

Machine Translated by Google

integriert mit Datenfreigabe- und Veröffentlichungsmechanismen. Diese Unabhängigkeit und Flexibilität machen die Vorverarbeitung entscheidend, um Datenqualität und Fairness sicherzustellen, bevor analytische oder prädiktive Modellierungen erfolgen. Schließlich konzentrieren wir uns auf neue benutzerdefinierte Samplingmethoden, die die geschützte Kategorie im Datentraining abtasten, um ein faires Modell z

4. Methodik Der

Schwerpunkt unserer Arbeit liegt auf der Erforschung von Samplingmethoden, um die Fairness in ML zu verbessern, ohne dass die entsprechende Vorhersageleistung darunter leidet, und so den Kompromiss zwischen **Fairness und Leistung** zu mildern . **Daher werden vier neue Samplingmethoden vorgeschlagen, die sich auf das Erreichen dieses Ziels konzentrieren. Diese Samplingmethoden befassen sich mit dem Problem der unausgewogenen Klassen, das durch die neue Multikategorie entsteht, die durch die Kombination der geschützten Kategorien generiert wird. Es werden benutzerdefinierte Samplingmethoden benötigt, da die vorhandenen Methoden Daten auf Grundlage von Minderheits- und Mehrheitsklassen abtasten, aber um die Fairness zu mildern, werden die neuen Multikategorien so abgetastet, dass sie gleich sind. Dies wiederum erfordert eine Änderung der vorhandenen Samplingmethoden, um Daten auf Grundlage dieser neuen Kategorien abzutasten. Dies führt zu vier neuen Samplingmethoden: Oversampling geschützter Kategorien, proportionales Sampling geschützter Kategorien, SMOTE geschützter Kategorien (PC- SMOTE) und ADASYN geschützter Kategorien (PC-ADASYN).**

4.1. Oversampling geschützter Kategorien

Beim Oversampling geschützter Kategorien besteht der erste Schritt darin, die geschützten Kategorien im Datensatz zu kombinieren und die Kombination zu kodieren, um unsere neue Multikategorie zu erzeugen. Beispielsweise werden im Datensatz „Erwachseneneinkommen“ Alter (jung und erwachsen), Rasse (weiß und andere) und Geschlecht (männlich und weiblich) kombiniert, um acht neue Kategorien zu erzeugen, die jeweils **ADULTWHITEMALE, ADULTOTHERSMALE, YOUNGWHITEMALE, YOUNGOTHERSMAIL, ADULTWHITEFEMALE , ADULTOTHERSFEMALE , YOUNGWHITER Diese neuen Kategorien haben unterschiedliche Stichprobengrößen, und das Ziel unseres Oversampling für geschützte Kategorien besteht darin, diese neue Kategorie so auszugleichen, dass die Stichprobengröße jeder der neuen Kategorien der Größe der Kategorie mit der größten Stichprobengröße entspricht. Um Datenlecks zu vermeiden, wird der Datensatz in Training und Test aufgeteilt, wobei Oversampling nur auf die Trainingsdaten angewendet und dann mit einem ungesampelten Testsatz getestet wird.**

Der Pseudocode in Algorithmus 1 zeigt unsere Methode zum Oversampling geschützter Kategorien im Detail. Im Algorithmus wurde die größte Kategorie als Basis verwendet, da dies die Kategorie mit der größten Stichprobengröße ist. Der Sampling-Prozess führt zu neuen Trainingsdaten mit einer ausgewogenen Stichprobengröße für die neue Kategorie. Der Algorithmus funktioniert, indem er die restlichen geschützten Kategorien so abtastet, dass sie der Stichprobengröße der Basis entsprechen. Dieses Sampling wird durchgeführt, indem die Kategorien zusammen mit ihren Klassenbezeichnungen mehrmals

4.2. Proportionale Stichprobenziehung mit geschützten Kategorien

Die Methode der proportionalen Stichprobennahme für geschützte Kategorien ist eine Verallgemeinerung der Überstichprobennahme für geschützte Kategorien, da der Prozess mit der Festlegung einer Zielstichprobengröße beginnt (die ein anzupassender Hyperparameter ist und nicht nur die Größe der größten Mehrfachkategorie), die als Zielstichproben bezeichnet wird. **Dies entspricht der gewünschten Anzahl von Instanzen, die für jede Kategorie benötigt werden. Dieses Ziel stellt Einheitlichkeit über alle Kategorien hinweg sicher und verringert das Risiko einer Modellverzerrung zugunsten häufigerer Kategorien. Das typische Ergebnis der Anwendung dieser Methode ist, dass einige Kategorien, die mehr Stichproben als die** Zielstichproben haben, **unterabgetastet werden, während andere überabgetastet werden, um einen gleichen Anteil von ihnen im Trainingsdatensatz zu erhalten . Der Pseudocode in Algorithmus 2 zeigt die schrittweise Vorgehensweise der Methode der proportionalen Stichprobennahme für geschützte Kategorien.**

Machine Translated by Google

Algorithmus 1: Oversampling geschützter Kategorien

```
baselineCount ÿ Summe der Einträge in „Largest_Category“ von Xtrain
totalCount ÿ Anzahl der Einträge in Xtrain
Basislinienanteil ÿ Basislinienanzahl/Gesamtanzahl
balancedData ÿ Initialisiere einen leeren Datensatz
Kategorien ÿ Liste der Spaltennamen in Xtrain, die mit „combined_category_“ beginnen
für jede Kategorie in Kategorien tun
    categoryData ÿ Einträge in Xtrain auswählen , bei denen Kategorie = 1
    categoryData ÿ kombiniere categoryData mit entsprechenden Labels von ytrain
    categoryCount ÿ Anzahl der Einträge in categoryData
    targetCount ÿ ganzzahliger Teil von totalCount × baselineProportion
    wenn Kategorieanzahl < Zielanzahl dann
        sampledData ÿ Beispiel targetCount aus categoryData mit Ersetzung
        balancedData ÿ sampledData an balancedData anhängen
    anders
        balancedData ÿ Kategoriedaten an balancedData anhängen
Ende, wenn
Ende für
returniere ausgeglichene Daten
```

Algorithmus 2: Proportionale Stichprobennahme mit geschützter Kategorie

```
Zielproben ÿ 5000
sampledBalanced ÿ Initialisieren Sie einen leeren Datensatz
für jede Spalte in new_categories.columns machen
    categoryRows ÿ Zeilen in new_categories auswählen , in denen Spalte = 1
sampledRows ÿ Beispiel targetSamples Einträge aus categoryRows mit Ersetzung
    für jede Spalte in oneHotEncodedBalanced.columns
        abgetasteteZeilen[Spalte] ÿ 0
Ende für
    sampledRows[Spalte] ÿ 1
    sampledBalanced ÿ sampledRows an sampledBalanced anhängen
Ende für
returniere sampledBalanced
```

4.3. Geschützte Kategorie SMOTE

Die Protected-Category Synthetic Minority Oversampling-Technik (PC-SMOTE)

Die Sampling-Methode ist ein komplexer Prozess, der SMOTE nachahmt, jedoch modifiziert wurde für Stichproben unserer neuen Kategorien statt Klassenbezeichnungen. Bei diesem Ansatz bestand der erste Schritt darin, Ändern Sie SMOTE so, dass eine feste Anzahl von Nachbarn verwendet wird und ein Nachbar zufällig ausgewählt wird für die Interpolation, anstatt alle zu mitteln. Der Pseudocode in Algorithmus 3 zeigt das Verfahren für den PC-SMOTE. Da die Absicht der Methode darin besteht, sie für die neue zu verwenden Kategorie-Sampling, befasst sich nicht direkt mit der Generierung von Klassenbezeichnungen. Daher wird ein neuer Es wird eine Funktion benötigt, die eine neue Klassenbezeichnung für die synthetischen Daten generieren kann. Dazu wird eine neue Es wird eine Funktion definiert, die Klassenbezeichnungen basierend auf der Anzahl der neuen synthetischen Daten generiert generiert und ein vorgewähltes Gleichgewichtsverhältnis zwischen den beiden Klassen. Algorithmus 4 zeigt, wie Unsere neue Funktion generiert Labels für unsere synthetischen Proben. Der Algorithmus bestimmt zunächst die Anzahl der Proben für jede Klasse auf Grundlage des Bilanzverhältnisses und generiert die Probe für jede Klasse erforderlich. Die Klassenbezeichnungen dienen der Umstellung, um eine algorithmische Verzerrung in der Klassen generiert.

Diese beiden Algorithmen werden zu PC-SMOTE kombiniert, wie in Algorithmus 5 gezeigt.

Bei dem Ansatz zur Erzielung eines Gleichgewichts in mehreren Kategorien wird jede einzelne Kategorie iteriert über so dass die Teilmenge der Daten identifiziert wird, die dieser Kategorie zugeordnet ist. Die Nummer Die Anzahl der synthetischen Samples, die zum Erreichen einer vordefinierten Maximalgröße pro Kategorie erforderlich sind, beträgt dann berechnet. Wenn zusätzliche Proben benötigt werden, werden die Daten mit PC-SMOTE generiert,

Machine Translated by Google

die zwischen vorhandenen Datenpunkten und ihren nächsten Nachbarn interpoliert. Gleichzeitig
Eine ausgewogene Verteilung der synthetischen Klassenbezeichnungen wird mit einem festgelegten Balanceverhältnis erzeugt durch
unter Verwendung von Algorithmus 4. Diese synthetischen Merkmale und Bezeichnungen werden dann in die Trainingsteilmenge für jede Kategorie. Der Vorgang wird für alle Kategorien wiederholt, was zu einem ausgeglichener Datensatz. Die Hyperparameter in diesem Algorithmus 5 sind die Anzahl der Nachbarn
und Gleichgewichtsverhältnis.

Algorithmus 3: Benutzerdefinierte Synthetic Minority Oversampling-Technik (SMOTE)

```
1: Prozedur CUSTOMSMOTE(Daten, n_Beispiele)
2:     syntheticSamples ÿ Nullmatrix der Größe (n_samples× Anzahl der Spalten in den Daten)
3: nn ÿ Nächste Nachbarn(n_Nachbarn = 7).fit(Daten)
4:     Nachbarn ÿ nn.kneighbors(Daten, return_distance = False)
5:     für i ÿ 1 bis n_samples mache
6:         sampleIdx ÿ Zufallszahl von 0 bis (Anzahl der Zeilen in den Daten ÿ 1)
7:         nnIdx ÿ zufällige Auswahl aus Nachbarn[sampleIdx, 1 :]
8:         di ff ÿ Daten[nnIdx] ÿ Daten[BeispielIdx]
9:         Gewicht ÿ Zufallszahl aus Gleichverteilung zwischen 0 und 1
10:        syntheticheSamples[i] ÿ Daten[SampleIdx] + Gewicht × di ff
11: Ende für
12:    synthetische Proben zurückgeben
13: Prozedur beenden
```

Algorithmus 4: Generieren Sie ausgewogene synthetische Etiketten

```
1: Prozedur GENBALSYNTHLABELS(n_samplesNeeded, balanceRatio)
2:     nClass1 ÿ int(n_benötigte Samples × Bilanzverhältnis)
3:     nKlasse0 ÿ n_benötigte Proben ÿ nKlasse1
4:     syntheticheLabels ÿ [0] × nKlasse0 + [1] × nKlasse1
5:     SHUFFLE(syntheticheLabels) gibt                ÿ Die Etiketten nach dem Zufallsprinzip mischen
6:     syntheticheLabels zurück
7: Prozedur beenden
```

Algorithmus 5 Geschützte Kategorie SMOTE

```
1: balancedDataList ÿ eine leere Liste initialisieren
2: für jede Kategorie in Kategorien tun
3:     categorySubset ÿ Zeilen in train_data auswählen st 'combined_category' == Kategorie
4:     Funktionen ÿ entferne „Klasse“, „kombinierte_Kategorie“ aus der Kategorie-Untermenge
5:     nSamples ÿ max_sizeÿ Anzahl der Zeilen im Kategorie-Subset
6:     wenn nSamples > 0 dann
7:         syntheticheFeatures ÿ PCSmote( features, nSamples)
8:         syntheticheLabels ÿ GenBalSynthLabels(nSamples, balanceRatio)
9:         syntheticheEigenschaften[ÿ Klasseÿ ] ÿ syntheticheLabels
10:        syntheticheEigenschaften[ÿ kombinierte_Kategorieÿ ] ÿ Kategorie
11:        categorySubsetBalanced ÿ verketten categorySubset und syntheticFeatures
12:    anders
13:        KategorieSubsetBalanced ÿ KategorieSubset
14: Ende, wenn
15: categorySubsetBalanced an balancedDataList anhängen
16: Ende für
17: balancedData ÿ balancedDataList anhängen und Index zurücksetzen
```

4.4. Geschützte Kategorie ADASYN

Die Methode der geschützten Kategorie ADASYN imitiert die adaptive synthetische Minderheit (ADASYN) Sampling, wird aber leicht modifiziert, um unser Ziel des geschützten Kategorie-Samplings zu erfüllen. Unser PC-ADASYN-Algorithmus ist in Algorithmus 6 dargestellt. Er erweitert ADASYN, indem er sich a

Machine Translated by Google

Dichte statt Klassenungleichgewicht. Genauer gesagt funktioniert diese Funktion, indem sie die nächste Nachbarn der Daten und dann die Berechnung der Dichte der Kategorie jedes Datenpunkts innerhalb seiner unmittelbare Nachbarschaft. Es gewichtet diese Dichten umgekehrt, um Minderheitenkategorien zu priorisieren, wodurch es wahrscheinlicher wird, synthetische Stichproben aus unterrepräsentierten Kategorien zu generieren. Synthetische Stichproben werden durch Interpolation zwischen ausgewählten Datenpunkten und ihren Nachbarn erstellt. ähnlich wie SMOTE, jedoch mit zufälliger Gewichtung zur Variation der Interpolation, wodurch eine vielfältiger synthetischer Datensatz. Dieser Ansatz hilft, das Ungleichgewicht auf Kategorieebene zu beheben und bereichert die Varianz des Datensatzes und verbessert potenziell die Robustheit und Fairness von ML-Modellen trainiert. Da diese Sampling-Methode durch Interpolation auch neue Samples erzeugt, Algorithmus 4 wird verwendet, um Klassenbezeichnungen für die neuen synthetischen Proben zu generieren.

Algorithmus 6 PC-ADASYN für kategoriebasiertes Balancing

```
1: Prozedur PCADASYNCATEGORIES (Daten, Beschriftungen, n_benötigte Proben, n_Nachbarn)
2:     n_Nachbarn ÿ n_Nachbarn + 1 3: nn ÿ                ÿ Einbeziehung des Datenpunkts selbst
NächsteNachbarn(n_Nachbarn).fit(Daten)
4:     Distanzen, Indizes ÿ nn.kneighbors(data)
5:     Dichten ÿ Null-Array mit Länge (Daten)
6:     für i ÿ 0 bis Länge(Daten) ÿ 1 mache
7:         aktuelle_Kategorie ÿ Beschriftungen[i]
8:         Nachbarindizes ÿ Indizes[i][1 :] Dichten[i] ÿ          ÿ Überspringen Sie den Selbstindex
9:         SUMME(Beschriftungen[Nachbarindizes] == aktuelle_Kategorie)
10: Ende für
11:    Gewichte ÿ 1/(Dichten + 1) Gewichte          ÿ Addiere 1, um eine Division durch Null zu verhindern
12:    ÿ Gewichte/SUMME(Gewichte) synthetische                ÿ Gewichte normalisieren
13:    Proben ÿ leere Liste
14:    sampleIndices ÿ zufällige Auswahl mit Ersetzung aus length(data) unter Verwendung von Gewichten
15:    für jeden idx in sampleIndices
16:        BasisIdx ÿ idx
17:        Nachbar-Idx ÿ RANDOMCHOICE(Indizes[Basis-Idx][1 :])
18:        di ff ÿ Daten[Nachbar-Idx] ÿ Daten[Basis-Idx]
19:        syntheticheProbe ÿ Daten[BasisIdx] + RANDOM() × di ff
20:        SyntheticSample an SyntheticSamples anhängen
21: Ende für
22:    Array zurückgeben (synthetische Samples)
23: Ende der Prozedur
```

Die Algorithmen 4 und 6 werden kombiniert, um unser geschütztes ADASYN-Sampling zu bilden. wie in Algorithmus 7 gezeigt. Für jede Kategorie werden die dieser Kategorie entsprechenden Daten isoliert und das Größendefizit relativ zur größten Kategorie berechnet. Wenn zusätzliche Proben benötigt werden, wird die PC-ADASYN-Methode angewendet, die synthetische Merkmale erzeugt, die die Verteilungseigenschaften der Kategorie zu berücksichtigen. Diese Merkmale werden dann ergänzt mit synthetisch generierten Labels, wobei ein vordefiniertes Klassengleichgewichtsverhältnis beibehalten wird. Der Prozess korrigiert nicht nur Kategorieungleichgewichte, sondern bereichert auch den Datensatz, möglicherweise Verbesserung der Vorhersagegenauigkeit und Fairness von Modellen, die mit diesen Daten trainiert werden.

4.5. Datensatz- und Hyperparameter-Tuning

Um die vier Stichprobenverfahren zu testen, wurde ein Klassifikator benötigt, um die Auswirkungen auf Fairness und Leistung. Letztlich ist der Typ des Klassifikators nicht direkt relevant, da die Ziel ist es, den Kompromiss zwischen Fairness und Leistung zu mildern, anstatt den besten Klassifikator zu finden. Aus diesem Grund haben wir uns für die Verwendung univariater Entscheidungsbäume auf Basis von CART [56] entschieden , da diese Robustheit gegenüber Rauschen und fehlenden Daten. Die spezifische Implementierung, die wir gewählt haben, war aus der sklearn-Bibliothek (Version 1.5.1) [57]. Darüber hinaus ermöglichen Entscheidungsbäume Steuern Sie die Stärke der Anpassung, indem Sie die Baumtiefe des gelernten Baums festlegen. Dies ermöglicht es uns, Vergleichen Sie Fairness und Leistung auf verschiedenen Anpassungsebenen.

Machine Translated by Google

Der Prozess beginnt mit der Kombination geschützter Kategorien innerhalb jedes Datensatzes, der Anwendung einer One-Hot-Kodierung zum Erstellen neuer Multikategoriefunktionen und der anschließenden Durchführung einer Beschriftungsk

Die Datensätze wurden dann mithilfe einer geschichteten 10-fachen Kreuzvalidierung aufgeteilt, um eine repräsentative Verteilung der Klassen in jeder Faltung sicherzustellen. Für jede Faltung wurde das Training mit Stichprobendaten unter Verwendung der zuvor beschriebenen Methoden durchgeführt, während die Klassifizierung mit den entsprechenden nicht gesampelten Testsätzen getestet wurde. Die Konsistenz beim Modelltraining wurde durch Anwenden identischer Baumtiefen auf alle Stichprobenmethoden gewahrt. Die bereitgestellten Ergebnisse sind Durchschnittswerte der 10-fachen Durchläufe mit den entsprechender

Algorithmus 7 Geschützte Kategorie ADASYN

```
1: balancedDataList ÿ Initialisiere eine leere Liste. 2: Für
jede Kategorie in Kategorien führe
3:     categorySubset aus ÿ Wähle aus new_data3 st 'combined_category' == category f eatures ÿ
4:     Entferne 'class', 'combined_category' aus categorySubset categoryLabels ÿ Extrahiere
5:     'combined_category' aus categorySubset nSamplesNeeded ÿ max_size minus
6:     Anzahl der Zeilen in categorySubset, wenn nSamplesNeeded > 0, dann syntheticFeatures
7:     ÿ PCAdasyn(f eatures,
8:         categoryLabels, nSamplesNeeded) syntheticLabels ÿ GenBalSynthLabels(nSamplesNeeded,
9:         balanceRatio) syntheticFeatures[ ÿ classÿ ] ÿ syntheticLabels syntheticFeatures[ ÿ
10:        combined_categoryÿ ] ÿ Kategorie categorySubsetBalanced
11:        ÿ verkette categorySubset mit syntheticFeatures, sonst
12:        categorySubsetBalanced ÿ categorySubset 15: Ende wenn 16: hänge categorySubsetBalanced
13:    an
14:        balancedDataList an 17: Ende für 18: balancedData ÿ
verkette
balancedDataList und setze Index zurück
```

Für unsere Analyse **wurden drei Datensätze aus dem UCI-Repository [58]** ausgewählt : **der Datensatz „Adult Income“ [59], der Datensatz „German Credit “ [60] und der Datensatz „Correctional Offender Management Profiling for Alternative Sanctions“ (COMPAS). Der Datensatz „Adult Income“ soll vorhersagen, ob eine Person mehr als 50.000 US-Dollar verdient. Er enthält acht kategorische und vier numerische Attribute mit geschützten Variablen entsprechend Alter (jung oder erwachsen), Geschlecht (männlich oder weiblich) und Rasse (weiß oder andere). Der Datensatz „Adult Income“ wurde der UCI 1996 gespendet. Der Datensatz „German Credit“, der zur Vorhersage der Kreditwürdigkeit verwendet wird, umfasst 20 kategorische und zwei numerische Attribute mit geschützten Variablen für Alter und Geschlecht. Der Datensatz „German Credit“ wurde der UCI 1994 gespendet. Der COMPAS-Datensatz [61], der die Rückfallquoten in den USA bewertet, umfasst sechs kategorische und sechs numerische Merkmale mit geschützten Variablen für Alter, Rasse und Geschlecht. Der Datensatz wurde 2018 veröffentlicht. Diese Datensätze wurden ausgewählt, weil sie den neuesten Stand der Technik zur Messung von Voreingenommenheit und Diskriminierung darstellen und in anderen Studien zu algorithmischer Voreingenommenheit und Fairness weit verbreitet sind (siehe Abschnitt 3). Außerdem haben die Datensätze unterschiedliche Größen, von klein bis groß, was sie zum Testen unserer Stichprobenmethoden geeignet macht.**

Die **Hyperparameter-Optimierung wurde mithilfe der Grid-Suche [62]** durchgeführt , **um ein breites Spektrum an Parametern zu untersuchen. Ergänzt wurde dies durch visuelle Bewertungen, um optimale Einstellungen zu ermitteln, die Equalized Odds Difference (EOD) und Genauigkeit ausbalancieren. Für den Datensatz „Adult Income“ umfassten die optimalen Hyperparameter eine maximale Baumtiefe von 3 und für PC-SMOTE und PC-ADASYN eine Nearest-Neighbor-Einstellung von 5 mit einem ausgeglichenen Verhältnis von 0,34. Diese Parameter waren für den Datensatz „German Credit“ ähnlich effektiv. Für den COMPAS-Datensatz war eine maximale Baumtiefe von 2 für alle Stichprobenverfahren optimal. PC-SMOTE und PC-ADASYN wurden auf die Nearest-Neighbor-Einstellung von 3 und ein ausgeglichenes Verhältnis von 0,60 eingestellt.**

Machine Translated by Google

5. Ergebnisse

Die vier Stichprobenstrategien wurden auf die drei oben beschriebenen Datensätze angewendet und evaluierten ihre Auswirkungen mithilfe eines einfachen univariaten Entscheidungsbaum-Klassifikators. Die Ergebnisse in Die Tabellen 1–3 zeigen erhebliche Unterschiede in der Modellleistung bei fünf Sampling-Strategien: kein Sampling, Oversampling, proportionales Sampling, PC-SMOTE und PC-ADASYN ein unsere drei Datensätze. Jede Methode wurde anhand von Genauigkeit, Makro F1, Equalized Odds Difference (EOD) und Statistische Parität (SP). Die Ergebnisse wurden in Genauigkeit gemessen und Makro F1, da diese beiden Metriken die beliebtesten Klassifizierungsmetriken sind. Außerdem Durch die Beschränkung der Metrik auf zwei werden die Ergebnisse für die statistische Analyse vergleichbar.

Um die statistische Signifikanz der Ergebnisse jeder Methode zu ermitteln, verwendeten wir den Friedman Test, eine nichtparametrische Alternative zur einfaktoriellen ANOVA mit wiederholten Messungen. Da der Friedman-Test signifikante Ergebnisse lieferte, führten wir den Nemenyi-Post-hoc-Test durch. Test. Dieser Test wird verwendet, um paarweise Vergleiche zwischen den Methoden zur Ermittlung durchzuführen welche Methoden sich statistisch voneinander unterscheiden. Der Nemenyi-Test ist vorteilhaft in dieser Einstellung, weil es mehrere Vergleiche berücksichtigt, ohne normale Verteilungen und bieten somit eine robuste Möglichkeit zum Verständnis spezifischer paarweiser Unterschiede.

Für den Datensatz „Einkommen Erwachsener“ (Tabelle 1) ergab die No-Sampling-Methode eine Genauigkeit von 0,82, was eine hohe Vergleichsbasis darstellt. Es zeigte jedoch eine leicht verzerrte Vorhersage mit einem EOD von 0,36 und minimaler Disparität in den Vorhersageraten (SP = 0,02). Im Gegensatz dazu behielt Oversampling die gleiche Genauigkeit bei, verringerte aber die Makro-F1 leicht auf 0,65, was auf mögliche Überanpassungsprobleme hindeutet und gleichzeitig die Fairness verschlechtert (EOD = 0,66) und zunehmende Disparität in den Vorhersageraten (SP = 0,09). Die proportionale Stichprobennahme verringerte Genauigkeit auf 0,79, verbesserte aber das Makro F1 auf 0,79, was auf eine bessere Balance zwischen Präzision und Rückruf. Allerdings erhöhte sich der SP signifikant auf 0,71, was auf eine erhebliche Ungleichheit bei den positiven Vorhersageraten, was Zweifel an der Fairness des Modells aufkommen lässt. Zwei benutzerdefinierte Ansätze für SMOTE und ADASYN wurden speziell entwickelt, um auf diesen Metriken. PC-SMOTE zeigte eine mäßige Leistung mit einer Genauigkeit von 0,81 und ein verbesserter EOD von 0,25, was auf eine verbesserte Fairness gegenüber einfachem Oversampling hindeutet und die No-Sampling-Methode. Allerdings wurde immer noch ein niedrigerer Makro-F1-Wert (0,63) aufgezeichnet, was darauf hindeutet, Falschdarstellungsprobleme bei der Generierung synthetischer Daten. PC-ADASYN erwies sich als ausgewogenster Ansatz, der eine hohe Genauigkeit (0,82) und eine bessere Handhabung der Klasse beibehält Ungleichgewichte, mit einer moderaten Verbesserung der Fairness (EOD = 0,28) und einer kontrollierten Zunahme der Vorhersageraten-Disparität (SP = 0,09). Insgesamt ist die Basisgenauigkeit statistisch deutlich besser als die proportionale Stichprobe, aber nicht statistisch signifikant, da im Vergleich zu den anderen Stichprobenverfahren. Für das Makro F1 ist die proportionale Stichprobennahme statistisch signifikant besser als andere Probenahmemethoden. Für den EOD ist PC-ADASYN ist statistisch besser als andere Stichprobenverfahren, während die No-Sampling-Methode SP statistisch signifikant besser als andere Stichprobenmethoden.

Tabelle 1. Ergebnisse der Stichprobenmethoden für den Datensatz „Einkommen Erwachsener“ mit 95ÿ%-Konfidenzintervallen.

Probenahme Methode	Genauigkeit	Makro F1	EOD	SP
Keine Stichprobennahme	0,82 ± 0,00	0,66 ± 0,01	0,36 ± 0,18	0,02 ± 0,00
Überstichprobe	0,82 ± 0,02	0,65 ± 0,06	0,66 ± 0,2	0,09 ± 0,02
Eigenstichprobe	0,79 ± 0,05	0,79 ± 0,06	0,46 ± 0,12	0,71 ± 0,10
PC-SMOTE	0,81 ± 0,05	0,63 ± 0,07	0,25 ± 0,21	0,07 ± 0,00
PC-ADASYN	0,82 ± 0,04	0,64 ± 0,07	0,28 ± 0,19	0,09 ± 0,02

Auch die Ergebnisse der Experimente mit dem deutschen Kreditdatensatz zeigen unterschiedliche Auswirkungen jeder Stichprobenstrategie, insbesondere im Hinblick auf Fairness und Genauigkeit, wie in Tabelle 2 dargestellt. Ohne Stichproben erreichte das Basismodell eine Genauigkeit von 0,72, wies aber erhebliche Verzerrung in der Vorhersage, mit einem EOD von 0,88, was auf eine erhebliche Fehlerdisparität hinweist Raten zwischen Gruppen in den geschützten Attributen. Die Implementierung von Oversampling hielt Genauigkeit bei gleichzeitiger Verbesserung des Makro-F1 auf 0,68 und deutlicher Reduzierung des EOD auf 0,37, wenn auch bei

Machine Translated by Google

die Kosten für die Erhöhung des SP auf 0,35, was einen möglichen Kompromiss zwischen unterschiedlichen Fairness-Maßnahmen. Die proportionale Stichprobennahme reduzierte die Genauigkeit leicht auf 0,68, erreichte aber die beste F1-Score von 0,69. Es senkte auch den EOD auf 0,32, was darauf hindeutet, dass es die Vorhersage effektiv ausgleicht Qualität mit Fairness. PC-SMOTE zeigt eine Verbesserung der Genauigkeit mit 0,73, aber eine geringere Makro-F1 von 0,55; die Fairness des Modells zeigt eine enorme Verbesserung gegenüber der Baseline mit ein EOD von 0,15 und ein moderater SP von 0,1. PC-ADASYN zeigt eine ähnliche Genauigkeit wie die Basislinie bei 0,72, allerdings mit dem niedrigsten Makro-F1 von 0,48, was auf einen möglichen Kompromiss hindeutet in Präzision und Rückruf. Das Modell weist jedoch die beste Fairnessvorhersage auf mit ein EOD von 0,13 und ein SP von 0,06. Insgesamt zeigt das Ergebnis, dass die Genauigkeit von No-Sampling ist im Vergleich zu anderen Stichprobenverfahren mit Ausnahme von proportionalen Stichproben nicht statistisch signifikant während für den EOD die Ergebnisse aller Probenahmemethoden statistisch signifikant sind in Vergleich mit der No-Sampling-Methode.

Tabelle 2. Ergebnisse der Stichprobenmethoden für den deutschen Kreditdatensatz mit 95 %-Konfidenzintervallen.

Probenahme Methode	Genauigkeit	Makro F1	EOD	SP
Keine Abtastung	0,72 ± 0,03	0,62 ± 0,06	0,88 ± 0,10	0,07 ± 0,01
Überabtastung	0,72 ± 0,05	0,68 ± 0,04	0,37 ± 0,16	0,35 ± 0,11
Eigenabtastung	0,69 ± 0,04	0,69 ± 0,07	0,32 ± 0,26	0,29 ± 0,10
Abtastung PC-SMOTE	0,73 ± 0,05	0,55 ± 0,09	0,15 ± 0,09	0,1 ± 0,02
PC-ADASYN	0,72 ± 0,03	0,48 ± 0,11	0,13 ± 0,02	0,06 ± 0,00

Tabelle 3 zeigt die Ergebnisse unserer Erfahrungen mit dem COMPAS-Datensatz. Diese Ergebnisse offenbaren erhebliche Unterschiede in der Modellleistung zwischen den verschiedenen Sampling-Strategien. Der Baseline-Ansatz ohne Stichproben erreichte eine Genauigkeit und einen Makro-F1-Score von 0,89, zeigte aber größere Unterschiede in den Fairnessmetriken, mit einer ausgeglichenen Quotendifferenz (EOD) von 0,39 und einer statistischen Parität (SP) von 0,29. Dies unterstreicht mögliche Verzerrungen, die unangepasste Modelle können gegenüber geschützten Gruppen eine Tendenz aufweisen. Die Anwendung von Oversampling leicht verbesserte Genauigkeit auf 0,90, aber auch verbesserte Fairness deutlich, wodurch EOD auf 0,26. Dies deutet darauf hin, dass es effektiv ist, Ergebnisunterschiede zu verringern, ohne Kompromisse einzugehen SP. Umgekehrt, proportionale Abtastung, während die Genauigkeit und Makro F1 auf 0,90 und 0,91 bzw. erreichte auch einen EOD von 0,26 und verbesserte sich damit gegenüber dem Ausgangswert, während auch ein höherer SP von 0,36, was auf eine mögliche Zunahme der Ungleichheit positiver Ergebnisse über alle Gruppen hinweg. PC-SMOTE und PC-ADASYN, mit identischen Werten in Genauigkeit, Makro F1 und SP konnten die Fairnessverbesserungen mit einem EOD von 0,30 aufrechterhalten, obwohl diese Methoden auch den SP auf 0,47 erhöhten. Insgesamt zeigen die Ergebnisse, dass sowohl die Genauigkeit und EOD unserer Probenahmemethoden sind statistisch signifikant besser als die No-Sampling-Methode.

Tabelle 3. Ergebnisse der Stichprobenmethoden im COMPAS-Datensatz mit 95 %-Konfidenzintervallen.

Probenahme Methode	Genauigkeit	Makro F1	EOD	SP
Keine Abtastung	0,89 ± 0,04	0,89 ± 0,04	0,39 ± 0,15	0,29 ± 0,17
Oversample	0,90 ± 0,03	0,90 ± 0,05	0,26 ± 0,14	0,25 ± 0,07
Prop. Sample	0,90 ± 0,05	0,91 ± 0,04	0,26 ± 0,12	0,36 ± 0,10
PC-SMOTE	0,91 ± 0,02	0,91 ± 0,02	0,30 ± 0,11	0,47 ± 0,19
PC-ADASYN	0,91 ± 0,02	0,91 ± 0,03	0,30 ± 0,13	0,47 ± 0,21

6. Diskussion

Die Ergebnisse der Experimente mit den drei Datensätzen belegen, dass geschützte Kategorien Sampling kann die Modellgerechtigkeit deutlich verbessern, oft ohne die Vorhersagegenauigkeit wesentlich zu beeinträchtigen. In einigen Fällen wurden auch Verbesserungen bei der Genauigkeit und der Makro-F1 demonstriert. Konzentriert man sich auf die Ergebnisse des Adult Income-Datensatzes, so

Machine Translated by Google

zeigte bemerkenswerte Verbesserungen bei EOD und hielt moderate SP-Niveaus aufrecht. Die Wirksamkeit dieser Methoden ist größtenteils auf ihre ausgefeilten Interpolationstechniken zurückzuführen. Eine visuelle Untersuchung der ohne Sampling und PC-ADASYN generierten Entscheidungsbäume liefert beispielsweise aufschlussreiche Kontraste. Beispiele aus einer einzelnen repräsentativen Falte sind in Abbildung 1 bzw. 2 dargestellt. Der ohne Sampling gelernte Entscheidungsbaum wählte seine Wurzel mit einem Merkmal, das eng mit geschützten Attributen verbunden ist und somit als Proxy-Attribut fungierte. Dies führte zu einer ausgeprägten Vorh
Im Gegensatz dazu begann der mit PC-ADASYN generierten Daten trainierte Entscheidungsbaum mit einer Funktion, die Vorhersagen sehr gut verallgemeinerte und Verzerrungen milderte, was durch eine deutliche Verbesserung der Modellfairness und eine höhere Gini-Unreinheit belegt wurde, was auf eine reinere anfängliche Aufte

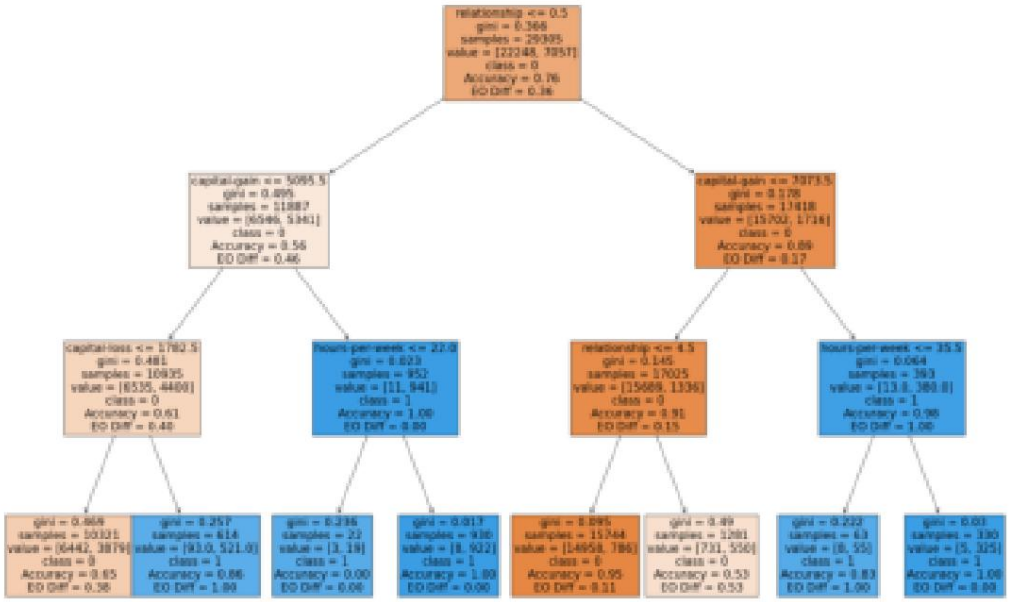


Abbildung 1. Beispiel eines Entscheidungsbaums, der ohne Stichprobennahme auf das Einkommen Erwachsener trainiert wurde.

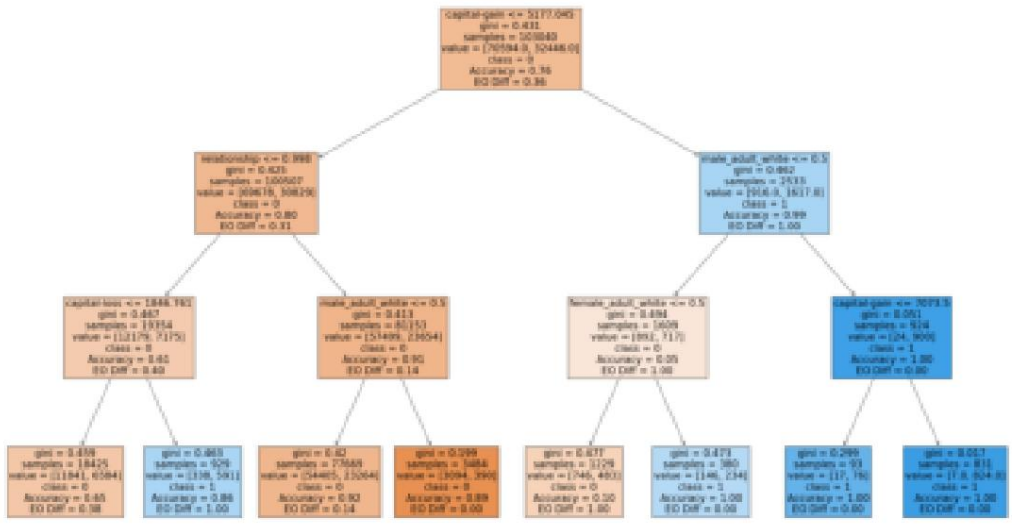


Abbildung 2. Entscheidungsbaum von PC-ADASYN zum Einkommen von Erwachsenen.

Machine Translated by Google

6.1. Vergleich von Fairness und Leistung

Beim Vergleich von Oversampling und proportionalem Sampling waren die Ansätze der Methoden zur Vergrößerung der Stichprobengröße durch Duplizierung vorhandener Datenzeilen unkompliziert und führten zu keinen wesentlichen Verbesserungen bei EOD. Dieses Ergebnis ist nachvollziehbar, da diese Methoden dazu neigen, vorhandene Verzerrungen zu replizieren, was Fairnessprobleme möglicherweise verschärfen kann, anstatt sie zu lindern. Dies wird insbesondere deutlich, wenn man den Datensatz „Erwachseneneinkommen“ betrachtet, bei dem die Klassen extrem unausgewogen sind. Diesen naiven Replikationsstrategien fehlt die Interpolationskapazität von PC-SMOTE und PC-ADASYN, um Stichproben in der Nähe von Entscheidungsgrenzen anzupassen, was für die Minderung der Verzerrung im Datensatz entscheidend ist. Im Gegensatz dazu erweitern die von SMOTE und ADASYN verwendeten Interpolationsstrategien den Datensatz und erhöhen seine Vielfalt. Dies ist besonders effektiv bei Stichproben in der Nähe von Entscheidungsgrenzen, bei denen leichte Verschiebungen der Merkmale die Fairness der Vorhersagen erheblich beeinträchtigen können. Durch Interpolation zwischen den Stichproben verschieben SMOTE und ADASYN diese Randstichproben effektiv in Richtung gleichmäßigerer Bereiche des Merkmalsraums. Auf diese Weise begegnen und reduzieren sie Verzerrungen effektiver als Methoden, die das Stichprobenvolumen erhöhen, ohne die Datenstruktur zu ändern. Die Klassengenerierungsfunktion (Algorithmus 4) trägt auch dazu bei, die allgemeine Klassenverteilung zu verbessern. Die Ergebnisse unserer PC-SMOTE- und PC-ADASYN-Studie zum Einkommen von Erwachsenen sind ebenfalls den in **[41] erzielten Ergebnissen überlegen , wo eine Genauigkeit von 0,59 und ein SP von 0,17 erzielt wu** Darüber hinaus sind die Ergebnisse von PC-SMOTE und PC-ADASYN den **in [40] erzielten Ergebnissen überlegen , wo ein EOD von 0,89 und eine etwas bessere Genauigkeit von 0,84 erreicht wurden.**

Bei der Untersuchung der Ergebnisse des deutschen Kreditdatensatzes haben wir einen ähnlichen Trend wie beim Erwachseneneinkommensdatensatz festgestellt: Die No-Sampling-Methode weist eine sehr hohe Verzerrung in Bezug auf EOD auf. Der niedrige SP von 0,07 weist auf minimale Unterschiede bei den positiven Vorhersageraten zwischen den Gruppen hin, aber dies allein ist kein gutes Maß für die Fairness, da die bevorzugte Gruppe mehr Stichproben hat als die benachteiligte. Dies hat zur Folge, dass die Berechnung des SP verzerrt wird, da nur positive Entscheidungen in jeder Gruppe gezählt werden, die von der Stichprobengröße beeinflusst werden. Eine Erkenntnis ist, wie wichtig es ist, mehrere Fairnessmetriken einzusetzen, um die Auswirkungen eines Modells auf alle Beteiligten umfassend zu betrachten. Beim Oversampling haben wir eine Verbesserung des EOD mit einem ähnlichen SP festgestellt; dies zeigt, dass eine Erhöhung der Anzahl der Stichproben für jedes der geschützten Attribute der Multikategorie die Fairness in Bezug auf EOD verbessert. Darüber hinaus spiegelt der aktualisierte SP wider, wie es aussehen würde, wenn für jede Multikategorie eine gleichmäßigere Anzahl von Stichproben vorhanden wäre, anders als in der Baseline, wo die bevorzugte Gruppe fünfmal mehr Stichproben hat als die unbegünstigte Gruppe. Die Genauigkeit der proportionalen Stichprobennahme sinkt, da die nach der Hyperparameter-Optimierung ausgewählte Baseline-Anzahl der Stichproben nicht ausreichte, damit das Modell den ungesampelten Testsatz verallgemeinern konnte, was zu einer Überanpassung führte. Die Überanpassung wurde durch Berücksichtigung der Trainingsgenauigkeit bestätigt. Interessanterweise tauscht das Modell nicht wie andere Modelle Genauigkeit gegen Rückruf ein, und dies verleiht der proportionalen Stichprobennahme die höchste Makro-F1.

In Bezug auf die Fairness stellten wir einen Anstieg des EOD im Vergleich zum Ausgangswert und zur Überabtastung fest . Dies liegt daran, dass jede Mehrfachkategorie auf denselben Ausgangswerten dargestellt wird. Dies kann die Fairness des Modells verbessern, da das Modell nun ein größeres Bild der Kategorien hat und bessere Vorhersagen und letztlich fairere Entscheidungen trifft. PC-SMOTE und PC-ADASYN spielen eine entscheidende Rolle bei der deutlichen Reduzierung der Verzerrung bei Modellvorhersagen.
Diese Konsistenz bestätigte die Robustheit dieser Methoden über verschiedene Datensätze hinweg. Bemerkenswerterweise geht keine der Methoden Kompromisse bei der Genauigkeit ein, während beide die Fairness verbessern, was ihre Wirksamkeit bei der Handhabung der Kompromisse zeigt, die typischerweise mit prädiktiver Diese Ergebnisse zeigen die starken Interpolationsfähigkeiten, die PC-SMOTE und PC-ADASYN innewohnen. Diese Methoden verteilen die Stichproben effektiv innerhalb des Merkmalsraums neu, insbesondere indem sie die Stichproben in unterprivilegierten Regionen von negativen zu positiveren Entscheidungsgrenzen verschieben . Solche Anpassungen sind entscheidend, um verzerrte Ergebnisse abzumildern und die Gerechtigkeit in automatisierten Entscheidungsprozessen zu fördern. Die Makro-F1 sinkt in beiden Modellen im Vergleich zur Basislinie, da eine höhere Anzahl von Stichproben erforderlich ist, damit das Modell bei der Generalisierung besser abschneidet, was dieser Datensatz nicht unterstützt. Insbesondere hat der Datensatz 700 Stichproben für Klasse 0 und 300 Stichproben für Klasse 1, was bedeutet, dass der Testsatz

Machine Translated by Google

hat nur 30 Proben für Klasse 1. Diese kleine Anzahl von Proben führte dazu, dass beide Modelle in Klasse 1 Rückruf gegen Präzision eintauschten. Insbesondere sahen wir einen niedrigen Rückruf für Klasse 1, was letztendlich zu einem niedrigen F1-Score für Klasse 1 führte, und da Makro F1 die beiden F1-Scores mittelt und sie gleich behandelt, wirkt sich dies auf die Leistung beider Modelle in Makro F1 aus. Insgesamt ergeben die beiden Modelle ein faireres Modell mit guter Genauigkeit im Vergleich zur Basislinie und den beiden anderen Stichprobenmethoden.

Die Auswertung des COMPAS-Datensatzes bestätigt die Wirksamkeit unserer Stichprobenmethoden weiter . Die deutlichen Muster, die sich herauskristallisieren, stimmen mit denen überein, die in den Datensätzen „Erwachseneneinkommen“ und „Deutscher Kredit“ beobachtet wurden, was die Robustheit unserer Ergebnisse unterstreicht. Insbesondere haben Oversampling- und proportionale Stichprobenverfahren erhebliche Verbesserungen bei der ausgeglichenen Quotendifferenz (EOD) und der Genauigkeit gezeigt, während Oversampling auch bei SP deutlich verbessert. Diese Verbesserung ist wahrscheinlich auf die einzigartige Zusammensetzung und Ausgewogenheit innerhalb des COMPAS-Datensatzes zurückzuführen, im Gegensatz zu den anderen Datensätzen, in denen die Klassen unausgewogen sind. Der Erfolg von Oversampling und proportionalem Sampling in diesem Zusammenhang ist auf die ausgewogene Natur des Datensatzes zurückzuführen, die eine wiederholte Duplizierung vorhandener Zeilen (von diesen Methoden verwendete Stichprobenverfahren) ermöglicht, um den Datensatz zu verbessern, ohne eine signifikante Schiefe zugunsten einer bestimmten Klasse einzuführen. Diese Methode erweitert effektiv die Darstellung aller Klassen und der geschützten Attribute in ausgewogener Form, wodurch diese Techniken besonders effektiv für Datensätze sind, bei denen die Merkmalsdomänen gleichermaßen zu Vorhersagen beitragen und bei denen die anfänglichen Klassenverteilungen nicht unter einem starken Ungleichgewicht leiden . Dies kann weiter anhand ihres Makros F1 verifiziert werden, da keines der Modelle Präzision gegen Rückruf eintauscht . Die Verbesserung von SP beim Oversampling kann auf die höhere Anzahl von Samples beim Oversampling im Vergleich zum proportionalen Sampling zurückgeführt werden. In Bezug auf PC-SMOTE und PC-ADASYN zeigen diese Algorithmen eine Verbesserung gegenüber der Baseline sowohl bei der Genauigkeit als auch beim EOD. Diese Trends folgen denen der vorherigen Ergebnisse. Ein bemerkenswerter Aspekt dieser Ergebnisse ist der große Rückgang von SP, der auf unser neues Label zurückgeführt werden kann, das generiert wurde, um den Datensatz in Richtung der negativen Klasse zu verzerre

6.2. Einfluss der Baumtiefe auf Fairness und Genauigkeit

In dieser Studie wurde auch der Einfluss der Tiefe des Entscheidungsbaums auf die Modellleistung untersucht, insbesondere wurde untersucht, wie Variationen der Baumtiefe die Genauigkeit und EOD beeinflussen. Das Verständnis des Tiefeneffekts ist wichtig, da es Einblick in die Effekte von Unteranpassung bis Überanpassung bietet und hilft, den optimalen Komplexitätsgrad zu ermitteln, bei dem sowohl Genauigkeit als auch Fairness maximiert werden. Zunächst durfte der Entscheidungsbaum ohne Einschränkungen bis zu seiner vollen Tiefe wachsen, die im Durchschnitt etwa 30 Zweige umfasste. Der Baum wurde dann visuell untersucht, um die maximale Tiefe ohne die nicht-aufteilenden Zweige zu ermitteln. Um die Effekte der Baumtiefe systematisch zu analysieren, durfte die maximale Tiefe der Bäume zwischen 1 und 30 variieren. Jede Tiefengrenze wurde mithilfe einer zehnfachen Kreuzvalidierung ausgewertet, um die Robustheit und Generalisierbarkeit der Erkenntnisse sicherzustellen.

Für jede Konfiguration der Baumtiefe wurden Genauigkeit und EOD im Testset gemessen . Zusätzlich wurden 95%-Konfidenzintervalle für die Metriken über die zehn Faltungen hinweg berechnet. Diese statistische Analyse hob die Tiefe hervor, in der der Entscheidungsbaum den Kompromiss zwischen Genauigkeit und Fairness ausbalancierte, während gleichzeitig der zugrunde liegende statistische Bias-Varianz-Kompromiss berücksichtigt wurde. Auf diese Weise war es möglich, den „Sweet Spot“ zu ermitteln – einen heiklen Punkt, an dem der Entscheidungsbaum eine hohe Vorhersagegenauigkeit beibehält, ohne Kompromisse bei der Fairness einzugehen, und so dem oft zitierten Komp

In den Abbildungen 3–7 werden die Diagramme für Genauigkeit und EOD gegenüber der maximalen Tiefe für jede der fünf Stichprobenmethoden im Datensatz „Erwachseneneinkommen“ dargestellt.

Machine Translated by Google

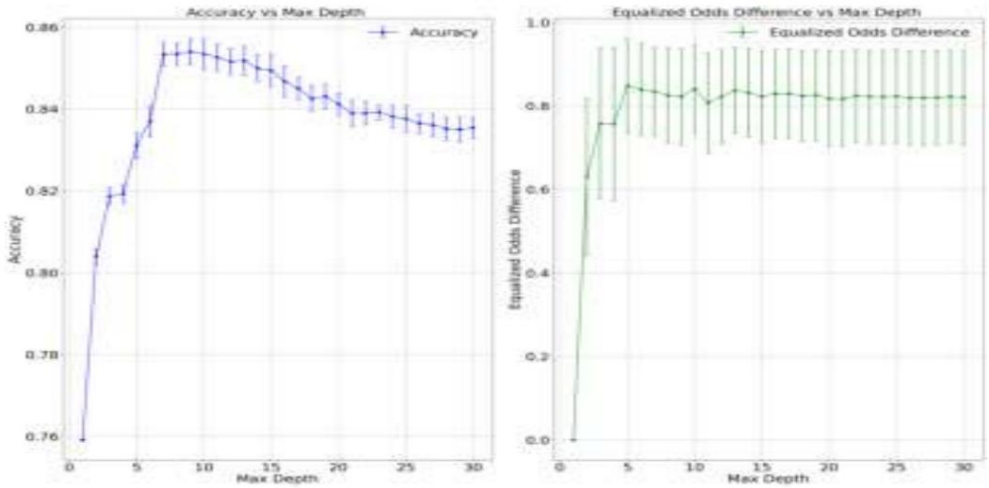


Abbildung 3. Diagramme des Erwachseneneinkommens ohne Stichprobennahme, die Genauigkeit und EOD mit 95 %-Konfidenzintervallen gegenüber der maximalen Baumtiefe im Bereich von 1 bis 30 zeigen.

Basierend auf Ergebnissen wie denen in den Abbildungen 3 und 4 ist sowohl bei der No-Sampling- als auch bei der Oversampling-Methode eine bemerkenswerte anfängliche Genauigkeitssteigerung mit zunehmender maximaler Tiefe zu verzeichnen. Beide Methoden weisen jedoch ab einer Tiefe von 10 einen Genauigkeitsrückgang auf, was auf den Beginn einer Überanpassung hindeutet. Entsprechend nimmt die EOD mit zunehmender Tiefe bis etwa Tiefe 10 stark ab, jenseits derer sie sich stabilisiert. Dieses Muster weist darauf hin, dass tiefere Bäume zwar zunächst die Fairness verbessern, schließlich aber einen Schwellenwert erreichen, jenseits dessen keine weiteren Verbesserungen mehr zu beobachten sind. Wenn man sich daran erinnert, dass das Fairnessziel darin bestand, die EOD zu minimieren, ist eine wichtige Beobachtung, dass die Festlegung der maximalen Tiefe zwischen drei und fünf ein optimales Gleichgewicht zwischen

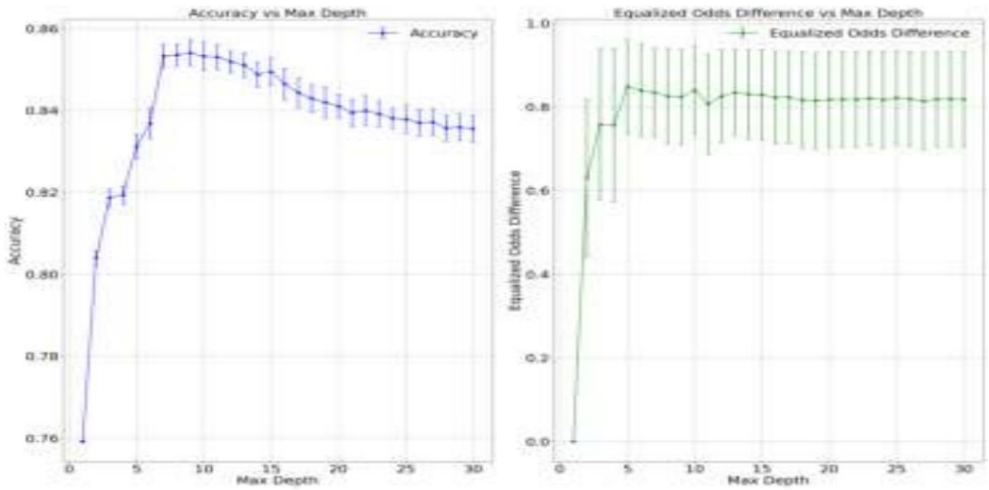


Abbildung 4. Diagramme des Erwachseneneinkommens unter Verwendung von Oversampling, die Genauigkeit und EOD mit 95 %-Konfidenzintervallen gegenüber der maximalen Baumtiefe im Bereich von 1 bis 30 zeigen.

Betrachtet man die in Abbildung 5 gezeigten Ergebnisse, so erhöht sich die Genauigkeit bei der proportionalen Stichprobenmethode kontinuierlich mit zunehmender Baumtiefe und erreicht bei einer Tiefe von etwa 26 ihren Höhepunkt. Umgekehrt steigt der EOD zunächst an, bevor er abnimmt und sich bei einer Tiefe von etwa 15 stabilisiert. Die in der EOD-Metrik beobachteten breiten Konfidenzintervalle deuten auf eine erhebliche Variabilität der Fairnessergebnisse hin. Dieses Ergebnis unterstreicht, wie wichtig es ist, eine Tiefe auszuwählen, die die Variabilität der Fairness minimiert und gleichzeitig die Genauigkeit

Machine Translated by Google

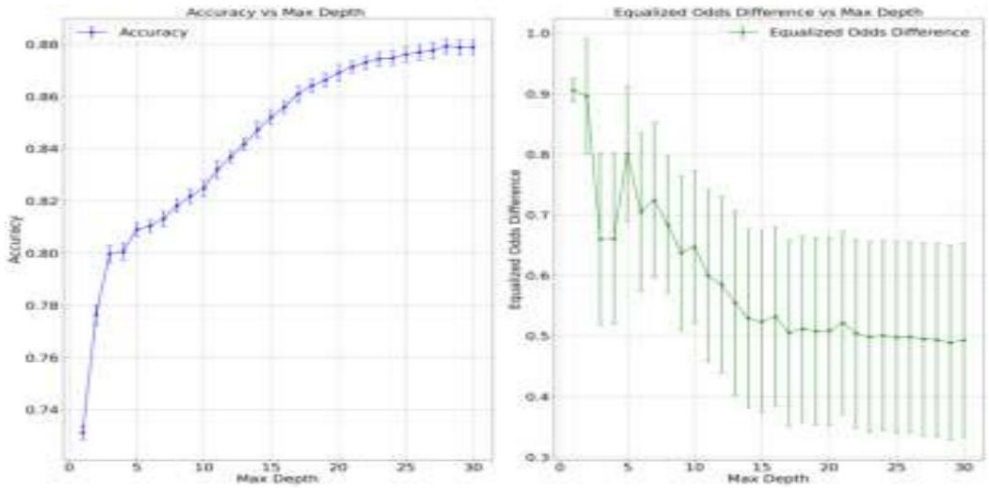


Abbildung 5. Diagramme des Einkommens Erwachsener unter Verwendung proportionaler Stichprobennahme, die Genauigkeit und EOD mit 95 % -Konfidenzintervallen gegenüber der maximalen Baumtiefe im Bereich von 1 bis 30 zeigen.

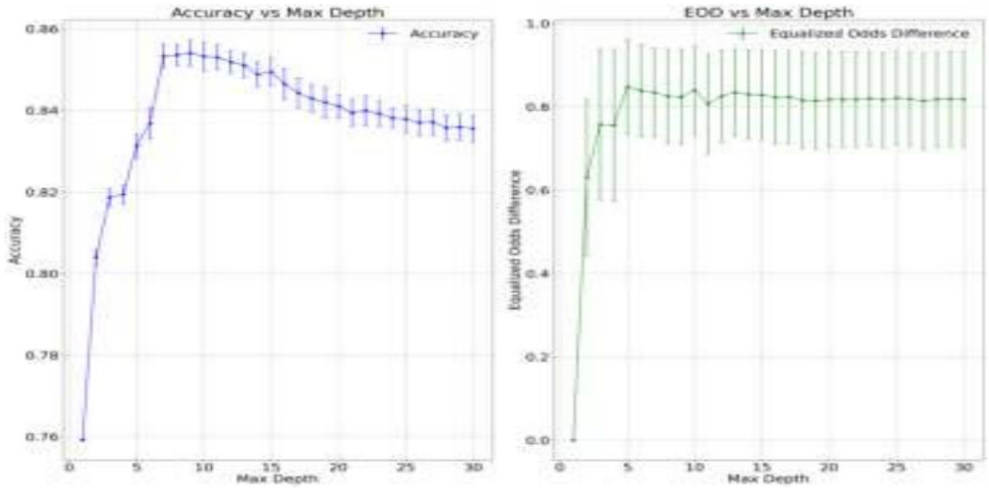


Abbildung 6. Diagramme des Erwachseneneinkommens unter Verwendung von PC-SMOTE, die Genauigkeit und EOD mit 95 %-Konfidenzintervallen gegenüber der maximalen Baumtiefe im Bereich von 1 bis 30 zeigen.

Abbildung 6 und 7, die die Ergebnisse mit PC-SMOTE bzw. PC-ADASYN darstellen, zeigen leichte Abwärtstrends bei der Genauigkeit, die sich kurzzeitig verbessern, bevor sie wieder abfallen – ein Muster, das auf Überanpassung bei größeren Tiefen hinweist. Die EOD-Metriken für diese Methoden zeigen anfängliche Stabilität bei geringeren Tiefen, Anstieg bei mittleren Tiefen und Rückgang, was auf komplexe Wechselwirkungen zwischen der Generierung synthetischer Proben und der Abgrenzung der Entscheidungsgrenzen hindeutet . **Angesichts dieser Beobachtungen wurde für unsere Experimente eine maximale Tiefe von 3 gewählt, da dies einen „Sweet Spot“ darstellt, bei dem sowo**

Angesichts dieser Ergebnisse ist eine Schlussfolgerung, den oft angenommenen Kompromiss zwischen Genauigkeit und Fairness in Frage zu stellen, indem gezeigt wird, dass unsere PC-ADASYN-Methode die Baselines in allen drei Datensätzen sowohl in Bezug auf Genauigkeit als auch Fairness durchgängig übertrifft. Dieses Ergebnis ist bedeutsam, da es darauf hindeutet, dass es möglich ist, die Modellfairness zu verbessern, ohne die Genauigkeit mit geeigneten Stichprobenmethoden und Modelloptimierung zu beeinträchtigen. Unsere Analyse zeigt jedoch auch Szenarien auf, in denen Anpassungen der Modellkomplexität, insbesondere der maximalen Tiefe von Entscheidungsbäumen, die Genauigkeit auf Kosten der Fairness verbessern können, wie durch Erhöhungen der Equalized Odds Difference (EOD) angezeigt wird. Es wird jedoch erwartet, dass die Kopplung von Stichprobenmethoden mit Inprocessing-Methoden wie fairnessbasierter Regularisierung diese Effekte ausgleichen kann. Diese Entscheidungen unterstreichen den Ermessensspielraum der Forscher bei der Ausbala Die Menge der Stichprobe und die zeitliche Komplexität sind wie bei jeder anderen Stichprobenmethode.

Machine Translated by Google

Mit zunehmender Stichprobenmenge erhöht sich auch die Zeitkomplexität, doch insgesamt weisen die Stichprobenmethoden aufgrund der gleichen Funktionalität die gleiche Zeitkomplexität auf wie die ihnen zugrunde liegenden Algorithmen .

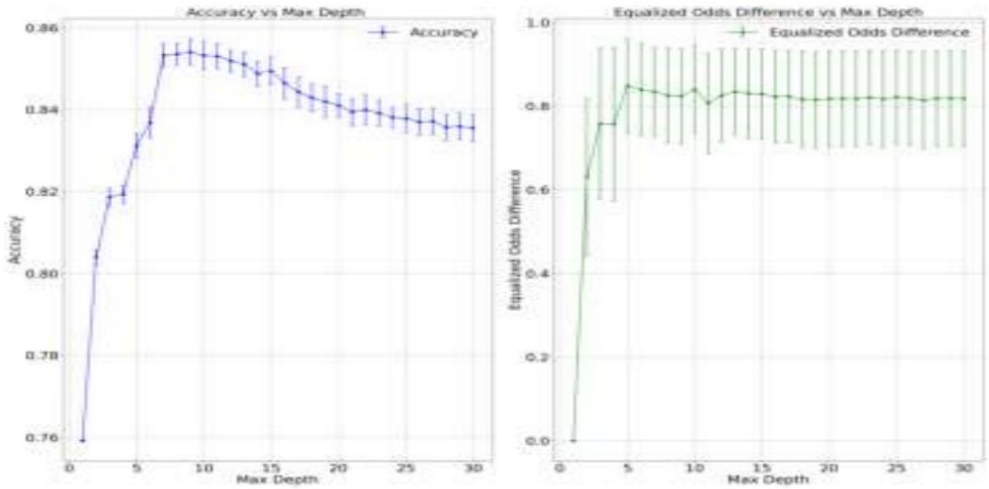


Abbildung 7. Diagramme des Erwachseneneinkommens unter Verwendung von PC-ADASYN, die Genauigkeit und EOD mit 95 %-Konfidenzintervallen gegenüber der maximalen Baumtiefe im Bereich von 1 bis 30 zeigen.

Darüber hinaus unterstreichen unsere Ergebnisse die Komplexität der gleichzeitigen Optimierung mehrerer Fairnessmetriken. So führten beispielsweise Bemühungen zur Verbesserung der statistischen Parität (SP) durch Bevorzugung positiverer Vorhersagen für jede geschützte Gruppe im COMPAS-Datensatz zu einer unbeabsichtigten Verringerung negativer Vorhersagen. Diese Verschiebung wirkte sich negativ auf die False Positive Rate (FPR), eine Komponente von EOD, aus und verschlechterte somit die EOD-Metrik, wenn sich die SP verbesserte. Dieses Phänomen veranschaulicht die inhärenten mathematischen Spannungen zwischen Fairnessmetriken, bei denen die Optimierung einer Metrik sich nachteilig auf eine andere auswirken kann. Der COMPAS-Datensatz mit seiner nahezu ausgeglichenen Klassenverteilung liefert ein konkretes Beispiel dafür, wie Datensatzmerkmale das Verhalten von Fairnessmetriken beeinflussen können. Die Optimierung der SP in diesem Zusammenhang impliziert eine verzerrte Messung der Fairness, insbesondere wenn inhärente Unterschiede zwischen Gruppen in geschützten Attributen bestehen. Dies wird durch die Literatur unterstützt, die darauf hinweist, dass die SP Gruppenunterschiede möglicherweise nicht ausreichend berücksichtigt, was möglicherweise zu irreführenden Schlussfolgerungen über di

7. Einschränkungen und zukünftige Arbeit

Aufgrund der Natur dieser Studie ist es nicht möglich, alle Fragen rund um Fairness und den sogenannten Kompromiss zwischen Fairness und Leistung zu behandeln. Daher sind der hier berichteten Arbeit Grenzen gesetzt. Dennoch hoffen und beabsichtigen wir, mit der hier berichteten Arbeit zusätzliche Forschungsansätze in diesem wichtigen Bereich aufzuzeigen.

Eine Einschränkung dieser Studie besteht darin, dass unsere Stichprobenmethode nicht speziell für die Optimierung beliebiger Fairnessmetriken entwickelt wurde. Anders ausgedrückt: Da zwischen den verfügbaren Fairnessmetriken häufig inhärente Kompromisse bestehen, wurde entschieden, sich auf einen metrikenunabhängigen Ansatz zu konzentrieren, da die Ergebnisse bei anderen Metriken anders ausgefallen sein könnten . Dies ist auch einer der Gründe, warum wir zwischen EOD und SP unterschiedliche Verhaltensweisen festgestellt haben.

Darüber hinaus wird anerkannt, dass die zugrunde liegende ML-Methode zwar für die vorgeschlagene Methode nicht relevant sein sollte, dies jedoch nicht tatsächlich getestet wurde. Daher wird diese Forschung in Zukunft erweitert, indem die Auswirkungen anderer ML-Algorithmen wie logistische Regression, Fuzzy ID3, K-Nearest Neighbor und Ensemblemethoden wie Random Forests oder Gradient-Boosted Trees berücksichtigt werden, um die Generalisierbarkeit unse Der Zweck einer solchen Studie wäre es, zu überprüfen, ob unsere Methoden unabhängig vom verwendeten ML-Algorithmus sind. Darüber hinaus würde dies dazu beitragen zu validieren, ob die beobachteten Verbesserungen bei Fairness und Genauigkeit modellspezifisch sind oder universell angewendet werden können.

Machine Translated by Google

Darüber hinaus wird anerkannt, dass nur drei unterschiedliche Datensätze berücksichtigt wurden – Datensätze , die in diesem Bereich umfassend untersucht wurden. Dies gibt Anlass zur Sorge, dass Methoden auf diese Daten zugeschnitten werden, anstatt sich mit dem umfassenderen Thema der Fairness im ML zu befassen. Um dieses Problem anzugehen, sind Experimente mit größeren und vielfältigeren Datensätzen geplant, um tiefere Einblicke in die Skalierbarkeit und Robustheit unserer Techniken zu erhalten. Ein weiterer Bereich für zukünftige Arbeiten ist die Verfeinerung unseres Multikategorien-Sampling-Ansatzes durch die Einbeziehung feinkörnigerer Unterteilungen geschützter Kategorien, wodurch möglicherweise subtilere Verzerrungen aufgedeckt und ein differenzierteres Verständnis von

Abschließend sei angemerkt, dass alternative Methoden zur Minderung von Verzerrungen vorgeschlagen wurden , die in dieser Arbeit jedoch überhaupt nicht untersucht wurden. Zukünftige Arbeiten würden Vergleiche mit weiteren Stichprobenstrategien erfordern. Es wird ein direkterer Vergleich der vorgeschlagenen Methoden mit In-Processing- und Post-Processing-Methoden durchgeführt. Beispielsweise wird die Einbeziehung von In-Processing-Methoden wie Regularisierung [22] oder einer Post-Processing- Methode wie dem Randomized Threshold Optimizer [64] als mögliche Mittel zur Erzielung weiterer Verbesserungen sowohl der Fairness als auch der Leistung untersucht .

8. Schlussfolgerungen

In dieser Studie wurde das Problem der Verzerrung bei ML-Vorhersagen untersucht und eine Methode entwickelt, die auf der Kombination geschützter Variablen in einer neuen Multikategorie basiert. Der Schwerpunkt lag insbesondere auf der in der Literatur aufgeworfenen Frage nach einem Kompromiss zwischen Verzerrung und Leistung und der Suche nach einer Methode zur Abmilderung dieses Kompromisses. Der vorgeschlagene neue Multikategorienansatz spiegelt die vielschichtige Identität von Personen wider und berücksichtigt das komplexe Zusammenspiel von Attributen, die reale Szenarien definieren. Angesichts des inhärenten Ungleichgewichts in dieser Multikategorie wurden vier Stichprobenmethoden entwickelt, die auf diese komplexen Kategorisierungen zugeschnitten sind, anstatt auf traditionelle Klassenbezeichnungen. Zum Zweck der Anwendung eines Basisklassifikators wurden Entscheidungsbäume trainiert und die Wirksamkeit dieser Methoden anhand von drei Datensätzen bewertet, die häufig in Fairnessstudien verwendet werden. Die Leistung der Methoden wurde mit Basismethoden ohne Stichproben verglichen, wobei Genauigkeit, Makro- F1, Equalized Odds Difference (EOD) und

Die Ergebnisse der Experimente zeigen, dass zwei der neu entwickelten Sampling- Techniken – PC-SMOTE und PC-ADASYN – die Fairness erfolgreich verbessern, ohne die Genauigkeit zu beeinträchtigen. Bemerkenswerterweise verbesserten diese Methoden in einigen Fällen auch die Genauigkeit und lieferten damit Beweise, die den gängigen Behauptungen eines Kompromisses zwischen Fairnes Eine weitere Analyse der Auswirkungen der maximalen Baumtiefe auf die Modellleistung ergab , dass eine zunehmende Tiefe zwar zunächst die Genauigkeit steigert, letztendlich jedoch zu einer Verschlechterung führt. Umgekehrt wirkt sich eine zunehmende Tiefe negativ auf die Fairness aus, was die Herausforderung unterstreicht, Komplexität und Gerechtigkeit in Einklang zu bringen. Es wurden jedoch optimale Baumtiefen identifiziert, die gleichzeitig Genauigkeit und EOD verbessern, was die Möglichkeit unterstreicht, Gerechtigkeit zu erreichen, ohne die Leistung zu beeinträchtigen.

Letztlich stellen diese Erkenntnisse die vorherrschenden Vorstellungen eines impliziten Kompromisses zwischen **Leistung und Fairness in der Forschung zur Vermeidung von Vorurteilen in Frage und legen nahe, dass sorgfältig konzipierte Strategien zur Vermeidung von Vorurteilen diesen Kompromiss umgehen können. Unser Ansatz schafft einen neuen Präzedenzfall für die Entwicklung gerechterer Vorhersagealgorithmen, indem er neu definiert, wie geschützte Attribute beim Modelltraining genutzt werden.**

Beiträge der Autoren: Konzeptualisierung, GP und JS; formale Analyse, GP und JS; Untersuchung, GP; Methodik, GP; Software, GP; Überwachung, JS; Validierung, GP und JS; Visualisierung, GP; Schreiben – Originalentwurf, GP; Schreiben – Überprüfung und Bearbeitung, JS. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Forschung erhielt keine externe Finanzierung.

Datenverfügbarkeitserklärung: Die in dieser Forschungsarbeit implementierten Datensätze und Codes wurden auf https://github.com/horlahsunbo/New-folder hochgeladen. **(abgerufen am 10. Juni 2024).**

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Machine Translated by Google

Verweise

1. **Maina, IW; Belton, TD; Ginzberg, S.; Singh, A.; Johnson, TJ Ein Jahrzehnt der Erforschung impliziter rassistischer/ethnischer Vorurteile im Gesundheitswesen Anbieter unter Verwendung des impliziten Assoziationstests.** Soc. Sci. Med. **2018,** 199, **219–229.** [CrossRef]
2. **Salimi, B.; Rodriguez, L.; Howe, B.; Suciu, D. Kausale Datenbankreparatur für algorithmische Fairness. In Proceedings der 2019 International Conference on Management of Data, Amsterdam, Niederlande, 30. Juni–5. Juli 2019; S. 793–810.**
3. **Kordzadeh, N.; Ghasemaghaei, M. Algorithmische Verzerrung: Überprüfung, Synthese und zukünftige Forschungsrichtungen.** Eur. J. Inf. Syst. **2022,** 31, **388–409.** [CrossRef]
4. **Pessach, D.; Shmueli, E. Eine Übersicht über Fairness im maschinellen Lernen.** ACM Comput. Surv. **2022,** 55, **1–44.** [CrossRef]
5. **Aghaei, S.; Azizi, MJ; Vayanos, P. Lernen optimaler und fairer Entscheidungsbäume für diskriminierungsfreie Entscheidungsfindung.** Proc. AAA Conf. Artif. Intell. **2019,** 33, **1418–1426.** [CrossRef]
6. **Kamiran, F.; Calders, T. Datenvorverarbeitungstechniken für eine Klassifizierung ohne Diskriminierung.** Knowl. Inf. Syst. **2012,** 33, **1–33.** [Querverweis]
7. **Calmon, F.; Wei, D.; Vinzamuri, B.; Ramamurthy, KN; Varshney, KR Optimierte Vorverarbeitung zur Diskriminierungsprävention. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4.–9. Dezember 2017.**
8. **Shahbazi, N.; Lin, Y.; Asudeh, A.; Jagadish, HV Darstellungsverzerrung in Daten: Eine Untersuchung zu Identifizierungs- und Lösungstechniken.** ACM Comput. Surv. **2023,** 55, **1–39.** [CrossRef]
9. **Chen, Z.; Zhang, JM; Sarro, F.; Harman, M. Eine umfassende empirische Studie zu Methoden zur Bias-Minderung beim maschinellen Lernen Klassifikatoren.** ACM Trans. Softw. Eng. Methodol. **2023,** 32, **1–30.** [CrossRef]
10. **Perzynski, A.; Berg, KA; Thomas, C.; Cemballi, A.; Smith, T.; Shick, S.; Gunzler, D.; Sehgal, AR Rassistische Diskriminierung und wirtschaftliche Faktoren bei der Redlining-Politik in Ohio.** Bois Rev. Soc. Sci. Res. Race **2023,** 20, **293–309.** [CrossRef]
11. **Steil, JP; Albright, L.; Rugh, JS; Massey, DS Die soziale Struktur der Hypothekendiskriminierung.** Hous. Stud. **2018,** 33, **759–776.** [Querverweis]
12. **Salgado, JF; Moscoso, S.; García-Izquierdo, AL; Anderson, NR** Gestaltung integrativer Arbeitsplätze durch sozialen Dialog; **Springer: Berlin/Heidelberg, Deutschland, 2017.**
13. **Leavy, S.; Meaney, G.; Wade, K.; Greene, D. Abschwächung geschlechtsspezifischer Voreingenommenheit in Datensätzen des maschinellen Lernens. In Proceedings of the Bias and Social Aspects in Search and Recommendation: Erster internationaler Workshop, BIAS 2020, Lissabon, Portugal, 14. April 2020; S. 12–26.**
14. **Hort, M.; Chen, Z.; Zhang, JM; Harman, M.; Sarro, F. Bias-Minderung für maschinelle Lernklassifikatoren: Eine umfassende Untersuchung.** ACM J. Responsible Comput. **2023, 1, 1–52.** [CrossRef]
15. **Fahse, T.; Huber, V.; van Giffen, B. Umgang mit Voreingenommenheit in Projekten zum maschinellen Lernen. In** Innovation Through Information Systems: Band II: Eine Sammlung neuester Forschung zu Technologiefragen; **Springer: Berlin/Heidelberg, Deutschland, 2021; S. 94–109.**
16. **Zhang, Z.; Neill, DB Identifizierung signifikanter prädiktiver Verzerrungen in Klassifikatoren.** arXiv **2017, arXiv:1611.08292.** [CrossRef]
17. **Zheng, Z.; Cai, Y.; Li, Y. Oversampling-Methode für unausgewogene Klassifizierung.** Comput. Inform. **2015,** 34, **1017–1037.**
18. **Chawla, NV; Bowyer, KW; Hall, LO; Kegelmeyer, WP SMOTE: Synthetische Minderheiten-Oversampling-Technik.** J. Artif. Intell. Res. **2002,** 16, **321–357.** [CrossRef]
19. **He, H.; Bai, Y.; Garcia, EA; Li, S. ADASYN: Adaptiver synthetischer Sampling-Ansatz für unausgewogenes Lernen. In Proceedings der 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hongkong, China, 1.–8. Juni 2008; S. 1322–1328.**
20. **Janssen, P.; Sadowski, BM Bias in Algorithms: Über den Kompromiss zwischen Genauigkeit und Fairness. In Proceedings der 23. Zweijahreskonferenz der International Telecommunications Society, Göteborg, Schweden, 21.–23. Juni 2021.**
21. **Hardt, M.; Price, E.; Srebro, N. Chancengleichheit beim überwachten Lernen. In Proceedings of the Advances in Neural Informationsverarbeitungssysteme, Barcelona, Spanien, 5.–10. Dezember 2016.**
22. **Kamishima, T.; Akaho, S.; Asoh, H.; Sakuma, J. Fairnessbewusster Klassifikator mit Vorurteilsbeseitiger-Regularisierer. In Proceedings of the Machine Learning and Knowledge Discovery in Databases: Europäische Konferenz, ECML PKDD 2012, Bristol, Großbritannien, 24.–28. September 2012; S. 35–50.**
23. **Calders, T.; Žliobaitė, I. Warum unvoreingenommene Rechenprozesse zu diskriminierenden Entscheidungsverfahren führen können. In** Diskriminierung und Datenschutz in der Informationsgesellschaft: Data Mining und Profiling in großen Datenbanken; **Springer: Berlin/Heidelberg, Deutschland, 2013; S. 23–33.**
24. **Burt, A. Wie man Diskriminierung in der KI bekämpft. Harvard Business Review. 2020 Online verfügbar: https://hbr.org/2020/08/how-to- fight-discrimination-in-ai (abgerufen am 12. Juli 2024).**
25. **Siegler, A.; Admussen, W. Entdeckung rassistischer Diskriminierung durch die Polizei.** Northwestern Univ. Law Rev. **2021,** 115, **987–1054.**
26. **Grabowicz, P.; Perello, N.; Mishra, A. Wie trainiert man Modelle, die keine Diskriminierung propagieren? Equate and Machine Learning Blog, University of Massachusets, Amherst. 2022. Online verfügbar: https://groups.cs.umass.edu/equate-ml/2022/04/07/how- to-train-models-that-do-not-propagate-discrimination/ (abgerufen am 12. Juli 2024).**
27. **Khani, F.; Liang, P. Von Diskriminierung im maschinellen Lernen zu Diskriminierung im Recht, Teil 1: Ungleiche Behandlung. Stanford AI Lab Blog. 2022. Online verfügbar: https://ai.stanford.edu/blog/discrimination_in_ML_and_law/ (abgerufen am 12. Juli 2024).**
28. **Shahriari, K.; Shahriari, M. IEEE-Standardüberprüfung – Ethisch ausgerichtetes Design: Eine Vision zur Priorisierung des menschlichen Wohlbefindens mit künstlicher Intelligenz und autonomen Systemen. In Proceedings der IEEE Canada International Humanitarian Technology Conference (IHTC), Toronto, ON, Kanada, 21.–22. Juli 2017; S. 197–201.**

Machine Translated by Google

29. Europäische Kommission; Generaldirektion Kommunikationsnetze, Inhalte und Technologien. Ethische Leitlinien für vertrauenswürdige KI. Amt für Veröffentlichungen. 2019. Online verfügbar: https://op.europa.eu/en/publication-detail/-/publication/d39885 69-0434-11ea-8c1f-01aa75ed71a1/Sprache-de/Format-PDF/Quelle-337437547 (abgerufen am 10. Juni 2024)

30. Ebers, M.; Hoch, VRS; Rosenkranz, F.; Ruschemeier, H.; Steinrötter, B. Der Vorschlag der Europäischen Kommission für ein Gesetz zur künstlichen Intelligenz – Eine kritische Bewertung durch Mitglieder der Robotics and AI Law Society (RAILS). J 2021, 4, 589–603. [CrossRef]

31. Hajian, S.; Domingo-Ferrer, J. Eine Methodologie zur direkten und indirekten Diskriminierungsprävention im Data Mining. IEEE Trans. Knowl. Data Eng. **2012,** 25, **1445–1459. [CrossRef]**

32. Fish, B.; Kun, J.; Lelkes, Á.D. Ein auf Vertrauen basierender Ansatz zum Ausgleich von Fairness und Genauigkeit. In Proceedings of the 2016 SIAM Internationale Konferenz zum Data Mining, Miami, FL, USA, 5.–7. Mai 2016; S. 144–152.

33. Hajian, S. Gleichzeitige Diskriminierungsprävention und Datenschutz bei der Veröffentlichung und dem Mining von Daten. arXiv **2013, arXiv:1306.6805. [CrossRef]**

34. Sondeck, LP; Laurent, M.; Frey, V. Die semantische Diskriminationsratenmetrik für Datenschutzmessungen, die den Nutzen von t-Nähe gegenüber l-Diversität in Frage stellt. In Proceedings der 14. Internationalen Konferenz für Sicherheit und Kryptographie, Madrid, Spanien, 24.–26. Juli 2017; Band 6, S. 285–294.

35. Ruggieri, S. Verwendung von T-Closeness-Anonymität zur Kontrolle von Nichtdiskriminierung. Trans. Data Priv. **2014, 2, 99–129.**

36. Romano, Y.; Bates, S.; Candes, E. Erreichen ausgeglichener Chancen durch erneutes Sampling sensibler Attribute. In Proceedings of the Advances in Neural Information Processing Systems, Online, 6.–12. Dezember 2020; S. 361–371.

37. Peng, K.; Chakraborty, J.; Menzies, T. Fairmask: Bessere Fairness durch modellbasierte Neugewichtung geschützter Attribute. IEEE Trans. Softw. Ing. **2022,** 49, **2426–2439. [CrossRef]**

38. Dhar, P.; Gleason, J.; Roy, A.; Castillo, CD; Chellappa, R. PASS: Geschütztes Attributunterdrückungssystem zur Minderung von Verzerrungen bei der Gesichtserkennung. In Proceedings der IEEE/CVF International Conference on Computer Vision, Montreal, QC, Kanada, 10.–17. Oktober 2021; S. 15087–15096.

39. Krasanakis, E.; Spyromitros-Xioufis, E.; Papadopoulos, S.; Kompatsiaris, Y. Adaptive sensible Neugewichtung zur Minderung von Verzerrungen bei fairnessbewusster Klassifizierung. In Proceedings der World Wide Web Conference, Lyon, Frankreich, 23.–27. April 2018; S. 853–8

40. Ranÿci´c, S.; Radovanovi´c, S.; Delibaši´c, B. Untersuchung von Oversampling-Techniken für faire Modelle des maschinellen Lernens. In Proceedings of the Decision Support Systems XI: Decision Support Systems, Analytics and Technologies in Response to Global Crisis Management: 7. Internationale Konferenz zu Decision Support System Technology, ICDSST 2021, Loughborough, Großbritannien, 26.–28. Mai 2021; Springer: Berlin/Heidelberg, Deutschland, 2021; S. 110–123.

41. Yan, S.; te Kao, H.; Ferrara, E. Fairer Klassenausgleich: Verbesserung der Modellfairness ohne Berücksichtigung sensibler Attribute. In Proceedings der 29. ACM International Conference on Information & Knowledge Management, virtuelle Veranstaltung, Irland, 19.–23. Oktober 2010; S. 1715–1724.

42. HuZhang, B.; Lemoine, B.; Mitchell, M. Unerwünschte Vorurteile durch kontroverses Lernen abmildern. In Proceedings der AAAI/ACM-Konferenz zu KI, Ethik und Gesellschaft, New Orleans, LA, USA, 2.–3. Februar 2018; S. 335–340.

43. Celis, LE; Huang, L.; Keswani, V.; Vishno, NK Klassifizierung mit Fairnessbeschränkungen: Ein Metaalgorithmus mit nachweisbaren Garantien. In Proceedings of the Conference on Fairness, Accountability, and Transparency, Atlanta, GA, USA, 29.–31. Januar 2019; S. 319–328.

44. Zafar, MB; Valera, I.; Rodriguez, MG; Gummadi, KP Fairness jenseits von ungleicher Behandlung und ungleicher Auswirkung: Klassifizierung lernen ohne ungleiche Misshandlung. In Proceedings der 26. International Conference on World Wide Web, Perth, Australien, 3.–7. April 2017; S. 1171–1180.

45. Agarwal, A.; Beygelzimer, A.; Dudík, M.; Langford, J.; Wallach, H. Ein Reduktionsansatz zur fairen Klassifizierung. In Proceedings of die Internationale Konferenz für maschinelles Lernen, Stockholm, Schweden, 10.–15. Juli 2018; S. 60–69.

46. Lowy, A.; Baharlouei, S.; Pavan, R.; Razaviyayn, M.; Beirami, A. Ein stochastisches Optimierungsframework zur fairen Risikominimierung. Trans. Mach. Learn. Res. **2022. [CrossRef]**

47. Spinelli, I.; Scardapane, S.; Hussain, A.; Uncini, A. Fairdrop: Biased Edge Dropout zur Verbesserung der Fairness beim Lernen von Graphdarstellungen . IEEE Trans. Artif. Intell. **2021, 3, 344–354. [CrossRef]**

48. Hort, M.; Zhang, JM; Sarro, F.; Harman, M. Fairea: Ein Modellverhaltensmutationsansatz zum Benchmarking von Methoden zur Minderung von Verzerrungen. In Proceedings des 29. gemeinsamen Treffens der ACM zur European Software Engineering Conference und zum Symposium über die Grundlagen des Software Engineering, Athen, Griechenland, 23.–28. August 2012; S. 994–1006.

49. Bhaskaruni, D.; Hu, H.; Lan, C. Verbesserung der Vorhersagefairness durch Modellensemble. In Proceedings der 31. IEEE-Internationalkonferenz zu Tools mit künstlicher Intelligenz (ICTAI), Portland, OR, USA, 4.–6. November 2019; S. 1810–1814.

50. Iosifidis, V.; Fetahu, B.; Ntoutsi, E. Fae: Ein fairnessbewusstes Ensemble-Framework. In Proceedings der IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9.–12. Dezember 2019; S. 1375–1380.

51. Dwork, C.; Hardt, M.; Pitassi, T.; Reingold, O.; Zemel, R. Fairness durch Bewusstsein. In Proceedings der 3. Innovations in Konferenz für Theoretische Informatik, Cambridge, MA, USA, 8.–10. Januar 2012; S. 214–226.

52. Pleiss, G.; Raghavan, M.; Wu, F.; Kleinberg, J.; Weinberger, KQ Über Fairness und Kalibrierung. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4.–8. Dezember 2017.

53. Karimi-Haghighi, M.; Castillo, C. Verbesserung eines Tools zur Rückfallvorhersage mit maschinellem Lernen: Wirksamkeit und algorithmische Fairness. In Proceedings der 18. Internationalen Konferenz zu Künstlicher Intelligenz und Recht, São Paulo, Brasilien, 21.–25. Juni 2017; S. 210–214.

Machine Translated by Google

54. Friedler, SA; Scheidegger, C.; Venkatasubramanian, S. Die (Un-)Möglichkeit von Fairness: Unterschiedliche Wertesysteme erfordern unterschiedliche Mechanismen für faire Entscheidungsfindung. Commun. ACM **2021,** 64, **136–143. [CrossRef]**

55. García, V.; Sánchez, JS; Mollineda, RA Über die Wirksamkeit von Vorverarbeitungsmethoden beim Umgang mit unterschiedlichen Klassenungleichgewichten . Knowl.-Based Syst. **2012,** 25, **13–21. [CrossRef]**

56. Breiman, L. Klassifikations- und Regressionsbäume; **Routledge: London, Großbritannien, 2017.**

57. Pedregosa, F.; Varoquaux, G.; Grafort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Maschinelles Lernen in Python. J. Mach. Learn. Res. **2011,** 12, **2825–2830.**

58. Kelly, M.; Longjohn, R.; Nottingham, K. The UCI Machine Learning Repository. 2024. Online verfügbar: https://archive.ics.uci.edu (abgerufen am 10. Juni 2024).

59. Becker, B.; Kohavi, R. Adult. UCI Machine Learning Repository. 1996. Online verfügbar: https://archive.ics.uci.edu/dataset/2/Erwachsene (abgerufen am 10. Juni 2024). [CrossRef]

60. Hofmann, H.; Statlog (Deutsche Kreditdaten). UCI Machine Learning Repository. 1994. Online verfügbar: https://archive.ics.uci.edu/dataset/144/statlog+german+credit+data (abgerufen am 10. Juni 2024). [CrossRef]

61. Dressel, J.; Farid, H. Genauigkeit, Fairness und Grenzen der Rückfallvorhersage. Sci. Adv. **2018, 4, eaao5580. [CrossRef]**

62. Huang, Q.; Mao, J.; Liu, Y. Ein verbesserter Rastersuchalgorithmus zur Optimierung von SVR-Parametern. In Proceedings der 2012 IEEE 14th International Conference on Communication Technology, Chengdu, China, 9.–11. November 2012; S. 1022–1026.

63. Caton, S.; Haas, C. Fairness im maschinellen Lernen: Eine Untersuchung. ACM Comput. Surv. **2023,** 56, **1–38. [CrossRef]**

64. Alabdulmohsin, I.; Lucic, M. Ein nahezu optimaler Algorithmus zur Entzerrung trainierter Machine-Learning-Modelle. In Proceedings of the 35. Konferenz zu neuronalen Informationsverarbeitungssystemen, online, 6.–14. Dezember 2021.

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultieren.

Machine Translated by Google

Elektronik

MDPI

Artikel

Sichere Verarbeitung und Verteilung von Daten, die auf privaten InterPlanetary File System mit Zero-Knowledge-Beweisen

Kyohei Shibano 1,*, Kensuke Ito 1, Changhee Han2, Tsz Tat Chu2, Wataru Ozaki 2 und Gento Mogi 1

1 Abteilung für Technologiemanagement für Innovation, Fakultät für Ingenieurwissenschaften, Universität Tokio,
2 Tokio 113-8656, Japan Callisto
Inc., Tokio 171-0022, Japan * Korrespondenz: shibano@tmi.tu-tokyo.ac.jp

Zusammenfassung: In dieser Studie wird eine neue Methode zum Teilen von Daten vorgeschlagen, die ein privates InterPlanetary File System – ein dezentrales Speichersystem, das innerhalb eines geschlossenen Netzwerks betrieben wird – verwendet, um Daten an externe Stellen zu verteilen und gleichzeitig deren Authentizität überprüfbar zu machen. Unter den beiden Betriebsmodi von IPFS, öffentlich und privat, konzentriert sich diese Studie auf die Methode zur Verwendung von privatem IPFS. Privates IPFS ist nicht für die breite Öffentlichkeit zugänglich; obwohl es bei der Verteilung von Daten an externe Parteien ein Risiko der Datenmanipulation birgt , stellt die vorgeschlagene Methode die Authentizität der empfangenen Daten sicher. Insbesondere wendet diese Methode eine Art Zero-Knowledge-Beweis an, nämlich das Groth16-Protokoll von zk-SNARKs, um sicherzustellen, dass die Daten dem Inhaltsbezeichner in einem privaten IPFS entsprechen. Darüber hinaus wird der Name des Empfängers in die verteilten Daten eingebettet, um eine unbefugte Sekundärverteilung zu verhindern. Experimente bestätigte In zukünftigen Studien wird die vorgeschlagene Methode auf größere und vielfältigere Datentypen angewendet.

Schlüsselwörter: IPFS; Zero-Knowledge-Beweis; circom; zk-SNARKs; privates IPFS; Datenverteilung; Datenverarbeitung ; Datensicherheit

check for updates

Zitat: Shibano, K.; Ito, K.; Han, C.; Chu, TT; Ozaki, W.; Mogi, G. Sichere Verarbeitung und Verteilung von Daten, die auf einem privaten interplanetaren Dateisystem verwaltet werden, unter Verwendung von Zero-Knowledge-Beweise. Electronics **2024,** 13, **3025. https://doi.org/10.3390/Elektronik13153025**

Wissenschaftliche Redakteurin: Aryya Gangopadhyay

Empfangen: 13. Juni 2024
Überarbeitet: 14. Juli 2024
Akzeptiert: 20. Juli 2024
Veröffentlicht: 31. Juli 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.
Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Einleitung

Dezentrale Systeme sind robust, da sie keinen einzelnen Ausfallpunkt aufweisen. Daher werden sie in vielen Unternehmensbereichen eingesetzt, darunter Kryptowährungen, Lieferkettenmanagement , Finanzdienstleistungen und digitale Identität. Um große Datenmengen wie Bilder zu speichern , benötigen diese Systeme Speicherfunktionen, die von Natur aus dezentralisiert sind. Blockchain, das häufig in Verbindung damit verwendet wird, verarbeitet normalerweise kleinere Datenmengen wie Transaktionshistorien und fungiert als Ledger-Datenbank. Das InterPlanetary File System (IPFS) ist ein bekanntes dezentrales Speichersystem, das Daten über mehrere Knoten hinweg speichert, um die Datenverfügbarkeit zu verbessern. Das IPFS gibt es in zwei Varianten: öffentliches IPFS, bei dem die Daten von jedem Benutzer mit uneingeschränktem Zugriff gespeichert werden können , und privates IPFS, bei dem ein geschlossenes Netzwerk eingerichtet wird, auf das nur bestimmte Organisationen oder Gruppen zugreifen können , was mehr Privatsphäre **und Sicherheit bietet .**

Beim Speichern von Daten in IPFS ist es wichtig, die Unterschiede zwischen öffentlichem und privatem IPFS zu verstehen. Öffentliches IPFS ermöglicht jedem den Zugriff auf Daten, während privates IPFS nur innerhalb bestimmter Organisationen oder Gruppen zugänglich ist, was Datenschutz und Sicherheit verbes Beim Speichern sensibler Informationen, wie vertraulicher Daten, in öffentlichen IPFS ist die Anwendung eines geeigneten Verschlüsselungsschemas unerlässlich, um den Datenschutz zu gewährleisten. Im Gegensatz dazu bietet privates IPFS eine höhere Sicherheit für die Datenspeicherung, da es nur innerhalb eines geschlossenen Netzwerks zugänglich ist.

Insbesondere für Organisationen wie Unternehmen oder Gesundheitseinrichtungen birgt die Speicherung von Daten in öffentlichen IPFS trotz der Verwendung starker Verschlüsselungstechnologien inhärente R Darüber hinaus besteht in solchen Fällen ständig die Gefahr von Datenlecks aufgrund von Betriebsfehlern . Die Daten sind zwar verschlüsselt, aber der Welt zugänglich und damit anfällig für Brute-Force-Angriffe und andere Sicherheitsbedrohungen.

Machine Translated by Google

In Bezug auf die Zugänglichkeit ermöglicht öffentliches IPFS allgemeinen Benutzern den direkten Zugriff und Abruf Daten. In privaten IPFS müssen Daten jedoch von Mitgliedern der Organisation oder Gruppe, die das Netzwerk bildet. Wenn während dieses Prozesses Daten manipuliert werden, können Benutzer nicht in der Lage, es zu erkennen. Daher ist es wichtig, den für den Umgang mit Daten verantwortlichen Vermittlern zu vertrauen Übertragung in solchen Fällen wird obligatorisch. Um das oben erwähnte Vertrauensproblem anzugehen, Hier wird eine Methode zur Verteilung von in einem privaten IPFS gespeicherten Daten an externe Einheiten vorgeschlagen. und gleichzeitig seine Authentizität überprüfbar zu machen. Das Groth16 [1] Protokoll von zk-SNARKs, einer Art von ZKP wird auf in einem privaten IPFS gespeicherte Daten angewendet, um die Authentizität der Daten sicherzustellen. Darüber hinaus werden die Informationen des Empfängers in die verteilten Daten eingebettet, um zu verhindern, unerlaubte Sekundärverteilung. Die vorgeschlagene Methode der Datenfreigabe ist wichtig weil es auf den Fall privater IPFS zugeschnitten ist.

Die Unterschiede in mehreren Aspekten, einschließlich Sicherheit und Zugänglichkeit, bei der Speicherung Daten in öffentlichen IPFS und privaten IPFS innerhalb der Unternehmensdomäne werden zusammengefasst in Tabelle 1. Diese Studie schlägt Lösungen für die mit privaten IPFS verbundenen Bedrohungen vor.

Tabelle 1. Vergleich zwischen öffentlichen und privaten IPFS im Unternehmensbereich.

	Öffentliches IPFS	Privates IPFS
Vertrauensmodell	**Vertrauenslos**	**Erfordert Vertrauen in die operative Gruppe**
Zugangsbeschränkungen	**Für jeden zugänglich**	**Nur innerhalb der Betriebsgruppe zugänglich**
Risiko von Datenlecks	**Ständiges Risiko von Datenlecks aufgrund von Benutzerfehlern.**	**Geringes Risiko von Datenlecks innerhalb eines geschlossenen Netzwerks.**
Umgang mit vertraulichen Information	**Erfordert eine ordnungsgemäße Verschlüsselung**	**Daten, die in IPFS gespeichert sind, erfordern selbst keine hochgradige Verschlüsselung; es gibt einen Vertrauenspunkt bei der Weitergabe von Daten an Benutzer**
Das Risiko eines Brute-Force-Angriffs ist immer vorhanden		**Niedrig**
Datenabrufmethode	**Direkter Zugriff durch Benutzer**	**Von Mitgliedern der Organisation oder Gruppe** erhaltene Daten
Bedrohungen	**Erfordert eine Verschlüsselung, die eine Entschlüsselung durch unbefugte Benutzer** verhindert	Bei der Übermittlung **von Daten an Benutzer besteht das Risiko einer Manipulation.**

Der Rest dieses Dokuments ist wie folgt gegliedert. Abschnitt 2 präsentiert verwandte frühere Forschung. Abschnitt 3 beschreibt die grundlegenden Technologien, d. h. ZKP und zk-SNARKs. Abschnitt 4 beschreibt die Struktur der vorgeschlagenen Methode, während Abschnitt 5 die mögliche Anwendungen dieser Methode. Abschnitt 6 stellt die Implementierung dieser Methode vor, Abschnitt 7 befasst sich mit den Experimenten, die zur Überprüfung der Wirksamkeit der Implementierung. Abschnitt 8 enthält eine Diskussion der experimentellen Ergebnisse, während Abschnitt 9 stellt die Schlussfolgerungen des Papiers und einen Überblick über künftige Herausforderungen dar.

2. Verwandte Studien

Bestehende dezentrale Systeme nutzen IPFS, insbesondere in Kombination mit Blockchain Technologie. Kumar et al. [2] schlugen eine Methode zur sicheren Verwaltung medizinischer Daten vor, indem Integration von IPFS in eine Blockchain. Azbeg et al. [3] schlugen speziell ein System vor, das Verwaltung und Speicherung medizinischer Daten mithilfe von privatem IPFS und einer autorisierten Blockchain durch Verwendung einer Proxy-Neuverschlüsselung, um eine sichere Entschlüsselung durch benannte Ärzte zu gewährleisten. Wenn ein Wenn ein Arzt einige Patientendaten erhält, bekommt er die neu verschlüsselten Daten über ein Krankenhaus. Hossan et al. [4] schlugen auch ein System zur sicheren Aufzeichnung von Informationen für Mitfahrgelegenheiten vor Dienste mit IPFS und einer privaten Blockchain.

Mit dem Schwerpunkt auf der Kontrolle der Verteilung der von IPFS verwalteten Daten, Lin et al. [5] schlug ein System zum Schutz privater Daten vor, das ein verbessertes IPFS kombiniert mit einem Blockchain. Dieses System zeichnete Dateimetadaten und Zugriffsberechtigungen auf der Blockchain auf, Damit können Benutzer die Dateifreigabe steuern. Darüber hinaus implementierte das System effiziente Verwaltungsfunktionen mithilfe von Smart Contracts und verbesserte so die Datensicherheit und -verwaltung.

Machine Translated by Google

Flexibilität. Battah et al. [6] entwickelten ein System, das Mehrparteienauthentifizierung (MPA), Proxy-Neuverschlüsselung und Smart Contracts auf einer Blockchain für die dezentrale Zugriffskontrolle von verschlüsselten, in IPFS gespeicherten Daten nutzte. Huang et al. [7] führten einen vertrauenswürdigen IPFS-Proxy ein, um Zugriffskontrolle und Gruppenschlüsselverwaltung für verschlüsselte, in IPFS gespeicherte Daten zu realisieren. Sun et al. [8] schlugen ein System vor, das nur Personen mit entsprechenden Attributen erlaubte, in IPFS gespeicherte verschlüsselte Daten unter Verwendung eines auf Chiffretextrichtlinienattributen basierenden Verschlüsselungssystems zu entschlüsseln, wodurch ein effizientes medizinisches Informationsmanagement erleichtert wurde. Kang et al. [9] ermöglichten die Verteilung von mit privatem IPFS und einer privaten Blockchain verwalteten Daten an externe Benutzer unter Verwendung eines Named Data Network (NDN). Zudem schlugen Uddin et al. [10] ein File-Sharing-System vor , das IPFS und Public-Key-Infrastruktur (PKI)-Technologie nutzte, ohne dass eine vertrauenswürdige Dritt

Mehrere Studien haben ZKP zur Datenverteilung verwendet. Beispielsweise schlugen Li et al. [11] ein datenschutzfreundliches Verkehrsmanagementsystem vor, das nicht interaktive Zero-Knowledge-Bereichsnachweise mit einer Blockchain kombinierte. Ein Prototyp mit Hyperledger Fabric und Hyperledger Ursa erfüllte die Datenschutzanforderungen für Echtzeit-Verkehrsmanagement.

Diese Studie schlägt eine Methode zur angemessenen Verarbeitung und Verteilung von Daten vor, die innerhalb eines privaten IPFS verwaltet werden, und zwar an Benutzer außerhalb des Netzwerks. Damit bietet sie einen anderen Ansatz als die in früheren Studien vorgeschlagenen. Einige Studien haben die Proxy-Neuverschlüsselung als geeignete Methode zur Datenspeicherung und -verteilung in IPFS übernommen **[3,6]. Mit dieser Methode können verteilte Daten erneut verschlüsselt werden, um mit dem privaten Schlüssel des Empfängers entschlüsselt zu werden. Darüber hinaus ermöglicht es beim Speichern von Daten in IPFS den Empfängern, die Richtigkeit ihrer empfangenen Daten nach der Entschlüsselung zu überprüfen, indem der Hashwert der vorverschlüsselten Daten in der Blockchain** aufgezeichnet **wird. Diese Methode kann jedoch keine Fälle verarbeiten, in denen Daten verarbeitet werden, wie z. B. das Einbetten des Namens des Empfänge**

3. Zero-Knowledge-Proof ZKP

ist ein kryptografisches Protokoll, das es einem Beweiser ermöglicht, einem Prüfer die Gültigkeit einer Aussage zu beweisen, ohne weitere Informationen außer der Gültigkeit der Aussage preiszugeben. Die Annahme dieser Studie ist, dass die einer externen Entität bereitgestellten Daten auf der Grundlage einer bestimmten CID generiert werden. Unser Ziel ist es, einem Mitglied eines privaten IPFS (Beweiser) zu ermöglichen, diese Aussage einer externen Entität (Prüfer als Empfänger der Daten) zu beweisen, ohne weitere wichtige Informationen (wie IPFS-Zugriffsrechte und Verschlüsselungsschlüssel) preiszugeben.

ZKP, insbesondere das in dieser Studie verwendete Groth16-Protokoll von zk-SNARKs, beginnt mit einem vertrauenswürdigen Setup, bei dem beide Parteien öffentliche Parameter festlegen, die für die sichere Generierung und Überprüfung von Beweisen von entscheidender Bedeutung sind. Im ZKP-Schema generieren wir zunächst einen Schaltkreis, der den Prozess beschreibt, für den ein Beweis vorgesehen ist. Der Schaltkreis enthält die zu überprüfenden Bedingungen, wie z. B. das Vorhandensein einer CID. Durch das ZKP-Schema werden kryptografische Schlüssel – insbesondere ein Beweisschlüssel und ein Überprüfungsschlüssel – erstellt. Diese Schlüssel sind für die Erstellung eines Beweises für den Schaltkreis und dessen Überprüfung von entscheidender Bedeutung. Unter Verwendung des Beweisschlüssels und der Eingabedaten generiert der Beweiser einen Beweis, der die Gültigkeit der Ausgabedaten anhand der im Schaltkreis angegebenen Bedingungen aufzeigt. Der Prüfer verwendet dann den Überprüfungsschlüssel, um den Beweis und die Ausgabedaten zu überprüfen. Wenn der Beweis gültig ist, bestätigt dies die Integrität der Daten, ohne zugrunde liegend

In dieser Studie wird die Groth16-Verarbeitung mit circom [12,13] und snarkjs [14] durchgeführt. Der Prozessablauf ist in Abbildung 1 zusammengefasst.

zk-SNARKs wird aufgrund seiner nicht-interaktiven Natur und Effizienz eingesetzt, die besonders für Systeme mit Smart Contracts von Vorteil sind, da der Rechenaufwand zur Überprüfung des Nachweises gering ist. Darüber hinaus ist zk-SNARKs für seinen hohen Rechenaufwand und die Notwendigkeit fortgeschrittener PC-Spezifikationen bekannt. Beispielsweise dauert in einer der repräsentativen Anwendungen von zk-SNARKs, Zcash [15], der Nachweisgenerierungsprozess für eine einzelne anonyme Transaktion über eine halbe Minute [16].

Machine Translated by Google

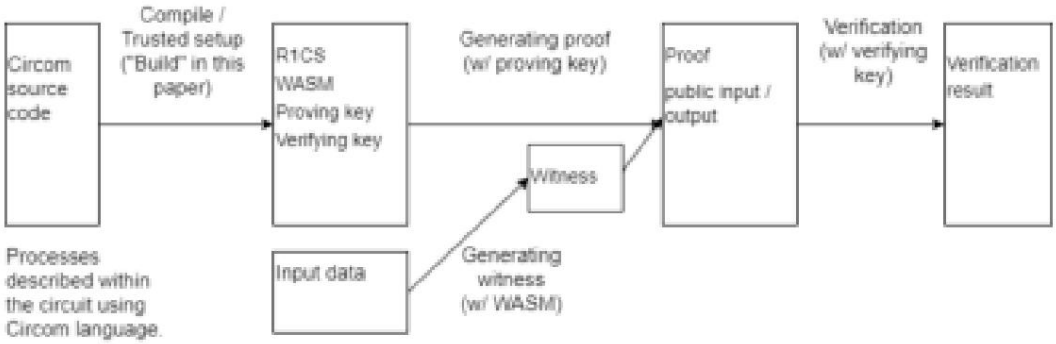


Abbildung 1. Zero-Knowledge-Beweis mit Circom.

4. Vorgeschlagenes

System Abbildung 2 bietet einen Überblick über das vorgeschlagene System.

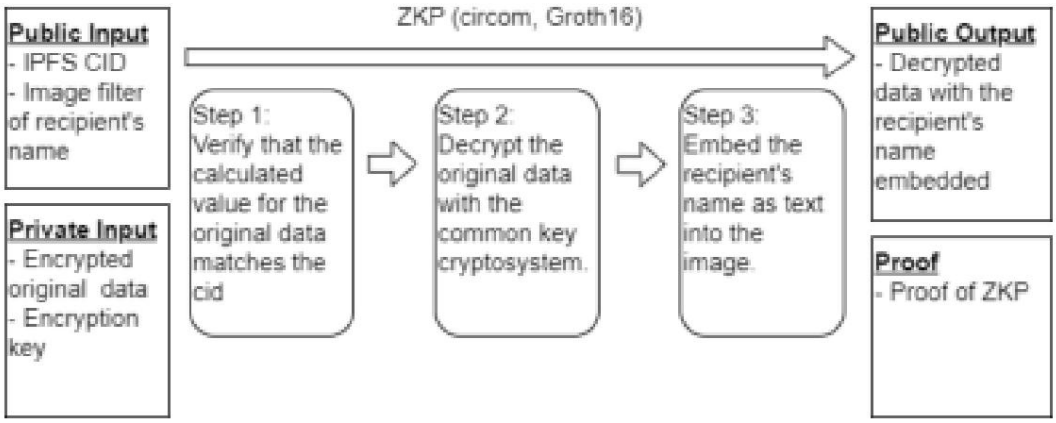


Abbildung 2. Prozessablauf von ZKP: ein Beispiel für die Bilddatenverarbeitung.

Dabei gehen wir von folgenden Annahmen aus: •

Private IPFS wird von einer begrenzten Anzahl von Mitgliedern betrieben. •
Daten werden in IPFS in einem verschlüsselten Format unter Verwendung symmetrischer Kryptografie gespeichert. • Der Verschlüsselungsschlüssel wird ausschließlich von einem einzelnen Mitglied gehalten, d. h. einem Administrator.

Das vorgeschlagene System erleichtert die Erstellung eines ZKP-Nachweises durch den Schaltkreis durch das Mitglied, das den Verschlüsselungsschlüssel besitzt (entspricht einem Administrator). Die Ein- und Ausgaben (außer Beweisen) dieses Prozesses sind wie

folgt: • Öffentliche

Eingabe: – CID der ursprünglich verschlüsselten
Daten; – Ein Filter zum Einbetten des Namens des Empfängers in die

Daten; • Private

Eingabe: – Verschlüsselte
Originaldaten; –

Verschlüsselungsschlüssel; • Ausgabe:

– Entschlüsselte Daten mit eingebettetem Namen des Empfängers.

Beachten Sie, dass die ZK-optimierte Implementierung (Abschnitt 6.1) dem öffentlichen Input mehr Informationen hinzufügt. Insbesondere umfasst der interne Prozess von ZKP die folgenden Schritte:

1. Berechnen der CID der ursprünglich verschlüsselten Daten, um eine Übereinstimmung mit den verschlüsselten Da
registrierte CID.

Machine Translated by Google

2. Entschlüsseln der Originaldaten mithilfe einer symmetrischen Kryptografie.
3. Einbettung des Empfängernamens in die entschlüsselten Bildrohdaten basierend auf dem Filter im öffentlichen Input bereitgestellt.

Im vorgeschlagenen System konzentrieren wir uns nicht nur auf Bilddaten, sondern verwenden sie als Beispieldaten, um die Anwendbarkeit des vorgeschlagenen Schemas zu überprüfen. Filter werden verwendet, um die Effizienz der Verarbeitung innerhalb des Schaltkreises zu verbessern. Da das Einbetten von Namensdaten in den Schaltkreis rechenintensiv ist, wird ein erheblicher Teil der Bildverarbeitung im Voraus außerhalb des Schaltkreises durchgeführt und ein Filter erstellt. Mithilfe öffentlicher Eingaben und Beweise kann der Empfänger überprüfen, ob die entschlüsselten Daten (d. h. die Ausgabe) mit ihrem eingebetteten Namen **aus den Originaldaten (die in der privaten Eingabe enthalten sind) generiert wurden, die** mit der CID verwaltet werden. Der Empfänger **kann bei mindestens einem Mitglied des Netzwerks nachfragen, um die Existenz der CID im privaten IPFS zu bestätigen.**

Zusammenfassend ermöglicht der oben genannte Prozess dem Empfänger die Überprüfung der empfangenen Daten, indem Sie die folgenden Aufgaben ausführen:

- **Überprüfen Sie, ob die empfangenen Daten aus den mit der CID von verwalteten Daten generiert wurden privates IPFS,**
- **bestätigen Sie anhand des Nachweises, dass der gesamte Vorgang korrekt durchgeführt wurde, ohne die verschlüsselten Daten oder den Verschlüsselungsschlüssel direkt zu kennen, und • überprüfen Sie durch Nachfrage bei mindestens einem Netzwerkmitglied, ob die CID konkret innerhalb des privaten IPFS existiert.**

Das Neue an dem vorgeschlagenen System ist, dass es die Überprüfung der Datenauthentizität ermöglicht, indem mindestens einem Mitglied des Netzwerks vertraut wird, selbst wenn der Empfänger nicht über den Verschlüsselungsschlüsse (Für den Empfänger ist es selbstverständlich, mindestens einem Mitglied eines bestimmten Mehrmitgliedersystems zu vertrauen . Wenn keinem der Mitglieder vertraut werden kann, besteht nur ein geringer Anreiz, von diesem Netzwerk verwaltete Daten zu empfangen.)

5. Potentielle Anwendungen In

der Medizinbranche werden Patientendiagnosedaten über mehrere medizinische Einrichtungen hinweg verwaltet. Mit dem vorgeschlagenen System können Patienten überprüfen, ob die Daten, die sie erhalten, tatsächlich in privaten IPFS verwaltet werden, um die Authentizität sicherzustellen. Außerdem wird hier eine All-in-One-Plattform für die Forschung und Entwicklung des maschinellen Lernens mit medizinischen Bildern vorgeschlagen [17]. Auf dieser Plattform werden anonymisierte medizinische Bilder in privaten IPFS verwaltet, die von einer Gruppe medizinischer Einrichtungen betrieben werden. Das System ermöglicht es Forschern im Bereich des maschinellen Lernens, die sich außerhalb des Netzwerks befinden, zu überprüfen, ob die Bilddaten tatsächlich im privaten IPFS verwaltet werden. Darüber hinaus können medizinische Einrichtungen durch die Einbettung der Informationen über die Forscher im Bereich des maschinellen Lernens in die Bilddaten das Risiko einer Sekundärverteilung verringern.

Wenn die Anwendung nicht auf die Einbettung des Empfängernamens beschränkt ist, können die potenziellen Anwendungen des vorgeschlagenen Systems noch weiter erweitert werden. Stellen Sie sich beispielsweise ein Szenario vor, in dem ein bestimmtes Unternehmen ein privates IPFS einrichtet, um vertrauliche Dokumente zwischen seinen Konzernunternehmen auszutauschen. Wenn Mitarbeiterdaten enthalten sind, können diese durch das Verbergen privater Daten und deren Weitergabe an externe Stellen die Zugehörigkeit der Mitarbeiter zum Unternehmen bestätigen und gleichzeitig Nehmen wir außerdem an, dass eine Universität ein privates IPFS eingerichtet hat, um nur dem akademischen Personal Zugriff auf die Leistungsdaten der Studenten zu gewähren. In diesem Fall können die Studenten überprüfen, ob ihre erhaltenen Leistungsdaten tatsächlich im privaten IPFS verwaltet werden.

Das vorgeschlagene System unterstützt also einen Hybridfall – die Verteilung interner Daten an spezifische externe Entitäten nach Bedarf – in realen Umgebungen weit verbreitet.

6. Implementierung Das

vorgeschlagene System wurde zur Verarbeitung von Bilddaten mit circom implementiert, einem renommierten Tool, das auf die Konstruktion von zk-SNARK-Schaltkreisen spezialisiert ist. Circom ermöglicht die Beschreibung von Rechenprozessen innerhalb eines Schaltkreises mithilfe seiner einzigartigen Sprache, und die

Machine Translated by Google

Die nach der Kompilierung generierte ausführbare Datei kann über die JavaScript-Bibliothek snarkjs aufgerufen werden. Diese Anordnung ermöglicht die Beschreibung von Schaltungsprozessen in circom, und die externe Verarbeitung und die Prüfung der Schaltungskorrektheit werden mit JavaScript durchgeführt. Für die zk-SNARKs Schema, Groth16 wurde verwendet; es ist bekannt für seine relativ schnellere Ausführungsgeschwindigkeit als andere zk-SNARKs-Schema.

Im Einzelnen haben wir an zwei Implementierungsarten für Bilddaten gearbeitet: einem Standard Implementierung mit allgemeinen kryptographischen Techniken und eine ZK-optimierte Implementierung mit ZK-freundlichen kryptographischen Techniken zur Reduzierung der Rechenz Schaltung. Diese Implementierungen wurden zum Vergleich der benötigten Rechenzeiten verwendet. ZKP-Schaltungen erfordern beträchtliche Rechenzeit, selbst für Berechnungen, die leicht durch Computer-Software gehandhabt werden (dies ist besonders spürbar bei der Arbeit mit Bilddaten). Daher ist die Rechenleistung für den praktischen Einsatz von entscheidender Bedeutung.

6.1. Standardimplementierung

Abschnitt 4 beschreibt die Dateneingabe in die Schaltung. Zur Vereinfachung der In-Circuit-Die ursprünglichen verschlüsselten Daten wurden als Bitmap-Bilddaten formatiert, die kompatibel sind mit mit OS/2-Standards. Die ersten 54 Bytes der Bilddaten enthalten Informationen wie die Breite, Höhe und Farbtiefe von Bildern [18]. Die Farbtiefe beträgt 8 Bit und jede Farbe Jeder Komponente in RGB wird ein Byte zugewiesen, was zu einer Darstellung von 3 Bytes pro Pixel führt.

Zunächst prüft das System, ob die verschlüsselten Daten, die als private Eingabe eingegeben wurden, stimmt mit der als öffentliche Eingabe angegebenen CID überein. Wenn sie nicht übereinstimmen, gibt das System an, ein Fehler und die als Ausgabe ausgegebenen Bilddaten sind eine Bytefolge, in der alle Werte sind 0x00. CID dient als entscheidender Mechanismus zur eindeutigen Identifizierung von Dateien und effizient Daten aus IPFS abrufen. CID gibt es in zwei Versionen: V0 und V1 [19]. Dabei ist die flexiblere Version CID V1 wurde verwendet. CID enthält einen Hash der jeweiligen Daten, wodurch unterschiedliche Daten haben unterschiedliche CIDs. Normalerweise wird CID V1 mithilfe des SHA256-Hashes berechnet Funktion, und die Standardimplementierung verwendet SHA256 zur Berechnung der CID.

Die Datenstruktur von CID V1 ist in Tabelle 2 dargestellt.

Tabelle 2. CID V1-Datenstruktur.

Byte Position	Beschreibung	Mehrwert bei der Umsetzung
Erstes Byte	CID-Version	Nummer:
Zweites Byte	Multibase-Präfix	0x55: Rohdaten
Drittes Byte	Hash-Funktionskennung	0x12: SHA-256
Viertes Byte	Hash-Länge	0x20: 32 Byte
Ab dem fünften Byte	Hashwert	SHA-256-Hashwert (32 Bytes)

Die Kodierung für CID erfolgt mit Base32. Base32 kodiert eine Folge von Bytes Basierend auf dieser Struktur erstellt, um CID zu generieren.

Innerhalb der Schaltung wird der eingegebene CID-Wert aus Base32 dekodiert und das System prüft, ob der extrahierte Hashwert mit dem SHA256-Hash übereinstimmt, der aus dem verschlüsselte Daten.

Anschließend werden die verschlüsselten Daten entschlüsselt. Als Verschlüsselungsverfahren wird AES-CTR verwendet. Algorithmus, der eine Art symmetrische Kryptographie ist. Die AES-CTR-Verschlüsselung und Entschlüsselung in circom-chacha20 [20] wurde verwendet. Zur Entschlüsselung wurde AES-CTR-Verschlüsselung verwendet, der Verschlüsselungsschlüssel und die Nonce, die bei der Verschlüsselung verwendet werden, sind erforderlich. Sie werden in die Schaltung als 256-Bit-Schlüssel bzw. 128-Bit-Nonce als private Eingänge. Außerdem AES-CTR verarbeitet Datenmengen in Vielfachen von 16. Wenn also die Länge des Bildes Daten vor der Verschlüsselung sind kein Vielfaches von 16, Nullen (0x00) werden an das Ende der Daten angehängt um es dieser Anforderung anzupassen.

Schließlich wird ein Filter auf die entschlüsselten Daten angewendet, um den Namen des Empfängers einzubetten. Die Implementierung der Texteinbettung direkt in der Schaltung kann die Rechenlast erheblich erhöhen ; daher wird außerhalb der Schaltung ein Filter erstellt, der eine erhebliche

Machine Translated by Google

Teil der Bildverarbeitung im Voraus. Die für den Text, der den Namen des Empfängers darstellt, verwendete Schriftart ist die Schriftart Misaki [21]. Der Filter wird dann verwendet, um die Verarbeitung innerhalb der Schaltung zu optimieren. Der Filter ist eine Liste von Zahlen, wobei Werte von 0 bis 255 verwendet werden, um die Farbe jedes Pixels zu ändern, falls sie von der Farbe der Pixel im Originalbild abweicht; außerdem gibt ein Wert von 300 an, dass die Farbe wie im Originalbild bleiben soll. Dieser Filter stellt die Position im Bild dar, an der der Name des Empfängers eingefügt werden soll.

Innerhalb der Schaltung werden die angegebenen Pixelfarben in den entschlüsselten BMP-Daten basierend auf diesem Filter geändert.

6.2. ZK-optimierte Implementierung

Die ZK-optimierte Implementierung ändert die Hash-Funktion, die Verschlüsselungstechnologie und die In-Circuit-Verarbeitung zur standardmäßigen ZK-freundlichen Verschlüsselungsimplementierung. Diese Implementierung verbessert die Rechenleistung und bewertet den Unterschied in der Rechengeschwindigkeit zwischen ZK-freundlicher Verschlüsselung und allgemeiner Verschlüsselung nicht. Daher wurde auch die In-Circuit-Verarbeitung geändert.

Als Hash-Funktion zur Berechnung der CID wurde **Poseidon-Hash [22] verwendet. Die Verwendung des Poseidon-Hash für CIDs wird nicht offiziell unterstützt; daher wurde er speziell für diese Studie entwickelt. Obwohl SHA256 häufig in allgemeinen Berechnungen verwendet wird, erfordert es in ZKP-Schaltungen beträchtliche Rechenzeit. Der Poseidon-Hash ist in circom und JavaScript implementiert (circomlib [23] bzw. circomlibjs [24]). Er wird über einen endlichen Körper mit Primzahlordnung berechnet und kann bis zu 16 Eingangsvariablen akzeptieren.**

Die verwendete Reihenfolge ist kleiner als das Maximum von 32 Bytes, aber größer als das Maximum von 31 Bytes. Dies bedeutet, dass jeder der 16 Eingänge Daten enthalten darf, die diese Reihenfolge nicht überschreiten. In dieser Implementierung werden die für die Hash-Berechnung vorgesehenen Daten als Eingabewerte in 31-Byte-Segmente unterteilt. Wenn die Unterteilung 16 Segmente überschreitet, wird der Poseidon-Hash für die ersten 16 Segmente berechnet. Dieses Ergebnis wird zu den nächsten 15 Datensegmenten für einen nachfolgenden Poseidon-Hash-Eingang addiert. Der Vorgang wird wiederholt, bis alle Eingabedaten für die Hash-Berechnung verwendet wurden. Wenn der letzte Eingang rechnerisch nicht 16 Segmente vervollständigt, werden die fehlenden Eingänge auf Null gesetzt, u

Beim Generieren von CID aus dem Poseidon-Hashwert sollte die Bytefolge der CID V1-Datenstruktur folgen und Base32-codiert sein. Um die Rechenzeit jedoch weiter zu verkürzen , wird bei dieser Implementierung die Base32-Codierung weggelassen und der Poseidon-Hashwert direkt als Ersatz für CID verwendet. Das Aufteilen der Eingabedaten in 16 Segmente innerhalb des Schaltkreises ist rechenintensiv; daher wird diese Aufteilung außerhalb des Schaltkreises durchgeführt und als Eingabe bereitgestellt. In diesem Fall werden die verschlüsselte Datenbytefolge und die Werteliste zur Berechnung des Poseidon-Hashes als öffentliche Eingaben bereitgestellt, wodurch überprüft werden kann, ob beide Datensätze dieselben Informationen darstellen. Empfänger können bestätigen, dass die für den Poseidon-Hash berechneten Daten und die im Schaltkreis entschlüsselten Daten identisch sind, indem sie diese beiden Werte gegenseitig konvertieren und überprüfen. Da Benutzer in diesem Fall die entschlüsselten Daten erhalten können , besteht die Sorge, dass Passwörter durch Brute-Force-Angriffe oder andere Mittel verloren gehen.

Als Verschlüsselungstechnologie haben wir anstelle der AES- CTR-Verschlüsselung die Poseidon-Verschlüsselung **[25]** übernommen . **Bei der Poseidon-Verschlüsselung, die in circom und TypeScript implementiert ist (poseidon- encryption-circom2 [26]), wird der öffentliche Schlüssel des Empfängers empfangen, ein gemeinsamer Schlüssel generiert und eine sichere Ver- und Entschlüsselung durch beide Parteien sichergestellt. In diesem Fall wird jedoch direkt ein gemeinsamer Schlüssel generiert und zur Ver- und Entschlüsselung verwendet. Der Schaltung werden zwei Werte, die die Koordinaten einer elliptischen Kurve darstellen, und ein Nonce-Wert als private Eingänge für die Verschlüsselung bereitgestellt . Darüber hinaus wird der Filter auf die gleiche Weise wie in der Standardimplementierung implementie**

7. Auswertung

Wir haben ein Beispielprogramm auf Basis der oben genannten Implementierungen erstellt, das circom zur Beschreibung der Schaltung und snarkjs zur Ausführung der Schaltung und zur Überprüfung verwendet

Machine Translated by Google

Beweise. Als kryptographische Bibliotheken stehen circom-chacha20 [20], circomlib [23], circomlibjs [24], und poseidon-encryption-circom2 [26] wurden verwendet.

Für jeden Schaltkreis wurden die Standard- und die ZK-optimierte Implementierung umgesetzt. und ihre Rechenzeiten wurden während der Ausführung verglichen. Weiße Bitmap-Bilder waren die Zielbilder, und die Experimente wurden mit dem Buchstaben „A“ als eingebettetes Zeichen. Da das Einbetten beliebiger Zeichen die Verarbeitung nicht verändert Durch das Einbetten eines einzelnen Zeichens durch den Filter wurde ein Vergleich der Berechnungszeiten ermöglicht. Darüber hinaus haben wir die Bildgrößen variiert, um die Ausführungszeiten für jeden Schaltkreis zu messen. Die verwendeten Größen waren 10 × 10, 15 × 15, 30 × 15, 30 × 30, 60 × 30, 60 × 60, 120 × 60, 120 × 120, und 180 × 120 Pixel. Als Ausführungsumgebung diente Windows 11 mit einem Ryzen 9 3950X CPU und 128 GB RAM, betrieben unter Ubuntu 22.04 in einer WSL2-Umgebung.

Abbildung 3 zeigt ein Beispielbild, das von der Schaltung erzeugt wird, speziell für das 60 × 30 Pixel große Größe mit der ZK-optimierten Implementierung. Die Ergebnisse für jede Bildgröße werden präsentiert in Tabelle 3, wo nichtlineare Beschränkungen die Anzahl der nichtlinearen Beschränkungen in der Schaltung, Build-Zeit ist die Zeit, die benötigt wird, um Circom zu kompilieren und die Schaltung auszugeben, und Beweis gen time ist die Zeit, die benötigt wird, um Beweise mit der Schaltung zu generieren. Da die Standardimplementierung AES-CTR-Verschlüsselung verwendet, werden am Ende Daten mit 0x00 angehängt, um sich Die Eingabegröße ist ein Vielfaches von 16.

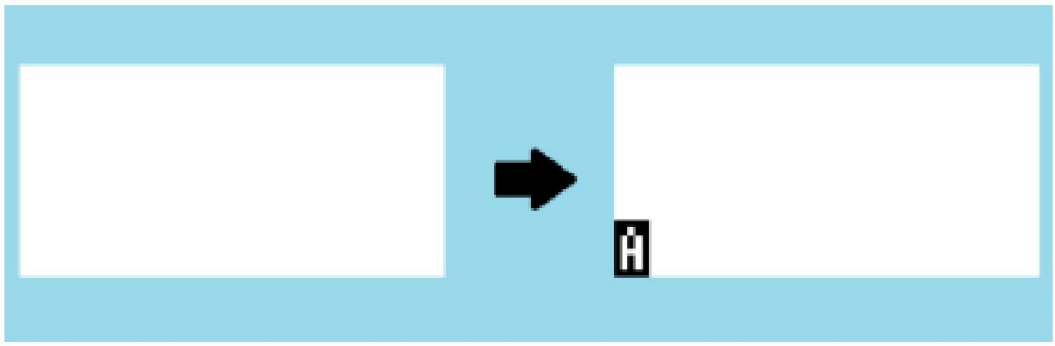


Abbildung 3. Ein vom Schaltkreis für eine Bildgröße von 60 × 30 durch ZK-optimierte Implementierung generiertes Bild.

Tabelle 3. Vergleich der Ausführungszeiten der Schaltung.

	Pixel	Bildgröße [Byte]	Nichtlinear Einschränkungen	Bauzeit [MS]	Beweis Gen Zeit [ms]
Standard					
	10 × 10	384	558.341	668.220	14.377
	15 × 15	784	1.095.292	1.316.666	25.545
	30 × 15	1440	1.980.688	1.696.370	35.161
	30 × 30	2816	3.867.989	3.657.578	65.785
	60 × 30	5456	7.453.300	7.864.583	126.262
ZK-optimiert					
	10 × 10	376	35.407	128.979	3076
	15 × 15	775	72.725	173.210	4032
	30 × 15	1435	134.663	275.630	5900
	30 × 30	2815	263.450	470.872	9733
	60 × 30	5455	509.375	850.612	17.571
	60 × 60	10.855	1.013.483	1.257.418	29.044
	120 × 120	43.255	4.036.913	6.754.928	96.580

In Standard- und ZK-optimierten Implementierungen für 60 × 60 Pixel und 180 × 120 Pixel Bildgrößen, dem System ging der Speicher aus und die Berechnung konnte nicht abgeschlossen werden. In der ZK-optimierten Implementierung wurde die Anzahl der nichtlinearen Einschränkungen reduziert auf ungefähr ein Zehntel der Standardimplementierung bei gleicher Bildgröße.

Machine Translated by Google

Dadurch verkürzte sich die Zeit für den Aufbau und die Proof-Generierung. Die maximal verarbeitbare Bildgröße betrug jedoch immer noch nur bis zu 120 × 120 Pixel, was für praktische Anwendungen erheblich klein ist.

8. Diskussion

Obwohl bei der ZK-optimierten Implementierung ZK-freundliche kryptografische Technologien verwendet und In-Circuit-Prozesse optimiert wurden, betrug die maximal verwaltbare Bildgröße ungefähr 120 × 120 Pixel. Dies beschränkt den praktischen Nutzen auf deutlich kleinere Bildgrößen. Die Forschung zur Verbesserung der Leistung von ZKPs läuft jedoch weiter, und zukünftige technologische Fortschritte könnten die Verarbeitung größerer Bildgrößen ermöglichen.
Beispielsweise erreichten Zhang et al. [27] eine zehnfache Beschleunigung von zk-SNARKs durch den Einsatz von ASICs. Ma et al. [16] verwendeten eine ähnliche Grafikverarbeitungseinheit, um die Beweisgenerierungszeit zu beschleunigen , und erreichten im Vergleich zu herkömmlichen Methoden eine bis zu 48,1-mal schnellere Leistung. Darüber hinaus wurden Methoden zur Vereinfachung von Rechenprozessen vorgeschlagen, wie beispielsweise die „Faltungsmethode“. Diese Methode komprimiert die zu beweisenden Aussagen [28]. Da zunehmend an Geschwindigkeitsverbesserungen geforscht wird, wird wahrscheinlich auch der Speicherverbrauch optimiert. Dies wird in Zukunft möglicherweise die Verarbeitung größerer Bildgrößen ermöglichen.

Darüber hinaus haben wir festgestellt, dass unsere vorgeschlagene Methode Datengrößen von etwa 10 KB verarbeiten kann. Obwohl die direkte Anwendung unseres Vorschlags auf realistische Bilddaten (von mehreren MB bis zu Dutzenden von MB) eine Herausforderung darstellt, könnte die Aufteilung der Daten in Blöcke durch Änderung der Verschlüsselung und eingebetteter Zeichenfolgen die Anwendung machbar mac

Darüber hinaus erfordert unsere Implementierung einen Wert basierend auf der Größe der zu verarbeitenden (verschlüsselnden) Originaldaten als Argument während der Schaltungsgenerierung. Daher muss für alle Daten eine Schaltung generiert werden. Die Schaltungsgenerierungszeit (Build-Zeit) erhöht sich mit der Bilddatengröße erheblich; beispielsweise dauert die Generierung einer Schaltung für eine Bildgröße von 120 × 120 selbst bei einer ZK-optimierten Implementierung mehr als 112 Minuten (6.754.928 ms). Sobald die Schaltung jedoch erstellt ist, ist die Proof-Erstellungszeit unter denselben Bedingungen kurz und beträgt etwa 97 s (96.580 ms). Mit anderen Worten: Sobald eine Schaltung erstellt ist, ist die Proof-Erstellung nicht mehr zeitintensiv. Diese Tatsache stellt keine praktischen Probleme dar, wenn dasselbe Bild an verschiedene Personen verteilt wird.

In ZK-freundlichen Implementierungen werden verschlüsselte Daten als öffentliche Eingabe eingegeben. Der Umgang mit Verschlüsselungsschlüsseln für Bilder erfordert sorgfältige Überlegung. In privaten IPFS verwaltete Daten sind verschlüsselt. Wenn jedoch Verschlüsselungsschlüssel durchsickern, könnten die verschlüsselten Daten entschlüsselt werden. Daher sollten bestimmte Benutzer, die private IPFS verwalten, Administratoren werden, um die Schlüssel sorgfältig zu verwalten, oder es könnte eine Blockchain vom Typ Konsortium im selben Netzwerk eingerichtet werden, um Zugriffsrechte entsprechend festzulege

9. Schlussfolgerungen

Hier wurde eine neue Methode vorgeschlagen, um in privatem IPFS gespeicherte Daten an externe Entitäten zu verteilen und gleichzeitig ihre Authentizität überprüfbar zu machen. Die Methode wendet eine Art ZKP, zk- SNARKs, an, um die CID der Daten zu überprüfen und den Namen des Empfängers einzubetten. Mit diesem Ansatz können externe Entitäten überprüfen, ob die empfangenen Daten aus den Originaldaten im privaten IPFS generiert wurden, ohne dass Details wie IPFS-Zugriffsrechte und Verschlüsselungsschlüssel erforde

Um **die Rechenleistung der vorgeschlagenen Methode zu verbessern,** wurden **eine Standardimplementierung mit herkömmlichen kryptografischen Techniken und eine ZK- optimierte Implementierung mit ZK-freundlichen kryptografischen Schemata implementiert .** Experimente mit einem **Beispielprogramm bestätigten die Wirksamkeit der vorgeschlagenen Methode für eine Bilddatengröße von bis zu 120 × 120 Pixeln.**

Diese vorgeschlagene Methode erweitert den nutzbaren Bereich dezentraler Speichersysteme auf einen Hybridfall – die Verteilung interner Daten an bestimmte externe Einheiten nach Bedarf. Diese Studie ebnet einen neuen Weg für den Austausch sensibler Informationen zwischen verschiedenen Sektoren innerhalb und außerhalb einer Gruppe. Für die breite praktische Anwendbarkeit der vorgeschlagenen Methode auf größere und vielfältigere Datentypen wie Bilder und Videos muss die Verarbeitungsgeschwindigkeit jedo

Machine Translated by Google

verbessert werden und es müssen Methoden zur Datenaufteilung eingesetzt werden, die im Rahmen unserer zukünftigen Studien liegen.

Beiträge der Autoren: Konzeptualisierung, KS, CH, TTC und WO; Schreiben – Vorbereitung des Originalentwurfs , KS; Schreiben – Überprüfung und Bearbeitung, KS und KI; Aufsicht, GM; Projektverwaltung, KS. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr

Finanzierung: Diese Forschung wurde als Gemeinschaftsforschungsprojekt zwischen der Universität Tokio und Callisto Inc. durchgeführt und von Callisto Inc. finanziert. Diese Arbeit wurde vom Stiftungslehrstuhl für Blockchain-Innovation und dem Mohammed bin Salman Center for Future Science and Technology for Saudi-Japan Vision 2030 (MbSC2030) an der Universität Tokio unterstützt.

Datenverfügbarkeitserklärung: Der für die Simulationen verwendete Quellcode ist auf GitHub verfügbar. https://github.com/blockchaininnovation/circom_image_processing (abgerufen am 21. März 2024).

Interessenkonflikte: Die Autoren Changhee Han, Tsz Tat Chu und Wataru Ozaki waren bei der Firma Callisto Inc. angestellt. Die übrigen Autoren erklären, dass die Forschung ohne kommerzielle oder finanzielle Beziehungen durchgeführt wurde, die als potenzieller Interessenkonflikt au

Verweise

1. **Groth, J. Zur Größe von paarungsbasierten nicht-interaktiven Argumenten. In Proceedings of the Advances in Cryptology–EUROCRYPT 2016: 35. jährliche internationale Konferenz zur Theorie und Anwendung kryptographischer Techniken, Wien, Österreich, 8.–12. Mai 2016; Proceedings, Teil II 35; Springer: Berlin/Heidelberg, Deutschland, 2016; S. 305–326.**
2. **Kumar, S.; Bharti, AK; Amin, R. Dezentrale sichere Speicherung von Krankenakten mit Blockchain und IPFS: Ein Vergleich Analyse mit zukünftigen Richtungen.** Secur. Priv. **2021, 4, e162. [CrossRef]**
3. **Azbeg, K.; Ouchetto, O.; Andaloussi, SJ BlockMedCare: Ein Gesundheitssystem basierend auf IoT, Blockchain und IPFS für Daten Managementsicherheit.** Ägypten. Inform. J. **2022,** 23, **329–343. [CrossRef]**
4. **Hossan, MS; Khatun, ML; Rahman, S.; Reno, S.; Ahmed, M. Sicherung von Mitfahrdiensten mit IPFS und Hyperledger basierend auf privater Blockchain. In Proceedings der 24. Internationalen Konferenz für Computer- und Informationstechnologie (ICCIT) 2021, Dhaka, Bangladesch, 18.–20. Dezember 2021; IEEE: Piscataway, NJ, USA, 2021; S. 1–6.**
5. **Lin, Y.; Zhang, C. Eine Methode zum Schutz privater Daten in IPFS. In Proceedings der 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Dalian, China, 5.–7. Mai 2021; IEEE: Piscataway, NJ, USA, 2021; S. 404–409.**
6. **Battah, AA; Madine, MM; Alzaabi, H.; Yaqoob, I.; Salah, K.; Jayaraman, R. Blockchain-basierte Mehrparteienautorisierung für den Zugriff auf IPFS-verschlüsselte Daten.** IEEE Access **2020, 8, 196813–196825. [CrossRef]**
7. **Huang, HS; Chang, TS; Wu, JY Ein sicheres Dateifreigabesystem basierend auf IPFS und Blockchain. In Proceedings der 2. International Electronics Communication Conference, Singapur, 8.–10. Juli 2020; S. 96–100.**
8. **Sun, J.; Yao, X.; Wang, S.; Wu, Y. Blockchain-basiertes sicheres Speicher- und Zugriffsschema für elektronische Krankenakten in IPFS.** IEEE Access **2020, 8, 59389–59401. [CrossRef]**
9. **Kang, P.; Yang, W.; Zheng, J. Blockchain-Methode zum Teilen privater Dateispeicher basierend auf IPFS.** Sensors **2022,** 22, **5100. [CrossRef] [PubMed]**
10. **Uddin, MN; Hasnat, AHMA; Nasrin, S.; Alam, MS; Yousuf, MA Sicheres Dateifreigabesystem mit Blockchain-, IPFS- und PKI-Technologien. In Proceedings der 5. Internationalen Konferenz für elektrische Informations- und Kommunikationstechnologie (EICT) 2021, Khulna, Bangladesch, 17.–19. Dezember 2021; IEEE: Piscataway, NJ, USA, 2021; S. 1–5.**
11. **Li, W.; Guo, H.; Nejad, M.; Shen, CC Datenschutzwahrendes Verkehrsmanagement: Ein von Blockchain und Zero-Knowledge-Proof inspirierter Ansatz.** IEEE Access **2020, 8, 181733–181743. [CrossRef]**
12. **Bellés-Muñoz, M.; Isabel, M.; Muñoz-Tapia, JL; Rubio, A.; Baylina, J. Circom: Eine Schaltungsbeschreibungssprache für den Bau Zero-Knowledge-Anwendungen.** IEEE Trans. Dependable Secur. Comput. **2022, 20, 4733–4751. [CrossRef]**
13. **Offizielle Circom-Website. Online verfügbar: https://iden3.io/circom (abgerufen am 24. März 2024).**
14. **Snarkjs Github Repository. Online verfügbar: https://github.com/iden3/snarkjs (abgerufen am 4. Juni 2024).**
15. **ZCash. Online verfügbar: https://z.cash/ (abgerufen am 12. Juli 2024).**
16. **Ma, W.; Xiong, Q.; Shi, X.; Ma, X.; Jin, H.; Kuang, H.; Gao, M.; Zhang, Y.; Shen, H.; Hu, W. Gzkp: Ein GPU-beschleunigtes Zero-Knowledge-Proof-System. In Proceedings der 28. ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Band 2, Vancouver BC Kanada, 25.–29. März 2023; S. 340–353.**
17. **Han, C.; Shibano, K.; Ozaki, W.; Osaki, K.; Haraguchi, T.; Hirahara, D.; Kimura, S.; Kobayashi, Y.; Mogi, G. All-in-one-Plattform für KI-Forschung und -Entwicklung in der medizinischen Bildgebung, die Datenerfassung, -auswahl, -annotation und -vorverarbeitung umfasst.** In Proceedings of the Medical Imaging 2024: Imaging Informatics for Healthcare, Research, and Applications, San Diego, CA, USA, 18.–23. Februar 2024; **SPIE: Bellingham, WA, USA, 2024; Band 12931, S. 311–315.**
18. **Miano, J.** Komprimierte Bilddateiformate: Jpeg, png, gif, xbm, bmp; **Addison-Wesley Professional: Boston, MA, USA, 1999.**

Machine Translated by Google

19. Inhaltskennungen (CIDs). Online verfügbar: https://docs.ipfs.tech/concepts/content-addressing/#cids-are-not-file-hashes (abgerufen am 24. März 2024). 20. circom-chacha20 Github Repository. Online verfügbar: https://github.com/reclaimprotocol/circom-chacha20 (aufgerufen am 24. März 2024).

21. Die 8 × 8 Punkte große japanische Schriftart „Misaki Font“. Online verfügbar: https://littlelimit.net/misaki.htm (abgerufen am 11. Juni 2024 (Auf Japanisch)

22. Grassi, L.; Khovratovich, D.; Rechberger, C.; Roy, A.; Schofnegger, M. Poseidon: Eine neue Hash-Funktion für {Zero-Knowledge} -Beweissysteme. In Proceedings des 30. USENIX Security Symposiums (USENIX Security 21), Vancouver, BC, Kanada, 11.–13. August 2021; S. 519–535.

23. Circomlib Github Repository. Online verfügbar: https://github.com/iden3/circomlib (abgerufen am 21. März 2024).

24. Circomlibjs Github Repository. Online verfügbar: https://github.com/iden3/circomlibjs (abgerufen am 21. März 2024).

25. Khovratovich, D. Verschlüsselung mit Poseidon. 2019. Online verfügbar: https://drive.google.com/file/d/1EVrP3DzoGbmzkRmYnyEDcIQcXVU7GlOd/Ansicht (abgerufen am 19. Juli 2024).

26. Poseidon-Encryption-Circom2 Github Repository. Online verfügbar: https://github.com/Shigoto-dev19/poseidon-encryption-c ircom2 (abgerufen am 21. März 2024).

27. Zhang, Y.; Wang, S.; Zhang, X.; Dong, J.; Mao, X.; Long, F.; Wang, C.; Zhou, D.; Gao, M.; Sun, G. Pipezk: Beschleunigung von Zero- Knowledge-Beweisen mit einer Pipeline-Architektur. In Proceedings des 2021 ACM/IEEE 48th Annual International Symposium on Computer Architecture (ISCA), Valencia, Spanien, 14.–18. Juni 2021; IEEE: Piscataway, NJ, USA, 2021; S. 416–428.

28. Kothapalli, A.; Setty, S.; Tzialla, I. Nova: Rekursive Zero-Knowledge-Argumente aus Faltungsschemata. In Proceedings der Annual International Cryptology Conference, Santa Barbara, CA, USA, 15.–18. August 2022; Springer: Cham, Schweiz, 2022; S. 359–388.

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultieren.

Machine Translated by Google

Elektronik

MDPI

Artikel

Untersuchung und Abmilderung des Kompromisses zwischen Leistung und Fairness durch Protected-Category-Sampling

Gideon Popoola und John Sheppard *

Gianforte School of Computing, Montana State University, Bozeman, MT 59717, USA;
gideon.popoola@student.montana.edu * Korrespondenz: john.sheppard@montana.edu

Zusammenfassung: Algorithmen für maschinelles Lernen sind bei alltäglichen Entscheidungen mittlerweile weit verbreitet und Entscheidungshilfesysteme sind in unserem Alltag allgegenwärtig. Daher hat die Forschung zur Vermeidung und Eindämmung potenzieller Verzerrungen und Ungerechtigkeiten bei den von diesen Algorithmen getroffenen Vorhersagen in den letzten Jahren zugenommen. Die meisten Forschungsarbeiten zu Fairness und Verzerrungsminderung beim maschinellen Lernen behandeln oft jede geschützte Variable separat, aber in Wirklichkeit ist es möglich, dass eine Person mehreren geschützten Kategorien angehört . Daher wurde in dieser Arbeit die Kombination einer Reihe geschützter Variablen und die Generierung neuer Spalten untersucht, die diese geschützten Variablen in Diese neuen Unterkategorien sind häufig extrem unausgewogen, sodass die Risikominderung als Problem der unausgewogenen Klassifizierung angegangen wurde. Konkret wurden vier neue benutzerdefinierte Stichprobenverfahren entwickelt und untersucht, um diese neuen Unterkategorien zu beproben. Diese neuen Stichprobenverfahren werden als „ Protected-category Oversampling“, „Protected-category Proportional Sampling“, „Protected-category Synthetic Minority Oversampling Technique“ (PC-SMOTE) und „Protected-category Adaptive Synthetic Sampling“ (PC-ADASYN) bezeichnet. Diese Stichprobenverfahren modifizieren das bestehende Stichprobenverfahren, indem sie sich bei der Stichprobennahme auf die neuen Unterkategorien statt auf die Klassenbezeichnung konzentrieren. Die Auswirkung dieser Stichprobenstrategien wurde dann anhand der klassischen Leistung und Fairness in Klassifizierungseinstellungen bewert Die Klassifizierungsleistung wurde anhand von Genauigkeit und F1 gemessen, basierend auf dem Training univariater Entscheidungsbäume, und die Fairness wurde anhand von ausgeglichenen ungeraden Differenzen und statistischer Parität gemessen. Um die Auswirkung von Fairness gegenüber Leistung zu bewerten, wurden diese Maßnahmen anhand der Tiefe des Entscheidungsbaums bewertet. Die Ergebnisse zeigen, dass die vorgeschlagenen Methoden in der Lage waren, optimale Punkte zu bestimmen, wodurch die Fairness erhöht wurde, ohne die Leistung zu verringern, wodurch jeder potenzielle Kompromiss zwischen Leistung und Fairness gemildert wurde.

Schlüsselwörter: Fairness, geschützte Kategorien, maschinelles Lernen, Sampling

check for updates

Zitat: Popoola, G.; Sheppard, J. Untersuchung und Eindämmung der Kompromiss zwischen Leistung und Fairness durch Stichprobennahme geschützter Kategorien. Electronics **2024,** 13, **3024. https:// doi.org/ 10.3390/electronics13153024**

Akademische Herausgeber: Niusha Shafiabady und Jianlong Zhou

Empfangen: 11. Juni 2024
Überarbeitet: 23. Juli 2024
Akzeptiert: 24. Juli 2024
Veröffentlicht: 31. Juli 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.
Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

Da Algorithmen des maschinellen Lernens (ML) Entscheidungsfindungs- und Entscheidungsunterstützungssysteme zunehmend dominieren, wirft ihr weitverbreiteter Einsatz in verschiedenen Sektoren dringende Fragen hinsichtlich der Fairness und Transparenz ihrer Vorhersagen auf **[1]. Das Potenzial dieser Algorithmen, bestehende gesellschaftliche Vorurteile aufrechtzuerhalten oder zu verschärfen, hat eine bedeutende Forschungsarbeit ausgelöst, die sich mit der Untersuchung und Eindämmung algorithmischer Ungerechtigkeit befasst. Dies ist von entscheidender Bedeutung, da die von diesen Algorithmen beeinflussten Entscheidungen tiefgreifende Auswirkungen auf Einzelpersonen haben und sich auf Ergebnisse in Bereichen von Finanzen und Beschäftigung bis hin zu Strafjustiz und Gesundheitswesen a**

Die Quelle von Ungerechtigkeit und Verzerrung im ML ist vielschichtig [3]. Insbesondere ist es möglich, dass Ungerechtigkeit direkt aus den ML-Algorithmen selbst entsteht, aufgrund einer möglichen Fehlausrichtung der zugrunde liegenden induktiven Verzerrung **der Algorithmen gegenüber dem Zielkonzept** und der Datenverteilung. Dies wird als algorithmische Verzerrung **bezeichnet . Ein alternatives Problem liegt in einer potenziellen Verzerrung in den Daten, die zum Trainieren der Modelle verwendet werden, wo als direkte Folge der typischen Annahme „unabhängig und identisch verteilt“ (I**

Machine Translated by Google

Methoden besteht das Ergebnis des Lernens darin, die Verzerrung in den Vorhersagen so zu propagieren, dass sie der Verzerrung in den zugrunde liegenden Daten selbst entsprechen. Diese letztere Situation bildet den Schwerpunkt unserer Arbeit hier.

1.1. Voreingenommenheit und Ungerechtigkeit beim maschinellen Lernen

Unter Voreingenommenheit versteht man die nachteilige, unfaire oder ungleiche Behandlung einer Einzelperson oder Gruppe aufgrund bestimmter Merkmale, die oft als sensible oder geschützte Merkmale bezeichnet werden [4]. Beispiele für diese geschützten Merkmale sind Alter, Rasse, Behinderung, Geschlecht und Gender [5]. Voreingenommenheit im ML kann grob in ungleiche Behandlung (direkte Ungerechtigkeit) und wirksame Behandlung (indirekte Ungerechtigkeit) unterteilt werden **[6]. Direkte Ungerechtigkeit liegt vor, wenn geschützte Merkmale explizit bei der Entscheidungsfindung verwendet werden. Indirekte Ungerechtigkeit kommt heute immer häufiger vor. Bei dieser Art der Ungerechtigkeit werden geschützte Attribute nicht explizit verwendet. Sie liegt stattdessen vor, wenn das Vertrauen auf Variablen, die mit diesen Attributen verbunden sind, zu erheblich unterschiedlichen Ergebnissen für die geschützten Gruppen führt. Diese anderen Variablen werden als Proxy-Merkmale bezeichnet. Beispiele für Voreingenommenheit in der realen Welt sind die historische Praxis des „Redlining“ in den USA, bei der Bewohnern von Postleitzahlengebieten mit überwiegend Minderheitenanteil Hypotheken verweigert wurden, geschlechtsspezifische Voreingenommenheit im Einstellungsverfahre**

Obwohl diese Entscheidungshilfe-Tools dabei helfen, den Entscheidungsprozess zu automatisieren, können sie zu einer unfairen Behandlung von Einzelpersonen oder Gruppen führen, sowohl direkt als auch indirekt [8]. Ungerechtigkeiten können in mehreren Bereichen der Modellierung auftreten, beispielsweise im Trainingsdatensatz . Dies kann passieren, wenn der Trainingsdatensatz keine faire Darstellung der geschützten Kategorien bietet, sodass die „Grundwahrheit“ schwer zu bestimmen ist. Betrachten Sie beispielsweise einen Datensatz eines Unternehmens, in dem eine bestimmte Gruppe in der Vergangenheit diskriminiert wurde. Nehmen wir beispielsweise an, weibliche Mitarbeiter in diesem Unternehmen wurden nicht so befördert wie ihre männlichen Kollegen, die im Gegensatz dazu Karriere machen konnten, obwohl beide Gruppen die gleiche Leistung erbrachten. In dieser Situation ist der wahre Wert der Beiträge weiblicher Mitarbeiter – die Grundwahrheit – nicht sichtbar. Infolgedessen erkennt und berücksichtigt ein ML-Algorithmus, der mit diesen Daten trainiert wird, diese Verzerrung wahrscheinlich und verewigt so bestehende Vorurteile. Dies könnte dazu führen, dass der Algorithmus diskriminierende Entscheidungen trifft, beispielsweise männliche Kandidaten häufiger für eine Einstellung oder Beförderung empfiehlt als gleich oder besser qualifizierte weibliche Kandidaten.

Ein weiterer Bereich, in dem Ungerechtigkeit auftreten kann, ist der ML-Algorithmus selbst [9]. ML-Algorithmen können immer noch diskriminierende Entscheidungen treffen, selbst wenn sie mit einem unvoreingenommenen Datensatz trainiert werden, in dem die „Grundwahrheit“ genau dargestellt ist. Diese Situation entsteht, wenn die Fehler des Systems Einzelpersonen einer bestimmten Gruppe oder Minderheit überproportio Betrachten wir beispielsweise einen Algorithmus zur Brustkrebserkennung, der bei schwarzen Personen deutlich mehr Falsch-Negativ-Ergebnisse liefert als bei weißen. Das heißt, er erkennt Brustkrebs bei schwarzen Patienten häufiger nicht als bei weißen. Würde dieser Algorithmus zur Abgabe von Behandlungsempfehlungen verwendet, würde er fälschlicherweise mehr schwarzen als weißen Personen von einer Behandlung abraten, was zu rassistischen Ungleichheiten bei den Behandlungsergebnissen führen würde. Dies unterstreicht die dringende Notwendigkeit, sicherzustellen, dass die Algorithmen hinsichtlich ihrer Trainingsdaten und der Auswirkungen ihrer Fehler auf unterschiedliche Populationen für alle Gruppen gleich gut funktionieren. Ergebnisse aus der bisherigen Literatur berichten von mehreren Fällen, in denen Algorithmen zu unfairer Behandlung führten, z. B. Redlining und Racial Profiling [10], Diskriminierung bei Hypotheken [11], Einstellung und Personalauswahl [12].

Während erhebliche Anstrengungen unternommen wurden, um die Verzerrung bei ML-Vorhersagen zu beheben [13,14], konzentrierte sich ein Großteil der bestehenden Forschung auf die Abschwächung der Verzerrung für einzelne geschützte Attribute isoliert [15]. Beispielsweise können die meisten bestehenden Ansätze in einem Datensatz mit zwei geschützten Attributen, Rasse und Geschlecht, entweder ein faires Modell mit Rasse oder ein faires Modell mit Geschlecht lernen, aber kein faires Modell mit sowohl Rasse als auch Geschlecht [7]. Identitäten in der realen Welt sind jedoch nicht singulär; sie sind komplex und vielschichtig, wobei Einzelpersonen oft mehreren geschützten Gruppen gleichzeitig angehören [16]. Beispielsweise kann eine Person aufgrund mehrerer geschützter Attribute wie Alter, Rasse und Geschlecht gleichzeitig diskriminiert werden. Diese Intersektionalität kann zu zusammengesetzten Formen der Verzerrung führen und

Machine Translated by Google

Diskriminierung, die durch Fairnessmaßnahmen mit einer einzigen Variable nicht ausreichend angegangen werden. Daher ist es wichtig, Methoden zu entwickeln, die die mehrdimensionale Natur persönlicher Identitäten ganzheitlich berücksichtigen. Dieses Projekt versucht, diese Lücke zu schließen, indem es Kombinationen geschützter Kategorien in Betracht zieht und diese geschützten Kategorien so zu umfassenden Mehrkategoriengruppen zusammenfasst. Es zielt darauf ab, die vielschichtigen Komplexitäten der Voreingenommenheit mithilfe neuartiger Stichprobenmethoden für geschützte Kategorien effektiver anzugehen und so die Vielschichtigkeit persönlicher Identitäten und potenzielle

Die in diesem Artikel vorgestellte Arbeit ist motiviert durch das Problem der Verwendung von ML-Algorithmen für die Entscheidungsfindung in sozial sensiblen Bereichen wie Kreditbewertung, Einstellung oder Hypothekenbewertung , wobei mit dieser Situation gearbeitet wird, in der eine Person mehreren geschützten Kategorien angehören kann . Bei einem gekennzeichneten Trainingsdatensatz mit zwei oder mehr geschützten Merkmalen kombiniert die vorgeschlagene Methode diese geschützten Attribute und teilt sie dann in neue Multikategorien auf. Diese neuen Kategorien sind wahrscheinlich extrem unausgewogen und müssen ausgeglichen werden, um die Fairness der Vorhersage unserer ML-Algorithmen zu verbessern. Beliebte Stichprobenmethoden wie Oversampling [17], Synthetic Minority Oversampling Technique (SMOTE) [18], Adaptive Synthetic Sampling (ADASYN) [19] usw. ziehen Stichproben über Klassenbezeichnungen hinweg , was nicht mit dem Ziel unserer Forschung der Stichprobennahme über die neuen Multikategorien hinweg übereinstimmt. Daher wird eine neue Klasse von Modifikationen dieser Stichprobenmethoden vorgeschlagen, die Stichproben über die neue Kategorie statt über Klassenbezeichnungen hinweg ziehen können. Diese neue Klasse modifizierter Stichprobennahme wird als geschützte Kategoriestichprobennahme **bezeichnet .** Die daraus resultierenden vorgeschlagenen **Stichprobenverfahren für geschützte Kategorien** werden verwendet, um die neuen Kategorien zu prüfen und auszugleichen, bevor **die Klassifizierung durchgeführt wird. Das Neue an dieser Arbeit ist zweifacher Natur. Erstens kombiniert der vorgeschlagene Ansatz die geschützten Kategorien, um neue Multikategorien zu bilden, die nachahmen, wie die Identität des Menschen in der realen Welt aussieht. Zweitens werden vorhandene Stichprobenverfahren so modifiziert, dass sie mit der Stichprobenziehung**

Nur zu Demonstrationszwecken wurde ein univariater Entscheidungsbaum als Klassifizierungsalgorithmus gewählt. Ziel ist es, die Auswirkungen der verschiedenen Stichprobenverfahren auf die Leistung zu demonstrieren , wobei davon ausgegangen wird, dass sich unabhängig von der **zugrunde liegenden Lernmethode ähnliche Trends zeigen .** Das vorgeschlagene Stichprobenverfahren wurde mit der **Baseline (nicht abgetastete Daten) verglichen, wobei Genauigkeit und F1 als Leistungskennzahlen für die Klassifizierung sowie ausgeglichene Quotenunterschiede und statistische Parität als Fairnesskennzahlen verwendet wurden. Außerdem wurden mehrere Analysen durchgeführt, um zu zeigen, wie sich die maximale Tiefe des Entscheidungsbaums auf Genauigkeit und Fairness auswirkt.**

1.2. Forschungsfrage

Ausgehend von der empirischen Beobachtung, dass manchmal ein Kompromiss zwischen Fairness und ML-Leistung besteht [20], versucht diese Forschung, mehrere Fragen zu beantworten, beispielsweise wie dieser Kompromiss gemildert werden könnte. Insbesondere wollen wir beantworten, ob die Methode der geschützten Kategorie-Stichprobennahme zur Bewältigung der Fairness diesen Kompromiss abmildern kann. Darüber hinaus können wir einen methodischen Rahmen entwickeln, der Verzerrungen über diese kombinierten geschützten Variablen hinweg effektiv abmildert, ohne die Vorhersagegenauigkeit von ML-Modellen zu beeinträchtigen? Schließlich planen wir, die Frage zu beantworten, wie sich die Tiefe eines Entscheidungsbaums sowohl auf die Genauigkeits- als auch auf die Fairnessmetriken auswirkt, und so die Beziehung zwischen dem Anpassungsgrad (Unteranpassung durch Überanpassung) und der

1.3. Hypothese

Wir gehen von der Hypothese aus, dass wir durch den Einsatz ausgefeilter geschützter Kategorienstichprobentechniken, die für diese neu formulierten Mehrkategoriengruppen entwickelt wurden, die Modellfairness in Bezug auf ausgeglichene Quotenunterschiede deutlich steigern können, ohne die Klassifizierungsleistung in Bezug auf Genauigkeit und F1 zu verringern. Darüber hinaus untersuchen wir das empfindliche Gleichgewicht zwischen Fairness und Genauigkeit und gehen davon aus, dass es möglich ist, strategische Punkte zu identifizieren, an denen die Fairness ohne nachteilige Auswirkungen auf die Leistung maxir Diese Forschung stellt bestehende Behauptungen über die Existenz von Kompromissen in Fairness und ML in Frage

Machine Translated by Google

Vorhersage. Es bereitet den Boden für zukünftige Untersuchungen der mehrdimensionalen Natur von Identität und Diskriminierung in automatisierten Entscheidungssystemen.

1.4. Beiträge

Das allgemeine Problem der Fairness beim maschinellen Lernen ist insofern bedeutsam, als die heutige Verbreitung von KI- und ML-Systemen erhebliche Auswirkungen auf das Leben und den Lebensunterhalt der Menschen Obwohl die Aufmerksamkeit für faires ML deutlich zugenommen hat, besteht weiterhin Bedarf an Methoden, um faires ML zu fördern, ohne die ML-Leistung negativ zu beeinflussen. Basierend auf einer gründlichen Literaturrecherche und dem oben genannten Bedarf an dieser Art von Arbeit leisten die hier vorgestellten Methoden die folgenden Beiträge: 1. Die weit verbreitete

Annahme, dass es beim maschinellen Lernen einen inhärenten Kompromiss zwischen Fairness und Leistung (d. h. Genauigkeit) gibt, wird durch Beweise in Frage gestellt, die diese Infragestellung untermauern. Insbesondere deuten die Ergebnisse in diesem Dokument darauf hin, dass ein solcher Kompromiss gemildert werden kann, was darauf hindeutet, dass jeder Kompromiss höchstwahrscheinlich mit der Art und Weise zusammenhängt, wie die Daten verwaltet werden.

2. Es werden vier neuartige Vorverarbeitungsmethoden für die Stichprobenentnahme von Daten vorgestellt, die auf der Anwendung einer Multikategorie-Stichprobenstrategie unter Verwendung von in geschützten Kategorien erfassten Daten basieren. Die Methoden gehen von der Annahme und der entsprechenden Hypothese aus, dass das Ausbalancieren der Daten auf der Grundlage dieser Multikategorieeigenschaften die Fairness erhöhen kann, ohne die Leistung des maschinellen Lernmodells zu be

3. Die experimentellen Ergebnisse werden anhand von drei Datensätzen präsentiert, die in der Fair-ML-Community ausführlich untersucht wurden. Die Experimente umfassen Vergleiche mit herkömmlichen Trainingsmethoden ohne Resampling, um die relativen Auswirkungen der vorgeschlagenen Methoden zu demonstrieren. Die Ergebnisse zeigen, dass zwei der vorgeschlagenen Methoden, Protected- Category Synthetic Minority Oversampling Technique (PC-SMOTE) und Protected- Category Adaptive Synthetic Sampling (PC-ADASYN), besonders effektiv bei der Verbesserung von Fairness und Leistung sind.

4. Eine detaillierte Analyse der möglichen Auswirkungen von Unter- und Überanpassung auf die Fairness wird durch die Untersuchung verschiedener Ebenen in einem Entscheidungsbaummodell mit und ohne Verwendung der vorgeschlagenen Stichprobenmethoden präsentiert. Die Ergebnisse zeigen, dass die vorgeschlagenen Methoden in der Lage sind, eine ideale Ebene des Baums zu identifizieren, auf der sowohl Fairness als auch Genauigkeit maximiert werden.

Als Ergebnis der oben genannten Beiträge stellt diese Arbeit einen bedeutenden Fortschritt bei der Lösung von Fairnessbedenken im maschinellen Lernen dar. Eine wichtige Erkenntnis aus den hier vorgestellten Methoden und Ergebnissen ist, dass Fairness ohne Beeinträchtigung der Modellleistung erreicht werden kann.

1.5. Organisation

Dieses Dokument ist wie folgt organisiert. Abschnitt 2 enthält eine detaillierte Erklärung von Fairness und eine Diskussion mehrerer technischer Fairnessmetriken. Anschließend wird in Abschnitt 3 frühere Literatur zu Strategien zur Minderung von Verzerrungen beschrieben. Abschnitt 4 beschreibt unsere vorgeschlagenen Sampling-Techniken, Datensätze und unseren Ansatz zur Hyperparameter-Optimierung. Abschnitt 5 präsentiert die Ergebnisse mehrerer Experimente zusammen mit statistischen Hypothesentests als Mittel zur Validierung dieser Ergebnisse. Abschnitt 6 diskutiert die experimentellen Ergebnisse und analysiert die Leistung jedes Algorithmus bei jedem Datensatz und jeder Metrik. Außerdem werden weitere Ergebnisse zum Einfluss der Baumtiefe auf Fairness und Genauigkeit präsentiert . Abschnitt 7 präsentiert die Grenzen dieser Arbeit und entsprechende Richtungen für künftige Arbeiten und Abschnitt 8 präsentiert eine Reihe von Schlussfolgerungen.

2. Hintergrund

Diese Studie untersucht die Fairness bei der Vorhersage eines Ergebnisses y ÿ **Y aus einer Reihe turen x ÿ X ÿ** Rd **und einige zusätzliche geschützte Attribute s ÿ S ÿ** Rp **Geschlecht und , von Merkmalen wie Rass Geschlecht. Bei der Kreditvorhersage beispielsweise** steht **x für die finanzielle Vergangenheit eines Antragstellers, s für die von ihm selbst angegebene Rasse und sein Geschlecht und** y **für die Frage, ob sein Kredit ge**

Machine Translated by Google

oder verweigert. Ein Vorhersagemodell gilt als fair, wenn seine Fehler gleichmäßig auf geschützte Gruppen wie unterschiedliche Rassen oder Geschlechter verteilt sind. Die Klassenvorhersagen aus den Trainingsdaten D werden als Yˆ D := h(x, s) **für ein** h : **X × S ÿ Y aus einer Klasse H bezeichnet. Die geschützten Attribute s ÿ S in unserer Studie werden als binär angenommen, wobei ein spezieller Wert** n **die nicht privilegierte Gruppe bezeichnet . Zum Beispiel** könnte **S die Rasse und** n **„nicht weiß“ sein; daher ist** die binäre **Natur von S** {w, **n}, wobei** w **weiße Bewerber darstellt, die die privilegierte Gruppe sind, und** n **nicht-weiße Bewerber darstellt, die die nicht privilegierte Gruppe sind. Die Definition lässt sich weiter auf nicht-binä**

X , Datensatz D, einen Merkmalssatz und einen geschützten Attributsatz S mit dem Domänenwert {w, **n}. Die Unterscheidung in D in Bezug auf die Gruppe S** = n , **bezeichnet als** diss=n(D), **ist definiert als** **Die Unterscheidung in beschrifteten Datensätzen kann wie folgt definiert werden: Es gibt einen**

$$\text{diss=n(D)} = \frac{|\{x \ddot{y} D : x(s) = w, h(x) = +\}|}{|\{x \ddot{y} D : x(s) = w\}|} \quad \frac{|\{x \ddot{y} D : x(s) = n, h(x) = +\}|}{\{x \ddot{y} D : x(s) = n\}|}$$

Die obige Definition kann als die Differenz der Wahrscheinlichkeit übersetzt werden, dass ein Bewerber für jeden geschützten Attributbereich {w, **n} zur positiven Klasse gehört. Unsere Studie erweitert die obige Definition, indem sie den Datensatz D betrachtet, der zwei oder mehr geschützte Attribute enthält.**

Es werden zwei gängige Fairnessmaße verwendet. Das erste ist die Equalized Odd Difference **(EOD), die misst, wie diskriminativ oder fair unsere Vorhersage ist. EOD besagt, dass ein binärer Klassifikator yˆ fair ist, wenn seine Falsch-Negativ-Rate (FNR) und seine Wahr-Positiv-Rate (TPR) über den gesamten Bereich von S** gleich sind **[21]. FPR und TPR in Bezug auf das geschützte Attribut s ÿ S mit dem Wert** n **können definiert werden als**

$$TPRn(\hat{y}) = P(\hat{y} = \mathbf{1}|\mathbf{y} = \mathbf{1}, S = n)$$

$$FPRn(\hat{y}) = P(\hat{y} = \mathbf{1}|\mathbf{y} = \mathbf{0}, S = n)$$

EOD wird dann mathematisch als die Differenz zwischen TPR und FPR über verschiedene Gruppen in einem geschützten Attribut definiert. Das heißt,

$$EOD = TPRn(\hat{y}) \ddot{y} FPRn(\hat{y}).$$

Ein fairer Klassifikator hat einen EOD von 0, während ein unfairer Klassifikator einen EOD von 1 hat. Obwohl es in der Praxis fast unmöglich ist, einen vollständig fairen Klassifikator zu erreichen, zielt diese Forschung darauf ab, den EOD zu verbessern, ohne die Genauigkeit zu verringern. Für den EOD gilt:

$$FPRn(\hat{y}) = P(\hat{y} = \mathbf{1}|\mathbf{y} = \mathbf{0}, S = n) = TPRn(\hat{y}) = P(\hat{y} = \mathbf{1}|\mathbf{y} = \mathbf{1}, S = n)$$

Und

$$FPRw(\hat{y}) = P(\hat{y} = \mathbf{1}|\mathbf{y} = \mathbf{0}, S = w) = TPRw(\hat{y}) = P(\hat{y} = \mathbf{1}|\mathbf{y} = \mathbf{0}, S = w)$$

Um die obige EOD-Definition auf unsere Multikategorie auszudehnen, wird der EOD für jede Spalte berechnet und dann der Makrodurchschnitt des EOD als endgültiger EOD dargestellt. Die zweite Metrik, die zur Messung der Fairness bei der ML-Vorhersage verwendet wird, ist die statistische Parität **(SP). SP definiert Fairness als gleiche Wahrscheinlichkeit, als positiv klassifiziert zu werden [22]. Dies kann so interpretiert werden, dass** jede Gruppe in einem geschützten Attribut die gleiche Wahrscheinlichkeit hat, mit einem **positiven Ergebnis klassifiziert zu werden .**

$$P(\hat{y} = \mathbf{1}|\mathbf{S} = w) = P(\hat{y} = \mathbf{1}|\mathbf{S} = n)$$

3. Literaturübersicht

ML-Algorithmen, die zunehmend für die Entscheidungsfindung in kritischen Anwendungen wie Rückfall, Kreditscoring, Kreditentscheidungen usw. eingesetzt werden, könnten zunächst als fair und frei von inhärenter Voreingenommenheit angesehen werden. In Wirklichkeit können sie jedoch jede Voreingenommenheit oder Diskriminierung übernehmen, die in den Daten vorhanden ist, mit denen sie trainiert werden, wie Burt [23] feststellt. Darüber hinaus reicht das bloße Entfernen geschützter Variablen aus dem Datensatz nicht aus, um indirekte D

Machine Translated by Google

könnte es sogar verbergen. Diese Erkenntnis hat den Bedarf an fortschrittlicheren Werkzeugen erhöht, wodurch die Entdeckung und Prävention von Diskriminierung zu einem wichtigen Forschungsgebiet geworden ist, wie [24–27] hervorhebt .

Voreingenommenheit in ML ist ein schnell wachsendes Thema in der Forschungsgemeinschaft für maschinelles Lernen. Voreingenommenheit in einem ML-Modell kann zu einer unfairen Behandlung von Menschen führen, die bestimmten geschützten Gruppen angehören. In letzter Zeit legen Branchenführer immer mehr Wert auf Voreingenommenheit in ML-Modellen und -Software. Das Institute for Electrical and Electronics Engineers (IEEE) [28], Microsoft [29] und die Europäische Union [30] haben kürzlich Grundsätze für faires KI-Verhalten veröffentlicht. Diese Organisationen haben erklärt, dass ML-Modelle in realen Anwendungen fair sein müssen. Strategien zur Minderung von Voreingenommenheit beinhalten die Änderung eines oder mehrerer der folgenden Elemente, um sicherzustellen, dass die vom ML-Algorithmus getroffenen Vorhersagen weniger voreingenommen sind: (a) die Trainingsdaten, (b) der ML-Algorithmus und (c) die daraus resultierenden Vorhersagen selbst. Diese werden jeweils in Vorverarbeitungsansätze [31], Inverarbe

Erstens können die Trainingsdaten vor dem Training des Modells vorverarbeitet werden, um Ungerechtigkeit oder Verzerrung zu verringern. Kamiran und Calders [6] schlagen vor, die Daten zu stichprobenartig zu erfassen oder neu zu gewichten, um Diskriminierung zu neutralisieren. Mit diesem Ansatz kann die Darstellung oder Bedeutung bestimmter Datenpunkte angepasst werden, um eine Klasse gegenüber einer anderen zu bevorzugen (oder zu verringern). Eine andere Methode besteht darin, einzelne Datensätze direkt zu ändern, um Diskriminierung zu verringern, wie in [33] untersucht wird. Bei diesem Ansatz werden beispielsweise Werte in einem Datensatz geändert, um erkennbare Verzerrungen gegenüber bestimmten Gruppen zu verringern. Darüber hinaus wird das von Sondeck et al. [34] eingeführte Konzept der t-Nähe in der Arbeit von [35] auf die Diskriminierungskontrolle angewendet . Die Verwendung der t-Nähe stellt sicher, dass die Verteilung sensibler Attribute in einer bestimmten Gruppe nahe an der Verteilung des Attributs im gesamten Datensatz liegt, wodurch die Privatsphäre gewahrt und Diskriminierung aufgrund sensibler Attribute verhindert wird. Ein gemeinsamer Nenner dieser Ansätze ist die Abwägung der Diskriminierungskontrolle mit dem Nutzen der verarbeiteten Daten, d. h. die Minimierung der Verzerrung, ohne die Genauigkeit, Repräsentativität und allgemeine Nützlichkeit der Daten für prädiktive Modellierung oder Analyse erheblich zu beeinträchtigen . Dieses Gleichgewicht ist von entscheidender Bedeutung, um sicherzustellen,

Insgesamt kann die Vorverarbeitungsmethode weiter in drei Kategorien unterteilt werden: (1) Datenänderung, (2) Datenentfernung und (3) Datenneuabtastung. Methoden der ersten Kategorie zielen darauf ab, die Werte der Trainingsdatenpunkte (einschließlich geschützter Attributwerte , Klassenwerte und Merkmalswerte) zu ändern, um die Verzerrung im Datensatz zu verringern. Ein Beispiel für diese Methode ist die von [15] vorgeschlagene Datenmassage . Ihr Ansatz ordnet die Trainingsdaten und Daten nahe der Entscheidungsgrenze in sowohl privilegierten als auch nicht privilegierten Gruppen werden umgedreht. Alternativ wurde von Fahse et al. [23] eine optimierte Vorverarbeitungsmethode vorgeschlagen, die eine probabilistische Transformation lernt, die die Klassen und Merkmale mit individueller Verzerrung und Gruppengerechtigkeit bearbeitet. In [36] werden die ursprünglichen Attributwerte durch Werte ersetzt, die unabhängig vom Klassenlabel gewählt werden, um ein Modell zu trainieren, das ungefähr ausgeglichene Chancen erreicht. In ähnlicher Weise ersetzen Peng et al. [37] die geschützten Attributwerte durch Werte, die basierend auf anderen Attributen vorhergesagt werden, ähnlich der Datenimputation.

Methoden der zweiten Kategorie zielen darauf ab, ein faires Modell zu trainieren, indem bestimmte Merkmale aus dem Trainingsdatensatz entfernt werden. Ein Beispiel für diese Methode ist die von Dhar et al. [38] vorgeschlagene Datenunterdrückung. In ihrem Artikel werden die geschützten Attribute und Merkmale, die stark mit geschützten Attributen korreliert sind (auch als Proxy-Attribute bekannt), aus dem Datensatz entfernt, um ein faires Modell zu trainieren.

Methoden der dritten Kategorie zielen darauf ab, ein faires Modell zu trainieren, indem entweder die Stichprobengewichte angepasst oder der Datensatz überabgetastet wird. Krasanakis et al. [39] schlugen beispielsweise eine Neugewichtungsmethode vor, bei der die Gewichte der Trainingsstichproben iterativ mit einem theoretisch fundierten Modell angepasst werden, um den Kompromiss zwischen Bias und Genauigkeit zu mildern. In [40] schlugen Chakraborty et al. FairSMOTE als Methode vor, um Trainingspunkte von Minderheitengruppen mit künstlichen Datenpunkten auf der Grundlage der Synthetic Minority Oversampling Technique (SMOTE) [18] zu überabtasten, um ausgewogene Klassenverteilungen zu erreichen. Auch Yan et al. [41] schlugen vor, die Trainingsdaten der Minderheitengruppen mit künstlichen Datenpunk

Machine Translated by Google

Im Gegensatz zu FairSMOTE konzentrierten sich die Autoren auf Szenarien, in denen geschützte Attribute unbekannt sind, und verwendeten eine Clustermethode, um verschiedene demografische Gruppen zu identifizieren.

Bei der Inprocessing-Methode werden Methoden eingesetzt, die die Art und Weise ändern, wie ein ML-Modell trainiert wird, um Verzerrungen zu reduzieren. In [42] wurde ein Ansatz zur kontradiktorischen Debiasing-Methode vorgeschlagen. Dieser Ansatz lernt einen Klassifikator, um die Genauigkeit und Fairness der Vorhersage zu erhöhen, indem eine Variable für die interessierte Gruppe einbezogen wird, indem gleichzeitig ein Prädiktor und ein Gegner gelernt werden. Dies führt zur Generierung eines unvoreingenommenen Klassifikators, da die Vorhersagen keine Informationen zur Gruppenunterscheidung enthalten , die der Gegner nutzen könnte. Alternativ wurde in **[43] ein Algorithmus vorgeschlagen, der eine Fairnessmetrik als Teil der Verlustfunktion verwendet und ein für diese Fairnessmetrik** trainiertes Modell zurückgibt. **Kamishima et al. [22] schlugen eine Regularisierungsmethode vor , die einen Strafterm in die Verlustfunktion eines Klassifikators einbezog, um eine unvoreingenommene Vorhersage zu erzielen. Zafar et al. [44] entwickelten eine neue Gewichtungsmethode, bei der sie das Stichprobengewicht für jedes Trainingsdatum anpassen, um ein bestimmtes Fairnessziel zu erreichen, wie z. B. ausgeglichene Chancen für die Validierungsdaten. In jüngster Zeit wurde die Bias-Minderung als ein Optimierungsproblem mit Nebenbedingungen angegangen, indem eine Fairness-Nebenbedingung hinzugefügt und der Verlust so optimiert wurde, dass er mit dieser Nebenbedingung übereinstimmt [45,46]. Einige Arbeiten modifizieren außerdem n**

Nachbearbeitungsmethoden verringern den Fehler nach der Anpassung eines ML-Modells und umfassen Ansätze wie Kalibrierung, Constraint-Optimierung und Transformationsschwellenwertbildu Solche Methoden schlagen einen Algorithmus vor, der benachteiligten Gruppen günstige Ergebnisse und begünstigten Gruppen ungünstige Ergebnisse innerhalb eines gegebenen Konfidenzintervalls um die Entscheidungsgrenze mit der höchsten Unsicherheit zuweist. Ein Ansatz beispielsweise modifiziert die Spitzenschwellen des Klassifikators, um ein bestimmtes Ziel gleicher Chancen oder gleicher Quoten zu erreichen. Ein weiterer Ansatz beinhaltet die zufällige Mutation der Klassen bestimmter Vorhersagen in verschiedene Klassen [48].

Mehrere neuere Studien [49,50] kombinierten entweder Vorverarbeitung, Zwischenverarbeitung oder Nachverarbeitung zu einer Ensemble-Methode. Beispielsweise kombinieren Bhaskaruni et al. **[50] die Überabtastung der ungleichgewichtsgeschützten Klasse mit einer Entscheidungsgrenzenverschiebung als Nachverarbeitungsmethode, um das Ungerechtigkeitsproblem anzugehen.**

Forscher haben sich mit verschiedenen Konzepten der Diskriminierung und Fairness bei **algorithmischen Entscheidungsprozessen befasst .** Ungleiche Auswirkungen (früher als indirekte Fairness bezeichnet) **werden beispielsweise durch statistische Parität und Gruppenfairness gemessen, wie von Bhaskaruni et al. [50] diskutiert. Andererseits betont das Konzept der individuellen Fairness, das ebenfalls von Bhaskaruni et al. eingeführt wurde, dass ähnliche Personen unabhängig von ihrer Gruppenzugehörigkeit ähnlich behandelt werden sollten. Dieser Ansatz konzentriert sich auf Fairness auf individueller Ebene und stellt sicher, dass Entscheidungen auf der Grundlage relevanter Attribute und nicht auf der Grundlage gruppenbasierter Stereotypen oder Vorurteile getroffen werden.**

Bei Klassifikatoren und anderen Vorhersagemodellen ist das Erreichen gleicher Fehlerraten über verschiedene Gruppen hinweg ein wichtiges Ziel, wie Zhang und Neill [16] hervorheben. Ebenso ist es von entscheidender Bedeutung, eine Kalibrierung **oder das Fehlen von Vorhersageverzerrungen in den Vorhersagen sicherzustellen, wie von Hardt et al. [21]** erörtert . **Die Spannung zwischen diesen Konzepten – Kalibrierung und gleiche Fehlerraten – wird jedoch** von Dwork et al. **[51] und Pleiss et al.** [52] **untersucht , was darauf hinweist, dass es eine Herausforderung sein kann, beide gleichzeitig zu erfüllen. Karimi-Haghighi und Castillo [53] präsentieren verwandte Arbeiten, in denen die Komplexität untersucht wird, die mit der Erreichung algorithmischer Fairness verbunden ist. Friedler et al. [54]** untersuchen **weiter die Kompromisse bei der Erfüllung verschiedener Definitionen algorithmischer Fairness, insbesondere aus der Perspektive der öffentlichen Sicherheit. Da sich unsere Arbeit eher auf die Vorverarbeitung als auf die Modellierung konzentriert, sind Überlegungen wie ausgeglichene Fehlerraten und Vorhersageverzerrungen weniger direkt anwend**

Aus unserer Überprüfung verschiedener Vorverarbeitungsmethoden geht hervor, dass keine Arbeiten durchgeführt wurden, die versuchten, die Fairness für zwei oder mehr geschützte Attribute gleichzeitig zu Außerdem konzentrierte sich die in früheren Arbeiten verwendete Sampling-Methode nur auf Sampling basierend auf Klassenbezeichnungen und nicht auf geschützte Kategorien. Daher wird in diesem Artikel die Vorverarbeitung betont, da sie den anpassungsfähigsten Aspekt der Data Science-Pipeline darstellt [55]. Die Vorverarbeitung ist insofern einzigartig, als sie nicht von der Wahl des Modellierungsalgorithm

Machine Translated by Google

integriert mit Datenfreigabe- und Veröffentlichungsmechanismen. Diese Unabhängigkeit und Flexibilität machen die Vorverarbeitung entscheidend, um Datenqualität und Fairness sicherzustellen, bevor analytische oder prädiktive Modellierungen erfolgen. Schließlich konzentrieren wir uns auf neue benutzerdefinierte Samplingmethoden, die die geschützte Kategorie im Datentraining abtasten, um ein faires Modell z

4. Methodik Der

Schwerpunkt unserer Arbeit liegt auf der Erforschung von Samplingmethoden, um die Fairness in ML zu verbessern, ohne dass die entsprechende Vorhersageleistung darunter leidet, und so den Kompromiss zwischen **Fairness und Leistung** zu mildern . **Daher werden vier neue Samplingmethoden vorgeschlagen, die sich auf das Erreichen dieses Ziels konzentrieren. Diese Samplingmethoden befassen sich mit dem Problem der unausgewogenen Klassen, das durch die neue Multikategorie entsteht, die durch die Kombination der geschützten Kategorien generiert wird. Es werden benutzerdefinierte Samplingmethoden benötigt, da die vorhandenen Methoden Daten auf Grundlage von Minderheits- und Mehrheitsklassen abtasten, aber um die Fairness zu mildern, werden die neuen Multikategorien so abgetastet, dass sie gleich sind. Dies wiederum erfordert eine Änderung der vorhandenen Samplingmethoden, um Daten auf Grundlage dieser neuen Kategorien abzutasten. Dies führt zu vier neuen Samplingmethoden: Oversampling geschützter Kategorien, proportionales Sampling geschützter Kategorien, SMOTE geschützter Kategorien (PC- SMOTE) und ADASYN geschützter Kategorien (PC-ADASYN).**

4.1. Oversampling geschützter Kategorien

Beim Oversampling geschützter Kategorien besteht der erste Schritt darin, die geschützten Kategorien im Datensatz zu kombinieren und die Kombination zu kodieren, um unsere neue Multikategorie zu erzeugen. Beispielsweise werden im Datensatz „Erwachseneneinkommen“ Alter (jung und erwachsen), Rasse (weiß und andere) und Geschlecht (männlich und weiblich) kombiniert, um acht neue Kategorien zu erzeugen, die jeweils **ADULTWHITEMALE, ADULTOTHERSMALE, YOUNGWHITEMALE, YOUNGOTHERSMAIL, ADULTWHITEFEMALE , ADULTOTHERSFEMALE , YOUNGWHITER Diese neuen Kategorien haben unterschiedliche Stichprobengrößen, und das Ziel unseres Oversampling für geschützte Kategorien besteht darin, diese neue Kategorie so auszugleichen, dass die Stichprobengröße jeder der neuen Kategorien der Größe der Kategorie mit der größten Stichprobengröße entspricht. Um Datenlecks zu vermeiden, wird der Datensatz in Training und Test aufgeteilt, wobei Oversampling nur auf die Trainingsdaten angewendet und dann mit einem ungesampelten Testsatz getestet wird.**

Der Pseudocode in Algorithmus 1 zeigt unsere Methode zum Oversampling geschützter Kategorien im Detail. Im Algorithmus wurde die größte Kategorie als Basis verwendet, da dies die Kategorie mit der größten Stichprobengröße ist. Der Sampling-Prozess führt zu neuen Trainingsdaten mit einer ausgewogenen Stichprobengröße für die neue Kategorie. Der Algorithmus funktioniert, indem er die restlichen geschützten Kategorien so abtastet, dass sie der Stichprobengröße der Basis entsprechen. Dieses Sampling wird durchgeführt, indem die Kategorien zusammen mit ihren Klassenbezeichnungen mehrmals

4.2. Proportionale Stichprobenziehung mit geschützten Kategorien

Die Methode der proportionalen Stichprobennahme für geschützte Kategorien ist eine Verallgemeinerung der Überstichprobennahme für geschützte Kategorien, da der Prozess mit der Festlegung einer Zielstichprobengröße beginnt (die ein anzupassender Hyperparameter ist und nicht nur die Größe der größten Mehrfachkategorie), die als Zielstichproben bezeichnet wird. **Dies entspricht der gewünschten Anzahl von Instanzen, die für jede Kategorie benötigt werden. Dieses Ziel stellt Einheitlichkeit über alle Kategorien hinweg sicher und verringert das Risiko einer Modellverzerrung zugunsten häufigerer Kategorien. Das typische Ergebnis der Anwendung dieser Methode ist, dass einige Kategorien, die mehr Stichproben als die** Zielstichproben haben, **unterabgetastet werden, während andere überabgetastet werden, um einen gleichen Anteil von ihnen im Trainingsdatensatz zu erhalten . Der Pseudocode in Algorithmus 2 zeigt die schrittweise Vorgehensweise der Methode der proportionalen Stichprobennahme für geschützte Kategorien.**

Machine Translated by Google

Algorithmus 1: Oversampling geschützter Kategorien

```
baselineCount ÿ Summe der Einträge in „Largest_Category“ von Xtrain
totalCount ÿ Anzahl der Einträge in Xtrain
Basislinienanteil ÿ Basislinienanzahl/Gesamtanzahl
balancedData ÿ Initialisiere einen leeren Datensatz
Kategorien ÿ Liste der Spaltennamen in Xtrain, die mit „combined_category_“ beginnen
für jede Kategorie in Kategorien tun
    categoryData ÿ Einträge in Xtrain auswählen , bei denen Kategorie = 1
    categoryData ÿ kombiniere categoryData mit entsprechenden Labels von ytrain
    categoryCount ÿ Anzahl der Einträge in categoryData
    targetCount ÿ ganzzahliger Teil von totalCount × baselineProportion
    wenn Kategorieanzahl < Zielanzahl dann
        sampledData ÿ Beispiel targetCount aus categoryData mit Ersetzung
        balancedData ÿ sampledData an balancedData anhängen
    anders
        balancedData ÿ Kategoriedaten an balancedData anhängen
Ende, wenn
Ende für
returniere ausgeglichene Daten
```

Algorithmus 2: Proportionale Stichprobennahme mit geschützter Kategorie

```
Zielproben ÿ 5000
sampledBalanced ÿ Initialisieren Sie einen leeren Datensatz
für jede Spalte in new_categories.columns machen
    categoryRows ÿ Zeilen in new_categories auswählen , in denen Spalte = 1
sampledRows ÿ Beispiel targetSamples Einträge aus categoryRows mit Ersetzung
    für jede Spalte in oneHotEncodedBalanced.columns
        abgetasteteZeilen[Spalte] ÿ 0
Ende für
    sampledRows[Spalte] ÿ 1
    sampledBalanced ÿ sampledRows an sampledBalanced anhängen
Ende für
returniere sampledBalanced
```

4.3. Geschützte Kategorie SMOTE

Die Protected-Category Synthetic Minority Oversampling-Technik (PC-SMOTE)

Die Sampling-Methode ist ein komplexer Prozess, der SMOTE nachahmt, jedoch modifiziert wurde für Stichproben unserer neuen Kategorien statt Klassenbezeichnungen. Bei diesem Ansatz bestand der erste Schritt darin, Ändern Sie SMOTE so, dass eine feste Anzahl von Nachbarn verwendet wird und ein Nachbar zufällig ausgewählt wird für die Interpolation, anstatt alle zu mitteln. Der Pseudocode in Algorithmus 3 zeigt das Verfahren für den PC-SMOTE. Da die Absicht der Methode darin besteht, sie für die neue zu verwenden Kategorie-Sampling, befasst sich nicht direkt mit der Generierung von Klassenbezeichnungen. Daher wird ein neuer Es wird eine Funktion benötigt, die eine neue Klassenbezeichnung für die synthetischen Daten generieren kann. Dazu wird eine neue Es wird eine Funktion definiert, die Klassenbezeichnungen basierend auf der Anzahl der neuen synthetischen Daten generiert generiert und ein vorgewähltes Gleichgewichtsverhältnis zwischen den beiden Klassen. Algorithmus 4 zeigt, wie Unsere neue Funktion generiert Labels für unsere synthetischen Proben. Der Algorithmus bestimmt zunächst die Anzahl der Proben für jede Klasse auf Grundlage des Bilanzverhältnisses und generiert die Probe für jede Klasse erforderlich. Die Klassenbezeichnungen dienen der Umstellung, um eine algorithmische Verzerrung in der Klassen generiert.

Diese beiden Algorithmen werden zu PC-SMOTE kombiniert, wie in Algorithmus 5 gezeigt.

Bei dem Ansatz zur Erzielung eines Gleichgewichts in mehreren Kategorien wird jede einzelne Kategorie iteriert über so dass die Teilmenge der Daten identifiziert wird, die dieser Kategorie zugeordnet ist. Die Nummer Die Anzahl der synthetischen Samples, die zum Erreichen einer vordefinierten Maximalgröße pro Kategorie erforderlich sind, beträgt dann berechnet. Wenn zusätzliche Proben benötigt werden, werden die Daten mit PC-SMOTE generiert,

Machine Translated by Google

die zwischen vorhandenen Datenpunkten und ihren nächsten Nachbarn interpoliert. Gleichzeitig
Eine ausgewogene Verteilung der synthetischen Klassenbezeichnungen wird mit einem festgelegten Balanceverhältnis erzeugt durch
unter Verwendung von Algorithmus 4. Diese synthetischen Merkmale und Bezeichnungen werden dann in die Trainingsteilmenge für jede Kategorie. Der Vorgang wird für alle Kategorien wiederholt, was zu einem ausgeglichener Datensatz. Die Hyperparameter in diesem Algorithmus 5 sind die Anzahl der Nachbarn
und Gleichgewichtsverhältnis.

Algorithmus 3: Benutzerdefinierte Synthetic Minority Oversampling-Technik (SMOTE)

```
1: Prozedur CUSTOMSMOTE(Daten, n_Beispiele)
2:     syntheticSamples ÿ Nullmatrix der Größe (n_samples× Anzahl der Spalten in den Daten)
3: nn ÿ Nächste Nachbarn(n_Nachbarn = 7).fit(Daten)
4:     Nachbarn ÿ nn.kneighbors(Daten, return_distance = False)
5:     für i ÿ 1 bis n_samples mache
6:         sampleIdx ÿ Zufallszahl von 0 bis (Anzahl der Zeilen in den Daten ÿ 1)
7:         nnIdx ÿ zufällige Auswahl aus Nachbarn[sampleIdx, 1 :]
8:         di ff ÿ Daten[nnIdx] ÿ Daten[BeispielIdx]
9:         Gewicht ÿ Zufallszahl aus Gleichverteilung zwischen 0 und 1
10:        syntheticSamples[i] ÿ Daten[SampleIdx] + Gewicht × di ff
11: Ende für
12:    synthetische Proben zurückgeben
13: Prozedur beenden
```

Algorithmus 4: Generieren Sie ausgewogene synthetische Etiketten

```
1: Prozedur GENBALSYNTHLABELS(n_samplesNeeded, balanceRatio)
2:     nClass1 ÿ int(n_benötigte Samples × Bilanzverhältnis)
3:     nKlasse0 ÿ n_benötigte Proben ÿ nKlasse1
4:     syntheticLabels ÿ [0] × nKlasse0 + [1] × nKlasse1
5:     SHUFFLE(syntheticLabels) gibt                    ÿ Die Etiketten nach dem Zufallsprinzip mischen
6:     syntheticLabels zurück
7: Prozedur beenden
```

Algorithmus 5 Geschützte Kategorie SMOTE

```
1: balancedDataList ÿ eine leere Liste initialisieren
2: für jede Kategorie in Kategorien tun
3:     categorySubset ÿ Zeilen in train_data auswählen st 'combined_category' == Kategorie
4:     Funktionen ÿ entferne „Klasse“, „kombinierte_Kategorie“ aus der Kategorie-Untermenge
5:     nSamples ÿ max_sizeÿ Anzahl der Zeilen im Kategorie-Subset
6:     wenn nSamples > 0 dann
7:         syntheticFeatures ÿ PCSmote( features, nSamples)
8:         syntheticLabels ÿ GenBalSynthLabels(nSamples, balanceRatio)
9:         syntheticEigenschaften[ÿ Klasseÿ ] ÿ syntheticLabels
10:        syntheticEigenschaften[ÿ kombinierte_Kategorieÿ ] ÿ Kategorie
11:        categorySubsetBalanced ÿ verketten categorySubset und syntheticFeatures
12:    anders
13:        KategorieSubsetBalanced ÿ KategorieSubset
14: Ende, wenn
15: categorySubsetBalanced an balancedDataList anhängen
16: Ende für
17: balancedData ÿ balancedDataList anhängen und Index zurücksetzen
```

4.4. Geschützte Kategorie ADASYN

Die Methode der geschützten Kategorie ADASYN imitiert die adaptive synthetische Minderheit (ADASYN) Sampling, wird aber leicht modifiziert, um unser Ziel des geschützten Kategorie-Samplings zu erfüllen. Unser PC-ADASYN-Algorithmus ist in Algorithmus 6 dargestellt. Er erweitert ADASYN, indem er sich a

Machine Translated by Google

Dichte statt Klassenungleichgewicht. Genauer gesagt funktioniert diese Funktion, indem sie die nächste Nachbarn der Daten und dann die Berechnung der Dichte der Kategorie jedes Datenpunkts innerhalb seiner unmittelbare Nachbarschaft. Es gewichtet diese Dichten umgekehrt, um Minderheitenkategorien zu priorisieren, wodurch es wahrscheinlicher wird, synthetische Stichproben aus unterrepräsentierten Kategorien zu generieren. Synthetische Stichproben werden durch Interpolation zwischen ausgewählten Datenpunkten und ihren Nachbarn erstellt. ähnlich wie SMOTE, jedoch mit zufälliger Gewichtung zur Variation der Interpolation, wodurch eine vielfältiger synthetischer Datensatz. Dieser Ansatz hilft, das Ungleichgewicht auf Kategorieebene zu beheben und bereichert die Varianz des Datensatzes und verbessert potenziell die Robustheit und Fairness von ML-Modellen trainiert. Da diese Sampling-Methode durch Interpolation auch neue Samples erzeugt, Algorithmus 4 wird verwendet, um Klassenbezeichnungen für die neuen synthetischen Proben zu generieren.

Algorithmus 6 PC-ADASYN für kategoriebasiertes Balancing

```
1: Prozedur PCADASYNCATEGORIES (Daten, Beschriftungen, n_benötigte Proben, n_Nachbarn)
2:     n_Nachbarn ÿ n_Nachbarn + 1 3: nn ÿ                ÿ Einbeziehung des Datenpunkts selbst
NächsteNachbarn(n_Nachbarn).fit(Daten)
4:     Distanzen, Indizes ÿ nn.kneighbors(data)
5:     Dichten ÿ Null-Array mit Länge (Daten)
6:     für i ÿ 0 bis Länge(Daten) ÿ 1 mache
7:         aktuelle_Kategorie ÿ Beschriftungen[i]
8:         Nachbarindizes ÿ Indizes[i][1 :] Dichten[i] ÿ          ÿ Überspringen Sie den Selbstindex
9:         SUMME(Beschriftungen[Nachbarindizes] == aktuelle_Kategorie)
10: Ende für
11:    Gewichte ÿ 1/(Dichten + 1) Gewichte          ÿ Addiere 1, um eine Division durch Null zu verhindern
12:    ÿ Gewichte/SUMME(Gewichte) synthetische                ÿ Gewichte normalisieren
13:    Proben ÿ leere Liste
14:    sampleIndices ÿ zufällige Auswahl mit Ersetzung aus length(data) unter Verwendung von Gewichten
15:    für jeden idx in sampleIndices
16:        BasisIdx ÿ idx
17:        Nachbar-Idx ÿ RANDOMCHOICE(Indizes[Basis-Idx][1 :])
18:        di ff ÿ Daten[Nachbar-Idx] ÿ Daten[Basis-Idx]
19:        syntheticheProbe ÿ Daten[BasisIdx] + RANDOM() × di ff
20:        SyntheticSample an SyntheticSamples anhängen
21: Ende für
22:    Array zurückgeben (synthetische Samples)
23: Ende der Prozedur
```

Die Algorithmen 4 und 6 werden kombiniert, um unser geschütztes ADASYN-Sampling zu bilden. wie in Algorithmus 7 gezeigt. Für jede Kategorie werden die dieser Kategorie entsprechenden Daten isoliert und das Größendefizit relativ zur größten Kategorie berechnet. Wenn zusätzliche Proben benötigt werden, wird die PC-ADASYN-Methode angewendet, die synthetische Merkmale erzeugt, die die Verteilungseigenschaften der Kategorie zu berücksichtigen. Diese Merkmale werden dann ergänzt mit synthetisch generierten Labels, wobei ein vordefiniertes Klassengleichgewichtsverhältnis beibehalten wird. Der Prozess korrigiert nicht nur Kategorieungleichgewichte, sondern bereichert auch den Datensatz, möglicherweise Verbesserung der Vorhersagegenauigkeit und Fairness von Modellen, die mit diesen Daten trainiert werden.

4.5. Datensatz- und Hyperparameter-Tuning

Um die vier Stichprobenverfahren zu testen, wurde ein Klassifikator benötigt, um die Auswirkungen auf Fairness und Leistung. Letztlich ist der Typ des Klassifikators nicht direkt relevant, da die Ziel ist es, den Kompromiss zwischen Fairness und Leistung zu mildern, anstatt den besten Klassifikator zu finden. Aus diesem Grund haben wir uns für die Verwendung univariater Entscheidungsbäume auf Basis von CART [56] entschieden , da diese Robustheit gegenüber Rauschen und fehlenden Daten. Die spezifische Implementierung, die wir gewählt haben, war aus der sklearn-Bibliothek (Version 1.5.1) [57]. Darüber hinaus ermöglichen Entscheidungsbäume Steuern Sie die Stärke der Anpassung, indem Sie die Baumtiefe des gelernten Baums festlegen. Dies ermöglicht es uns, Vergleichen Sie Fairness und Leistung auf verschiedenen Anpassungsebenen.

Machine Translated by Google

Der Prozess beginnt mit der Kombination geschützter Kategorien innerhalb jedes Datensatzes, der Anwendung einer One-Hot-Kodierung zum Erstellen neuer Multikategoriefunktionen und der anschließenden Durchführung einer Beschriftungsk

Die Datensätze wurden dann mithilfe einer geschichteten 10-fachen Kreuzvalidierung aufgeteilt, um eine repräsentative Verteilung der Klassen in jeder Faltung sicherzustellen. Für jede Faltung wurde das Training mit Stichprobendaten unter Verwendung der zuvor beschriebenen Methoden durchgeführt, während die Klassifizierung mit den entsprechenden nicht gesampelten Testsätzen getestet wurde. Die Konsistenz beim Modelltraining wurde durch Anwenden identischer Baumtiefen auf alle Stichprobenmethoden gewahrt. Die bereitgestellten Ergebnisse sind Durchschnittswerte der 10-fachen Durchläufe mit den entsprechender

Algorithmus 7 Geschützte Kategorie ADASYN

```
1: balancedDataList ÿ Initialisiere eine leere Liste. 2: Für
jede Kategorie in Kategorien führe
3:     categorySubset aus ÿ Wähle aus new_data3 st 'combined_category' == category f eatures ÿ
4:     Entferne 'class', 'combined_category' aus categorySubset categoryLabels ÿ Extrahiere
5:     'combined_category' aus categorySubset nSamplesNeeded ÿ max_size minus
6:     Anzahl der Zeilen in categorySubset, wenn nSamplesNeeded > 0, dann syntheticFeatures
7:     ÿ PCAdasyn(f eatures,
8:         categoryLabels, nSamplesNeeded) syntheticLabels ÿ GenBalSynthLabels(nSamplesNeeded,
9:         balanceRatio) syntheticFeatures[ ÿ classÿ ] ÿ syntheticLabels syntheticFeatures[ ÿ
10:        combined_categoryÿ ] ÿ Kategorie categorySubsetBalanced
11:        ÿ verkette categorySubset mit syntheticFeatures, sonst
12:        categorySubsetBalanced ÿ categorySubset 15: Ende wenn 16: hänge categorySubsetBalanced
13:    an
14:        balancedDataList an 17: Ende für 18: balancedData ÿ
verkette
balancedDataList und setze Index zurück
```

Für unsere Analyse **wurden drei Datensätze aus dem UCI-Repository [58]** ausgewählt : **der Datensatz „Adult Income“ [59], der Datensatz „German Credit “ [60] und der Datensatz „Correctional Offender Management Profiling for Alternative Sanctions“ (COMPAS). Der Datensatz „Adult Income“ soll vorhersagen, ob eine Person mehr als 50.000 US-Dollar verdient. Er enthält acht kategorische und vier numerische Attribute mit geschützten Variablen entsprechend Alter (jung oder erwachsen), Geschlecht (männlich oder weiblich) und Rasse (weiß oder andere). Der Datensatz „Adult Income“ wurde der UCI 1996 gespendet. Der Datensatz „German Credit“, der zur Vorhersage der Kreditwürdigkeit verwendet wird, umfasst 20 kategorische und zwei numerische Attribute mit geschützten Variablen für Alter und Geschlecht. Der Datensatz „German Credit“ wurde der UCI 1994 gespendet. Der COMPAS-Datensatz [61], der die Rückfallquoten in den USA bewertet, umfasst sechs kategorische und sechs numerische Merkmale mit geschützten Variablen für Alter, Rasse und Geschlecht. Der Datensatz wurde 2018 veröffentlicht. Diese Datensätze wurden ausgewählt, weil sie den neuesten Stand der Technik zur Messung von Voreingenommenheit und Diskriminierung darstellen und in anderen Studien zu algorithmischer Voreingenommenheit und Fairness weit verbreitet sind (siehe Abschnitt 3). Außerdem haben die Datensätze unterschiedliche Größen, von klein bis groß, was sie zum Testen unserer Stichprobenmethoden geeignet macht.**

Die **Hyperparameter-Optimierung wurde mithilfe der Grid-Suche [62]** durchgeführt , **um ein breites Spektrum an Parametern zu untersuchen. Ergänzt wurde dies durch visuelle Bewertungen, um optimale Einstellungen zu ermitteln, die Equalized Odds Difference (EOD) und Genauigkeit ausbalancieren. Für den Datensatz „Adult Income“ umfassten die optimalen Hyperparameter eine maximale Baumtiefe von 3 und für PC-SMOTE und PC-ADASYN eine Nearest-Neighbor-Einstellung von 5 mit einem ausgeglichenen Verhältnis von 0,34. Diese Parameter waren für den Datensatz „German Credit“ ähnlich effektiv. Für den COMPAS-Datensatz war eine maximale Baumtiefe von 2 für alle Stichprobenverfahren optimal. PC-SMOTE und PC-ADASYN wurden auf die Nearest-Neighbor-Einstellung von 3 und ein ausgeglichenes Verhältnis von 0,60 eingestellt.**

Machine Translated by Google

5. Ergebnisse

Die vier Stichprobenstrategien wurden auf die drei oben beschriebenen Datensätze angewendet und evaluierten ihre Auswirkungen mithilfe eines einfachen univariaten Entscheidungsbaum-Klassifikators. Die Ergebnisse in Die Tabellen 1–3 zeigen erhebliche Unterschiede in der Modellleistung bei fünf Sampling-Strategien: kein Sampling, Oversampling, proportionales Sampling, PC-SMOTE und PC-ADASYN ein unsere drei Datensätze. Jede Methode wurde anhand von Genauigkeit, Makro F1, Equalized Odds Difference (EOD) und Statistische Parität (SP). Die Ergebnisse wurden in Genauigkeit gemessen und Makro F1, da diese beiden Metriken die beliebtesten Klassifizierungsmetriken sind. Außerdem Durch die Beschränkung der Metrik auf zwei werden die Ergebnisse für die statistische Analyse vergleichbar.

Um die statistische Signifikanz der Ergebnisse jeder Methode zu ermitteln, verwendeten wir den Friedman Test, eine nichtparametrische Alternative zur einfaktoriellen ANOVA mit wiederholten Messungen. Da der Friedman-Test signifikante Ergebnisse lieferte, führten wir den Nemenyi-Post-hoc-Test durch. Test. Dieser Test wird verwendet, um paarweise Vergleiche zwischen den Methoden zur Ermittlung durchzuführen welche Methoden sich statistisch voneinander unterscheiden. Der Nemenyi-Test ist vorteilhaft in dieser Einstellung, weil es mehrere Vergleiche berücksichtigt, ohne normale Verteilungen und bieten somit eine robuste Möglichkeit zum Verständnis spezifischer paarweiser Unterschiede.

Für den Datensatz „Einkommen Erwachsener“ (Tabelle 1) ergab die No-Sampling-Methode eine Genauigkeit von 0,82, was eine hohe Vergleichsbasis darstellt. Es zeigte jedoch eine leicht verzerrte Vorhersage mit einem EOD von 0,36 und minimaler Disparität in den Vorhersageraten (SP = 0,02). Im Gegensatz dazu behielt Oversampling die gleiche Genauigkeit bei, verringerte aber die Makro-F1 leicht auf 0,65, was auf mögliche Überanpassungsprobleme hindeutet und gleichzeitig die Fairness verschlechtert (EOD = 0,66) und zunehmende Disparität in den Vorhersageraten (SP = 0,09). Die proportionale Stichprobennahme verringerte Genauigkeit auf 0,79, verbesserte aber das Makro F1 auf 0,79, was auf eine bessere Balance zwischen Präzision und Rückruf. Allerdings erhöhte sich der SP signifikant auf 0,71, was auf eine erhebliche Ungleichheit bei den positiven Vorhersageraten, was Zweifel an der Fairness des Modells aufkommen lässt. Zwei benutzerdefinierte Ansätze für SMOTE und ADASYN wurden speziell entwickelt, um auf diesen Metriken. PC-SMOTE zeigte eine mäßige Leistung mit einer Genauigkeit von 0,81 und ein verbesserter EOD von 0,25, was auf eine verbesserte Fairness gegenüber einfachem Oversampling hindeutet und die No-Sampling-Methode. Allerdings wurde immer noch ein niedrigerer Makro-F1-Wert (0,63) aufgezeichnet, was darauf hindeutet, Falschdarstellungsprobleme bei der Generierung synthetischer Daten. PC-ADASYN erwies sich als ausgewogenster Ansatz, der eine hohe Genauigkeit (0,82) und eine bessere Handhabung der Klasse beibehält Ungleichgewichte, mit einer moderaten Verbesserung der Fairness (EOD = 0,28) und einer kontrollierten Zunahme der Vorhersageraten-Disparität (SP = 0,09). Insgesamt ist die Basisgenauigkeit statistisch deutlich besser als die proportionale Stichprobe, aber nicht statistisch signifikant, da im Vergleich zu den anderen Stichprobenverfahren. Für das Makro F1 ist die proportionale Stichprobennahme statistisch signifikant besser als andere Probenahmemethoden. Für den EOD ist PC-ADASYN ist statistisch besser als andere Stichprobenverfahren, während die No-Sampling-Methode SP statistisch signifikant besser als andere Stichprobenmethoden.

Tabelle 1. Ergebnisse der Stichprobenmethoden für den Datensatz „Einkommen Erwachsener“ mit 95ÿ%-Konfidenzintervallen.

Probenahme Methode	Genauigkeit	Makro F1	EOD	SP
Keine Stichprobennahme	0,82 ± 0,00	0,66 ± 0,01	0,36 ± 0,18	0,02 ± 0,00
Überstichprobe	0,82 ± 0,02	0,65 ± 0,06	0,66 ± 0,2	0,09 ± 0,02
Eigenstichprobe	0,79 ± 0,05	0,79 ± 0,06	0,46 ± 0,12	0,71 ± 0,10
PC-SMOTE	0,81 ± 0,05	0,63 ± 0,07	0,25 ± 0,21	0,07 ± 0,00
PC-ADASYN	0,82 ± 0,04	0,64 ± 0,07	0,28 ± 0,19	0,09 ± 0,02

Auch die Ergebnisse der Experimente mit dem deutschen Kreditdatensatz zeigen unterschiedliche Auswirkungen jeder Stichprobenstrategie, insbesondere im Hinblick auf Fairness und Genauigkeit, wie in Tabelle 2 dargestellt. Ohne Stichproben erreichte das Basismodell eine Genauigkeit von 0,72, wies aber erhebliche Verzerrung in der Vorhersage, mit einem EOD von 0,88, was auf eine erhebliche Fehlerdisparität hinweist Raten zwischen Gruppen in den geschützten Attributen. Die Implementierung von Oversampling hielt Genauigkeit bei gleichzeitiger Verbesserung des Makro-F1 auf 0,68 und deutlicher Reduzierung des EOD auf 0,37, wenn auch bei

Machine Translated by Google

die Kosten für die Erhöhung des SP auf 0,35, was einen möglichen Kompromiss zwischen unterschiedlichen Fairness-Maßnahmen. Die proportionale Stichprobennahme reduzierte die Genauigkeit leicht auf 0,68, erreichte aber die beste F1-Score von 0,69. Es senkte auch den EOD auf 0,32, was darauf hindeutet, dass es die Vorhersage effektiv ausgleicht Qualität mit Fairness. PC-SMOTE zeigt eine Verbesserung der Genauigkeit mit 0,73, aber eine geringere Makro-F1 von 0,55; die Fairness des Modells zeigt eine enorme Verbesserung gegenüber der Baseline mit ein EOD von 0,15 und ein moderater SP von 0,1. PC-ADASYN zeigt eine ähnliche Genauigkeit wie die Basislinie bei 0,72, allerdings mit dem niedrigsten Makro-F1 von 0,48, was auf einen möglichen Kompromiss hindeutet in Präzision und Rückruf. Das Modell weist jedoch die beste Fairnessvorhersage auf mit ein EOD von 0,13 und ein SP von 0,06. Insgesamt zeigt das Ergebnis, dass die Genauigkeit von No-Sampling ist im Vergleich zu anderen Stichprobenverfahren mit Ausnahme von proportionalen Stichproben nicht statistisch signifikant während für den EOD die Ergebnisse aller Probenahmemethoden statistisch signifikant sind in Vergleich mit der No-Sampling-Methode.

Tabelle 2. Ergebnisse der Stichprobenmethoden für den deutschen Kreditdatensatz mit 95 %-Konfidenzintervallen.

Probenahme Methode	Genauigkeit	Makro F1	EOD	SP
Keine Abtastung	0,72 ± 0,03	0,62 ± 0,06	0,88 ± 0,10	0,07 ± 0,01
Überabtastung	0,72 ± 0,05	0,68 ± 0,04	0,37 ± 0,16	0,35 ± 0,11
Eigenabtastung	0,69 ± 0,04	0,69 ± 0,07	0,32 ± 0,26	0,29 ± 0,10
Abtastung PC-SMOTE	0,73 ± 0,05	0,55 ± 0,09	0,15 ± 0,09	0,1 ± 0,02
PC-ADASYN	0,72 ± 0,03	0,48 ± 0,11	0,13 ± 0,02	0,06 ± 0,00

Tabelle 3 zeigt die Ergebnisse unserer Erfahrungen mit dem COMPAS-Datensatz. Diese Ergebnisse offenbaren erhebliche Unterschiede in der Modellleistung zwischen den verschiedenen Sampling-Strategien. Der Baseline-Ansatz ohne Stichproben erreichte eine Genauigkeit und einen Makro-F1-Score von 0,89, zeigte aber größere Unterschiede in den Fairnessmetriken, mit einer ausgeglichenen Quotendifferenz (EOD) von 0,39 und einer statistischen Parität (SP) von 0,29. Dies unterstreicht mögliche Verzerrungen, die unangepasste Modelle können gegenüber geschützten Gruppen eine Tendenz aufweisen. Die Anwendung von Oversampling leicht verbesserte Genauigkeit auf 0,90, aber auch verbesserte Fairness deutlich, wodurch EOD auf 0,26. Dies deutet darauf hin, dass es effektiv ist, Ergebnisunterschiede zu verringern, ohne Kompromisse einzugehen SP. Umgekehrt, proportionale Abtastung, während die Genauigkeit und Makro F1 auf 0,90 und 0,91 bzw. erreichte auch einen EOD von 0,26 und verbesserte sich damit gegenüber dem Ausgangswert, während auch ein höherer SP von 0,36, was auf eine mögliche Zunahme der Ungleichheit positiver Ergebnisse über alle Gruppen hinweg. PC-SMOTE und PC-ADASYN, mit identischen Werten in Genauigkeit, Makro F1 und SP konnten die Fairnessverbesserungen mit einem EOD von 0,30 aufrechterhalten, obwohl diese Methoden auch den SP auf 0,47 erhöhten. Insgesamt zeigen die Ergebnisse, dass sowohl die Genauigkeit und EOD unserer Probenahmemethoden sind statistisch signifikant besser als die No-Sampling-Methode.

Tabelle 3. Ergebnisse der Stichprobenmethoden im COMPAS-Datensatz mit 95 %-Konfidenzintervallen.

Probenahme Methode	Genauigkeit	Makro F1	EOD	SP
Keine Abtastung	0,89 ± 0,04	0,89 ± 0,04	0,39 ± 0,15	0,29 ± 0,17
Oversample	0,90 ± 0,03	0,90 ± 0,05	0,26 ± 0,14	0,25 ± 0,07
Prop. Sample	0,90 ± 0,05	0,91 ± 0,04	0,26 ± 0,12	0,36 ± 0,10
PC-SMOTE	0,91 ± 0,02	0,91 ± 0,02	0,30 ± 0,11	0,47 ± 0,19
PC-ADASYN	0,91 ± 0,02	0,91 ± 0,03	0,30 ± 0,13	0,47 ± 0,21

6. Diskussion

Die Ergebnisse der Experimente mit den drei Datensätzen belegen, dass geschützte Kategorien Sampling kann die Modellgerechtigkeit deutlich verbessern, oft ohne die Vorhersagegenauigkeit wesentlich zu beeinträchtigen. In einigen Fällen wurden auch Verbesserungen bei der Genauigkeit und der Makro-F1 demonstriert. Konzentriert man sich auf die Ergebnisse des Adult Income-Datensatzes, so

Machine Translated by Google

zeigte bemerkenswerte Verbesserungen bei EOD und hielt moderate SP-Niveaus aufrecht. Die Wirksamkeit dieser Methoden ist größtenteils auf ihre ausgefeilten Interpolationstechniken zurückzuführen. Eine visuelle Untersuchung der ohne Sampling und PC-ADASYN generierten Entscheidungsbäume liefert beispielsweise aufschlussreiche Kontraste. Beispiele aus einer einzelnen repräsentativen Falte sind in Abbildung 1 bzw. 2 dargestellt. Der ohne Sampling gelernte Entscheidungsbaum wählte seine Wurzel mit einem Merkmal, das eng mit geschützten Attributen verbunden ist und somit als Proxy-Attribut fungierte. Dies führte zu einer ausgeprägten Vorh
Im Gegensatz dazu begann der mit PC-ADASYN generierten Daten trainierte Entscheidungsbaum mit einer Funktion, die Vorhersagen sehr gut verallgemeinerte und Verzerrungen milderte, was durch eine deutliche Verbesserung der Modellfairness und eine höhere Gini-Unreinheit belegt wurde, was auf eine reinere anfängliche Aufte

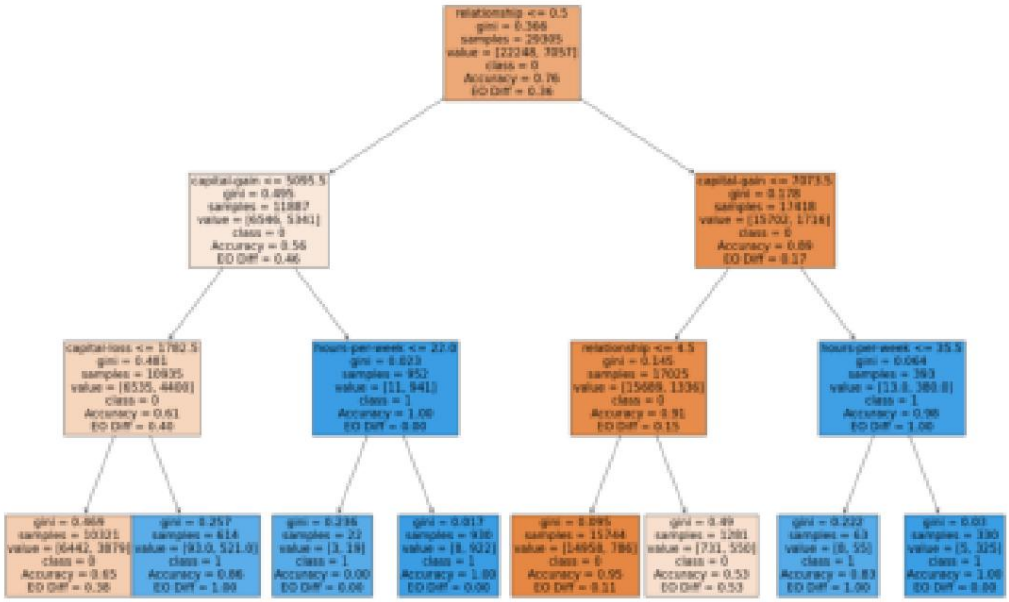


Abbildung 1. Beispiel eines Entscheidungsbaums, der ohne Stichprobennahme auf das Einkommen Erwachsener trainiert wurde.

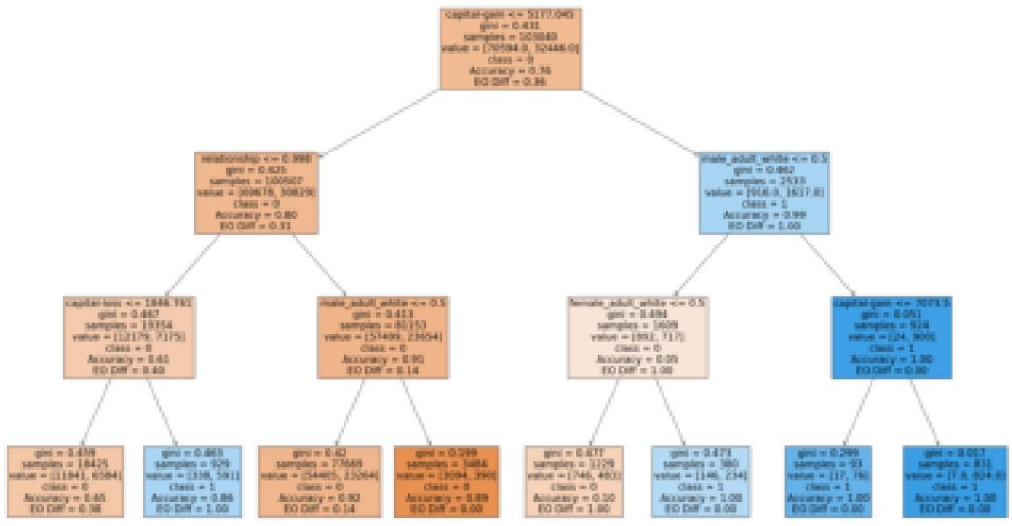


Abbildung 2. Entscheidungsbaum von PC-ADASYN zum Einkommen von Erwachsenen.

Machine Translated by Google

6.1. Vergleich von Fairness und Leistung

Beim Vergleich von Oversampling und proportionalem Sampling waren die Ansätze der Methoden zur Vergrößerung der Stichprobengröße durch Duplizierung vorhandener Datenzeilen unkompliziert und führten zu keinen wesentlichen Verbesserungen bei EOD. Dieses Ergebnis ist nachvollziehbar, da diese Methoden dazu neigen, vorhandene Verzerrungen zu replizieren, was Fairnessprobleme möglicherweise verschärfen kann, anstatt sie zu lindern. Dies wird insbesondere deutlich, wenn man den Datensatz „Erwachseneneinkommen“ betrachtet, bei dem die Klassen extrem unausgewogen sind. Diesen naiven Replikationsstrategien fehlt die Interpolationskapazität von PC-SMOTE und PC-ADASYN, um Stichproben in der Nähe von Entscheidungsgrenzen anzupassen, was für die Minderung der Verzerrung im Datensatz entscheidend ist. Im Gegensatz dazu erweitern die von SMOTE und ADASYN verwendeten Interpolationsstrategien den Datensatz und erhöhen seine Vielfalt. Dies ist besonders effektiv bei Stichproben in der Nähe von Entscheidungsgrenzen, bei denen leichte Verschiebungen der Merkmale die Fairness der Vorhersagen erheblich beeinträchtigen können. Durch Interpolation zwischen den Stichproben verschieben SMOTE und ADASYN diese Randstichproben effektiv in Richtung gleichmäßigerer Bereiche des Merkmalsraums. Auf diese Weise begegnen und reduzieren sie Verzerrungen effektiver als Methoden, die das Stichprobenvolumen erhöhen, ohne die Datenstruktur zu ändern. Die Klassengenerierungsfunktion (Algorithmus 4) trägt auch dazu bei, die allgemeine Klassenverteilung zu verbessern. Die Ergebnisse unserer PC-SMOTE- und PC-ADASYN-Studie zum Einkommen von Erwachsenen sind ebenfalls den in **[41] erzielten Ergebnissen überlegen , wo eine Genauigkeit von 0,59 und ein SP von 0,17 erzielt wu** Darüber hinaus sind die Ergebnisse von PC-SMOTE und PC-ADASYN den **in [40] erzielten Ergebnissen überlegen , wo ein EOD von 0,89 und eine etwas bessere Genauigkeit von 0,84 erreicht wurden.**

Bei der Untersuchung der Ergebnisse des deutschen Kreditdatensatzes haben wir einen ähnlichen Trend wie beim Erwachseneneinkommensdatensatz festgestellt: Die No-Sampling-Methode weist eine sehr hohe Verzerrung in Bezug auf EOD auf. Der niedrige SP von 0,07 weist auf minimale Unterschiede bei den positiven Vorhersageraten zwischen den Gruppen hin, aber dies allein ist kein gutes Maß für die Fairness, da die bevorzugte Gruppe mehr Stichproben hat als die benachteiligte. Dies hat zur Folge, dass die Berechnung des SP verzerrt wird, da nur positive Entscheidungen in jeder Gruppe gezählt werden, die von der Stichprobengröße beeinflusst werden. Eine Erkenntnis ist, wie wichtig es ist, mehrere Fairnessmetriken einzusetzen, um die Auswirkungen eines Modells auf alle Beteiligten umfassend zu betrachten. Beim Oversampling haben wir eine Verbesserung des EOD mit einem ähnlichen SP festgestellt; dies zeigt, dass eine Erhöhung der Anzahl der Stichproben für jedes der geschützten Attribute der Multikategorie die Fairness in Bezug auf EOD verbessert. Darüber hinaus spiegelt der aktualisierte SP wider, wie es aussehen würde, wenn für jede Multikategorie eine gleichmäßigere Anzahl von Stichproben vorhanden wäre, anders als in der Baseline, wo die bevorzugte Gruppe fünfmal mehr Stichproben hat als die unbegünstigte Gruppe. Die Genauigkeit der proportionalen Stichprobennahme sinkt, da die nach der Hyperparameter-Optimierung ausgewählte Baseline-Anzahl der Stichproben nicht ausreichte, damit das Modell den ungesampelten Testsatz verallgemeinern konnte, was zu einer Überanpassung führte. Die Überanpassung wurde durch Berücksichtigung der Trainingsgenauigkeit bestätigt. Interessanterweise tauscht das Modell nicht wie andere Modelle Genauigkeit gegen Rückruf ein, und dies verleiht der proportionalen Stichprobennahme die höchste Makro-F1.

In Bezug auf die Fairness stellten wir einen Anstieg des EOD im Vergleich zum Ausgangswert und zur Überabtastung fest . Dies liegt daran, dass jede Mehrfachkategorie auf denselben Ausgangswerten dargestellt wird. Dies kann die Fairness des Modells verbessern, da das Modell nun ein größeres Bild der Kategorien hat und bessere Vorhersagen und letztlich fairere Entscheidungen trifft. PC-SMOTE und PC-ADASYN spielen eine entscheidende Rolle bei der deutlichen Reduzierung der Verzerrung bei Modellvorhersagen.
Diese Konsistenz bestätigte die Robustheit dieser Methoden über verschiedene Datensätze hinweg. Bemerkenswerterweise geht keine der Methoden Kompromisse bei der Genauigkeit ein, während beide die Fairness verbessern, was ihre Wirksamkeit bei der Handhabung der Kompromisse zeigt, die typischerweise mit prädiktiver Diese Ergebnisse zeigen die starken Interpolationsfähigkeiten, die PC-SMOTE und PC-ADASYN innewohnen. Diese Methoden verteilen die Stichproben effektiv innerhalb des Merkmalsraums neu, insbesondere indem sie die Stichproben in unterprivilegierten Regionen von negativen zu positiveren Entscheidungsgrenzen verschieben . Solche Anpassungen sind entscheidend, um verzerrte Ergebnisse abzumildern und die Gerechtigkeit in automatisierten Entscheidungsprozessen zu fördern. Die Makro-F1 sinkt in beiden Modellen im Vergleich zur Basislinie, da eine höhere Anzahl von Stichproben erforderlich ist, damit das Modell bei der Generalisierung besser abschneidet, was dieser Datensatz nicht unterstützt. Insbesondere hat der Datensatz 700 Stichproben für Klasse 0 und 300 Stichproben für Klasse 1, was bedeutet, dass der Testsatz

Machine Translated by Google

hat nur 30 Proben für Klasse 1. Diese kleine Anzahl von Proben führte dazu, dass beide Modelle in Klasse 1 Rückruf gegen Präzision eintauschten. Insbesondere sahen wir einen niedrigen Rückruf für Klasse 1, was letztendlich zu einem niedrigen F1-Score für Klasse 1 führte, und da Makro F1 die beiden F1-Scores mittelt und sie gleich behandelt, wirkt sich dies auf die Leistung beider Modelle in Makro F1 aus. Insgesamt ergeben die beiden Modelle ein faireres Modell mit guter Genauigkeit im Vergleich zur Basislinie und den beiden anderen Stichprobenmethoden.

Die Auswertung des COMPAS-Datensatzes bestätigt die Wirksamkeit unserer Stichprobenmethoden weiter . Die deutlichen Muster, die sich herauskristallisieren, stimmen mit denen überein, die in den Datensätzen „Erwachseneneinkommen“ und „Deutscher Kredit“ beobachtet wurden, was die Robustheit unserer Ergebnisse unterstreicht. Insbesondere haben Oversampling- und proportionale Stichprobenverfahren erhebliche Verbesserungen bei der ausgeglichenen Quotendifferenz (EOD) und der Genauigkeit gezeigt, während Oversampling auch bei SP deutlich verbessert. Diese Verbesserung ist wahrscheinlich auf die einzigartige Zusammensetzung und Ausgewogenheit innerhalb des COMPAS-Datensatzes zurückzuführen, im Gegensatz zu den anderen Datensätzen, in denen die Klassen unausgewogen sind. Der Erfolg von Oversampling und proportionalem Sampling in diesem Zusammenhang ist auf die ausgewogene Natur des Datensatzes zurückzuführen, die eine wiederholte Duplizierung vorhandener Zeilen (von diesen Methoden verwendete Stichprobenverfahren) ermöglicht, um den Datensatz zu verbessern, ohne eine signifikante Schiefe zugunsten einer bestimmten Klasse einzuführen. Diese Methode erweitert effektiv die Darstellung aller Klassen und der geschützten Attribute in ausgewogener Form, wodurch diese Techniken besonders effektiv für Datensätze sind, bei denen die Merkmalsdomänen gleichermaßen zu Vorhersagen beitragen und bei denen die anfänglichen Klassenverteilungen nicht unter einem starken Ungleichgewicht leiden . Dies kann weiter anhand ihres Makros F1 verifiziert werden, da keines der Modelle Präzision gegen Rückruf eintauscht . Die Verbesserung von SP beim Oversampling kann auf die höhere Anzahl von Samples beim Oversampling im Vergleich zum proportionalen Sampling zurückgeführt werden. In Bezug auf PC-SMOTE und PC-ADASYN zeigen diese Algorithmen eine Verbesserung gegenüber der Baseline sowohl bei der Genauigkeit als auch beim EOD. Diese Trends folgen denen der vorherigen Ergebnisse. Ein bemerkenswerter Aspekt dieser Ergebnisse ist der große Rückgang von SP, der auf unser neues Label zurückgeführt werden kann, das generiert wurde, um den Datensatz in Richtung der negativen Klasse zu verzerre

6.2. Einfluss der Baumtiefe auf Fairness und Genauigkeit

In dieser Studie wurde auch der Einfluss der Tiefe des Entscheidungsbaums auf die Modellleistung untersucht, insbesondere wurde untersucht, wie Variationen der Baumtiefe die Genauigkeit und EOD beeinflussen. Das Verständnis des Tiefeneffekts ist wichtig, da es Einblick in die Effekte von Unteranpassung bis Überanpassung bietet und hilft, den optimalen Komplexitätsgrad zu ermitteln, bei dem sowohl Genauigkeit als auch Fairness maximiert werden. Zunächst durfte der Entscheidungsbaum ohne Einschränkungen bis zu seiner vollen Tiefe wachsen, die im Durchschnitt etwa 30 Zweige umfasste. Der Baum wurde dann visuell untersucht, um die maximale Tiefe ohne die nicht-aufteilenden Zweige zu ermitteln. Um die Effekte der Baumtiefe systematisch zu analysieren, durfte die maximale Tiefe der Bäume zwischen 1 und 30 variieren. Jede Tiefengrenze wurde mithilfe einer zehnfachen Kreuzvalidierung ausgewertet, um die Robustheit und Generalisierbarkeit der Erkenntnisse sicherzustellen.

Für jede Konfiguration der Baumtiefe wurden Genauigkeit und EOD im Testset gemessen . Zusätzlich wurden 95%-Konfidenzintervalle für die Metriken über die zehn Faltungen hinweg berechnet. Diese statistische Analyse hob die Tiefe hervor, in der der Entscheidungsbaum den Kompromiss zwischen Genauigkeit und Fairness ausbalancierte, während gleichzeitig der zugrunde liegende statistische Bias-Varianz-Kompromiss berücksichtigt wurde. Auf diese Weise war es möglich, den „Sweet Spot“ zu ermitteln – einen heiklen Punkt, an dem der Entscheidungsbaum eine hohe Vorhersagegenauigkeit beibehält, ohne Kompromisse bei der Fairness einzugehen, und so dem oft zitierten Komp

In den Abbildungen 3–7 werden die Diagramme für Genauigkeit und EOD gegenüber der maximalen Tiefe für jede der fünf Stichprobenmethoden im Datensatz „Erwachseneneinkommen“ dargestellt.

Machine Translated by Google

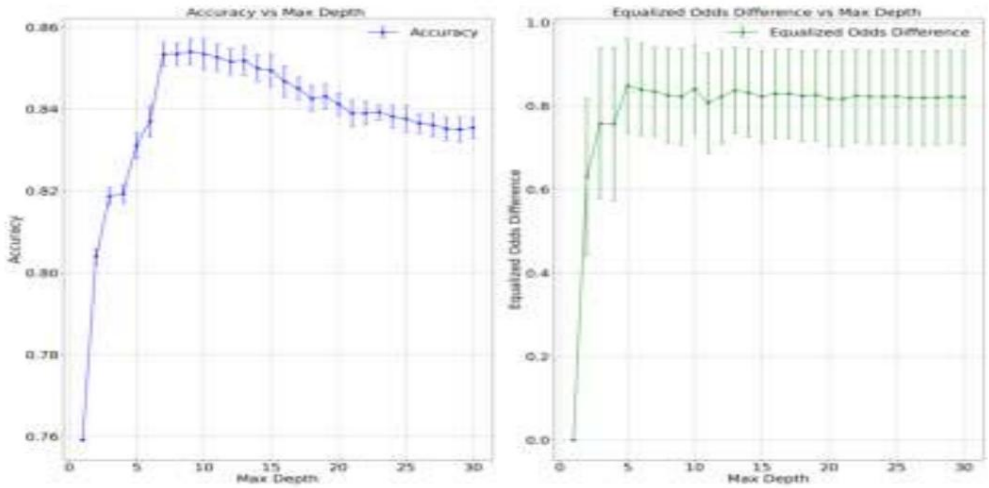


Abbildung 3. Diagramme des Erwachseneneinkommens ohne Stichprobennahme, die Genauigkeit und EOD mit 95 %-Konfidenzintervallen gegenüber der maximalen Baumtiefe im Bereich von 1 bis 30 zeigen.

Basierend auf Ergebnissen wie denen in den Abbildungen 3 und 4 ist sowohl bei der No-Sampling- als auch bei der Oversampling-Methode eine bemerkenswerte anfängliche Genauigkeitssteigerung mit zunehmender maximaler Tiefe zu verzeichnen. Beide Methoden weisen jedoch ab einer Tiefe von 10 einen Genauigkeitsrückgang auf, was auf den Beginn einer Überanpassung hindeutet. Entsprechend nimmt die EOD mit zunehmender Tiefe bis etwa Tiefe 10 stark ab, jenseits derer sie sich stabilisiert. Dieses Muster weist darauf hin, dass tiefere Bäume zwar zunächst die Fairness verbessern, schließlich aber einen Schwellenwert erreichen, jenseits dessen keine weiteren Verbesserungen mehr zu beobachten sind. Wenn man sich daran erinnert, dass das Fairnessziel darin bestand, die EOD zu minimieren, ist eine wichtige Beobachtung, dass die Festlegung der maximalen Tiefe zwischen drei und fünf ein optimales Gleichgewicht zwischen

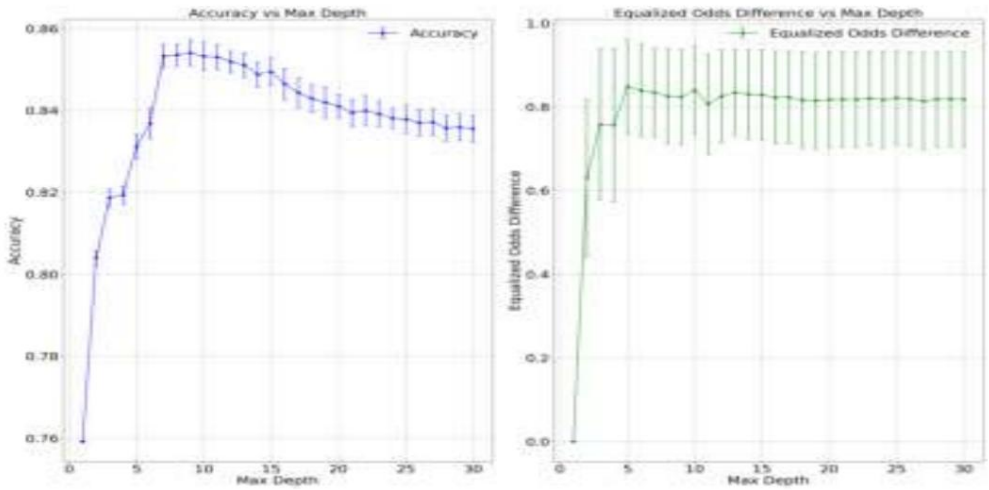


Abbildung 4. Diagramme des Erwachseneneinkommens unter Verwendung von Oversampling, die Genauigkeit und EOD mit 95 %-Konfidenzintervallen gegenüber der maximalen Baumtiefe im Bereich von 1 bis 30 zeigen.

Betrachtet man die in Abbildung 5 gezeigten Ergebnisse, so erhöht sich die Genauigkeit bei der proportionalen Stichprobenmethode kontinuierlich mit zunehmender Baumtiefe und erreicht bei einer Tiefe von etwa 26 ihren Höhepunkt. Umgekehrt steigt der EOD zunächst an, bevor er abnimmt und sich bei einer Tiefe von etwa 15 stabilisiert. Die in der EOD-Metrik beobachteten breiten Konfidenzintervalle deuten auf eine erhebliche Variabilität der Fairnessergebnisse hin. Dieses Ergebnis unterstreicht, wie wichtig es ist, eine Tiefe auszuwählen, die die Variabilität der Fairness minimiert und gleichzeitig die Genauigkeit

Machine Translated by Google

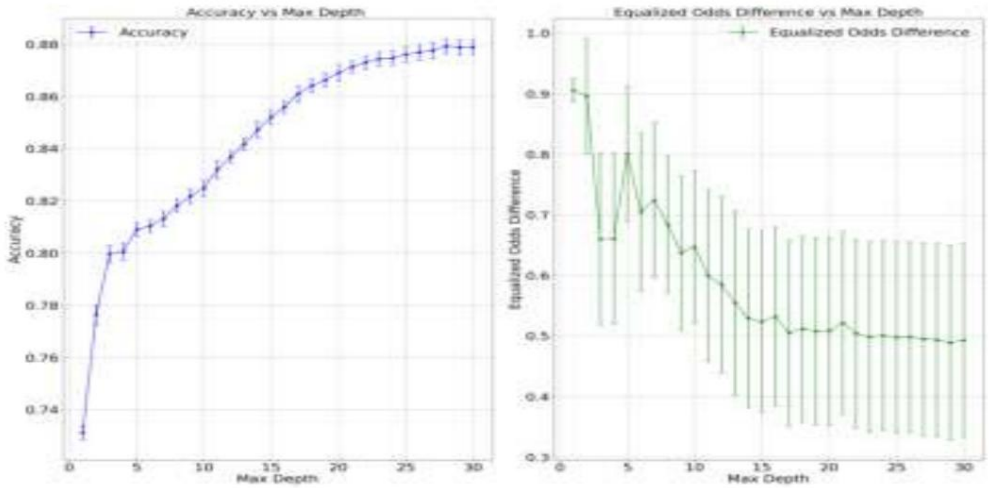


Abbildung 5. Diagramme des Einkommens Erwachsener unter Verwendung proportionaler Stichprobennahme, die Genauigkeit und EOD mit 95 % -Konfidenzintervallen gegenüber der maximalen Baumtiefe im Bereich von 1 bis 30 zeigen.

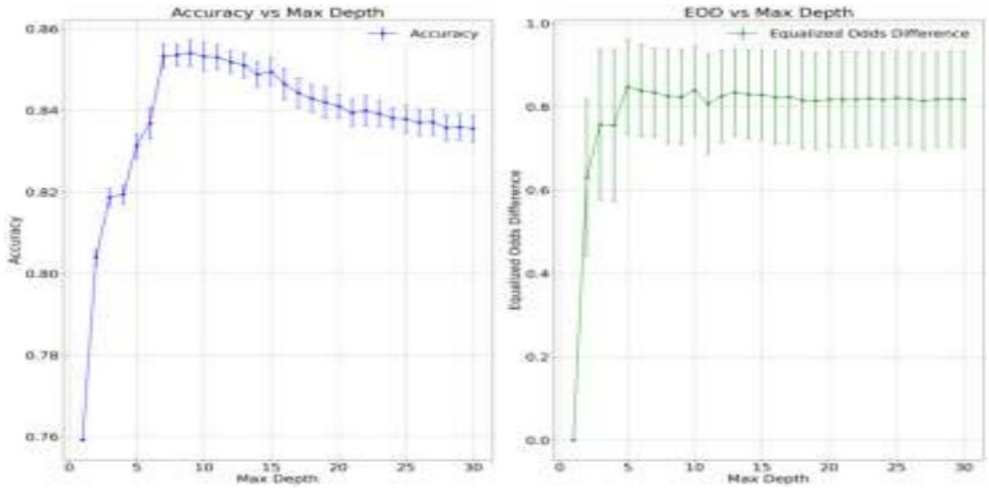


Abbildung 6. Diagramme des Erwachseneneinkommens unter Verwendung von PC-SMOTE, die Genauigkeit und EOD mit 95 %-Konfidenzintervallen gegenüber der maximalen Baumtiefe im Bereich von 1 bis 30 zeigen.

Abbildung 6 und 7, die die Ergebnisse mit PC-SMOTE bzw. PC-ADASYN darstellen, zeigen leichte Abwärtstrends bei der Genauigkeit, die sich kurzzeitig verbessern, bevor sie wieder abfallen – ein Muster, das auf Überanpassung bei größeren Tiefen hinweist. Die EOD-Metriken für diese Methoden zeigen anfängliche Stabilität bei geringeren Tiefen, Anstieg bei mittleren Tiefen und Rückgang, was auf komplexe Wechselwirkungen zwischen der Generierung synthetischer Proben und der Abgrenzung der Entscheidungsgrenzen hindeutet . **Angesichts dieser Beobachtungen wurde für unsere Experimente eine maximale Tiefe von 3 gewählt, da dies einen „Sweet Spot“ darstellt, bei dem sowo**

Angesichts dieser Ergebnisse ist eine Schlussfolgerung, den oft angenommenen Kompromiss zwischen Genauigkeit und Fairness in Frage zu stellen, indem gezeigt wird, dass unsere PC-ADASYN-Methode die Baselines in allen drei Datensätzen sowohl in Bezug auf Genauigkeit als auch Fairness durchgängig übertrifft. Dieses Ergebnis ist bedeutsam, da es darauf hindeutet, dass es möglich ist, die Modellfairness zu verbessern, ohne die Genauigkeit mit geeigneten Stichprobenmethoden und Modelloptimierung zu beeinträchtigen. Unsere Analyse zeigt jedoch auch Szenarien auf, in denen Anpassungen der Modellkomplexität, insbesondere der maximalen Tiefe von Entscheidungsbäumen, die Genauigkeit auf Kosten der Fairness verbessern können, wie durch Erhöhungen der Equalized Odds Difference (EOD) angezeigt wird. Es wird jedoch erwartet, dass die Kopplung von Stichprobenmethoden mit Inprocessing-Methoden wie fairnessbasierter Regularisierung diese Effekte ausgleichen kann. Diese Entscheidungen unterstreichen den Ermessensspielraum der Forscher bei der Ausbala Die Menge der Stichprobe und die zeitliche Komplexität sind wie bei jeder anderen Stichprobenmethode.

Machine Translated by Google

Mit zunehmender Stichprobenmenge erhöht sich auch die Zeitkomplexität, doch insgesamt weisen die Stichprobenmethoden aufgrund der gleichen Funktionalität die gleiche Zeitkomplexität auf wie die ihnen zugrunde liegenden Algorithmen .

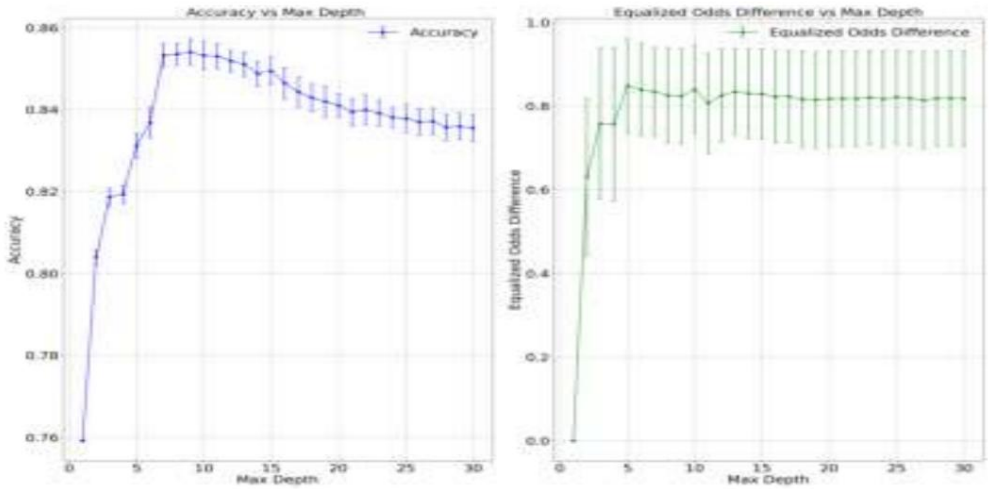


Abbildung 7. Diagramme des Erwachseneneinkommens unter Verwendung von PC-ADASYN, die Genauigkeit und EOD mit 95 %-Konfidenzintervallen gegenüber der maximalen Baumtiefe im Bereich von 1 bis 30 zeigen.

Darüber hinaus unterstreichen unsere Ergebnisse die Komplexität der gleichzeitigen Optimierung mehrerer Fairnessmetriken. So führten beispielsweise Bemühungen zur Verbesserung der statistischen Parität (SP) durch Bevorzugung positiverer Vorhersagen für jede geschützte Gruppe im COMPAS-Datensatz zu einer unbeabsichtigten Verringerung negativer Vorhersagen. Diese Verschiebung wirkte sich negativ auf die False Positive Rate (FPR), eine Komponente von EOD, aus und verschlechterte somit die EOD-Metrik, wenn sich die SP verbesserte. Dieses Phänomen veranschaulicht die inhärenten mathematischen Spannungen zwischen Fairnessmetriken, bei denen die Optimierung einer Metrik sich nachteilig auf eine andere auswirken kann. Der COMPAS-Datensatz mit seiner nahezu ausgeglichenen Klassenverteilung liefert ein konkretes Beispiel dafür, wie Datensatzmerkmale das Verhalten von Fairnessmetriken beeinflussen können. Die Optimierung der SP in diesem Zusammenhang impliziert eine verzerrte Messung der Fairness, insbesondere wenn inhärente Unterschiede zwischen Gruppen in geschützten Attributen bestehen. Dies wird durch die Literatur unterstützt, die darauf hinweist, dass die SP Gruppenunterschiede möglicherweise nicht ausreichend berücksichtigt, was möglicherweise zu irreführenden Schlussfolgerungen über di

7. Einschränkungen und zukünftige Arbeit

Aufgrund der Natur dieser Studie ist es nicht möglich, alle Fragen rund um Fairness und den sogenannten Kompromiss zwischen Fairness und Leistung zu behandeln. Daher sind der hier berichteten Arbeit Grenzen gesetzt. Dennoch hoffen und beabsichtigen wir, mit der hier berichteten Arbeit zusätzliche Forschungsansätze in diesem wichtigen Bereich aufzuzeigen.

Eine Einschränkung dieser Studie besteht darin, dass unsere Stichprobenmethode nicht speziell für die Optimierung beliebiger Fairnessmetriken entwickelt wurde. Anders ausgedrückt: Da zwischen den verfügbaren Fairnessmetriken häufig inhärente Kompromisse bestehen, wurde entschieden, sich auf einen metrikenunabhängigen Ansatz zu konzentrieren, da die Ergebnisse bei anderen Metriken anders ausgefallen sein könnten . Dies ist auch einer der Gründe, warum wir zwischen EOD und SP unterschiedliche Verhaltensweisen festgestellt haben.

Darüber hinaus wird anerkannt, dass die zugrunde liegende ML-Methode zwar für die vorgeschlagene Methode nicht relevant sein sollte, dies jedoch nicht tatsächlich getestet wurde. Daher wird diese Forschung in Zukunft erweitert, indem die Auswirkungen anderer ML-Algorithmen wie logistische Regression, Fuzzy ID3, K-Nearest Neighbor und Ensemblemethoden wie Random Forests oder Gradient-Boosted Trees berücksichtigt werden, um die Generalisierbarkeit unse Der Zweck einer solchen Studie wäre es, zu überprüfen, ob unsere Methoden unabhängig vom verwendeten ML-Algorithmus sind. Darüber hinaus würde dies dazu beitragen zu validieren, ob die beobachteten Verbesserungen bei Fairness und Genauigkeit modellspezifisch sind oder universell angewendet werden können.

Machine Translated by Google

Darüber hinaus wird anerkannt, dass nur drei unterschiedliche Datensätze berücksichtigt wurden – Datensätze , die in diesem Bereich umfassend untersucht wurden. Dies gibt Anlass zur Sorge, dass Methoden auf diese Daten zugeschnitten werden, anstatt sich mit dem umfassenderen Thema der Fairness im ML zu befassen. Um dieses Problem anzugehen, sind Experimente mit größeren und vielfältigeren Datensätzen geplant, um tiefere Einblicke in die Skalierbarkeit und Robustheit unserer Techniken zu erhalten. Ein weiterer Bereich für zukünftige Arbeiten ist die Verfeinerung unseres Multikategorien-Sampling-Ansatzes durch die Einbeziehung feinkörnigerer Unterteilungen geschützter Kategorien, wodurch möglicherweise subtilere Verzerrungen aufgedeckt und ein differenzierteres Verständnis von

Abschließend sei angemerkt, dass alternative Methoden zur Minderung von Verzerrungen vorgeschlagen wurden , die in dieser Arbeit jedoch überhaupt nicht untersucht wurden. Zukünftige Arbeiten würden Vergleiche mit weiteren Stichprobenstrategien erfordern. Es wird ein direkterer Vergleich der vorgeschlagenen Methoden mit In-Processing- und Post-Processing-Methoden durchgeführt. Beispielsweise wird die Einbeziehung von In-Processing-Methoden wie Regularisierung [22] oder einer Post-Processing- Methode wie dem Randomized Threshold Optimizer [64] als mögliche Mittel zur Erzielung weiterer Verbesserungen sowohl der Fairness als auch der Leistung untersucht .

8. Schlussfolgerungen

In dieser Studie wurde das Problem der Verzerrung bei ML-Vorhersagen untersucht und eine Methode entwickelt, die auf der Kombination geschützter Variablen in einer neuen Multikategorie basiert. Der Schwerpunkt lag insbesondere auf der in der Literatur aufgeworfenen Frage nach einem Kompromiss zwischen Verzerrung und Leistung und der Suche nach einer Methode zur Abmilderung dieses Kompromisses. Der vorgeschlagene neue Multikategorienansatz spiegelt die vielschichtige Identität von Personen wider und berücksichtigt das komplexe Zusammenspiel von Attributen, die reale Szenarien definieren. Angesichts des inhärenten Ungleichgewichts in dieser Multikategorie wurden vier Stichprobenmethoden entwickelt, die auf diese komplexen Kategorisierungen zugeschnitten sind, anstatt auf traditionelle Klassenbezeichnungen. Zum Zweck der Anwendung eines Basisklassifikators wurden Entscheidungsbäume trainiert und die Wirksamkeit dieser Methoden anhand von drei Datensätzen bewertet, die häufig in Fairnessstudien verwendet werden. Die Leistung der Methoden wurde mit Basismethoden ohne Stichproben verglichen, wobei Genauigkeit, Makro- F1, Equalized Odds Difference (EOD) und

Die Ergebnisse der Experimente zeigen, dass zwei der neu entwickelten Sampling- Techniken – PC-SMOTE und PC-ADASYN – die Fairness erfolgreich verbessern, ohne die Genauigkeit zu beeinträchtigen. Bemerkenswerterweise verbesserten diese Methoden in einigen Fällen auch die Genauigkeit und lieferten damit Beweise, die den gängigen Behauptungen eines Kompromisses zwischen Fairnes Eine weitere Analyse der Auswirkungen der maximalen Baumtiefe auf die Modellleistung ergab , dass eine zunehmende Tiefe zwar zunächst die Genauigkeit steigert, letztendlich jedoch zu einer Verschlechterung führt. Umgekehrt wirkt sich eine zunehmende Tiefe negativ auf die Fairness aus, was die Herausforderung unterstreicht, Komplexität und Gerechtigkeit in Einklang zu bringen. Es wurden jedoch optimale Baumtiefen identifiziert, die gleichzeitig Genauigkeit und EOD verbessern, was die Möglichkeit unterstreicht, Gerechtigkeit zu erreichen, ohne die Leistung zu beeinträchtigen.

Letztlich stellen diese Erkenntnisse die vorherrschenden Vorstellungen eines impliziten Kompromisses zwischen **Leistung und Fairness in der Forschung zur Vermeidung von Vorurteilen in Frage und legen nahe, dass sorgfältig konzipierte Strategien zur Vermeidung von Vorurteilen diesen Kompromiss umgehen können. Unser Ansatz schafft einen neuen Präzedenzfall für die Entwicklung gerechterer Vorhersagealgorithmen, indem er neu definiert, wie geschützte Attribute beim Modelltraining genutzt werden.**

Beiträge der Autoren: Konzeptualisierung, GP und JS; formale Analyse, GP und JS; Untersuchung, GP; Methodik, GP; Software, GP; Überwachung, JS; Validierung, GP und JS; Visualisierung, GP; Schreiben – Originalentwurf, GP; Schreiben – Überprüfung und Bearbeitung, JS. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Forschung erhielt keine externe Finanzierung.

Datenverfügbarkeitserklärung: Die in dieser Forschungsarbeit implementierten Datensätze und Codes wurden auf https://github.com/horlahsunbo/New-folder hochgeladen. **(abgerufen am 10. Juni 2024).**

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Machine Translated by Google

Verweise

1. **Maina, IW; Belton, TD; Ginzberg, S.; Singh, A.; Johnson, TJ Ein Jahrzehnt der Erforschung impliziter rassistischer/ethnischer Vorurteile im Gesundheitswesen Anbieter unter Verwendung des impliziten Assoziationstests.** Soc. Sci. Med. **2018,** 199, **219–229.** [CrossRef]
2. **Salimi, B.; Rodriguez, L.; Howe, B.; Suciu, D. Kausale Datenbankreparatur für algorithmische Fairness. In Proceedings der 2019 International Conference on Management of Data, Amsterdam, Niederlande, 30. Juni–5. Juli 2019; S. 793–810.**
3. **Kordzadeh, N.; Ghasemaghaei, M. Algorithmische Verzerrung: Überprüfung, Synthese und zukünftige Forschungsrichtungen.** Eur. J. Inf. Syst. **2022,** 31, **388–409.** [CrossRef]
4. **Pessach, D.; Shmueli, E. Eine Übersicht über Fairness im maschinellen Lernen.** ACM Comput. Surv. **2022,** 55, **1–44.** [CrossRef]
5. **Aghaei, S.; Azizi, MJ; Vayanos, P. Lernen optimaler und fairer Entscheidungsbäume für diskriminierungsfreie Entscheidungsfindung.** Proc. AAA Conf. Artif. Intell. **2019,** 33, **1418–1426.** [CrossRef]
6. **Kamiran, F.; Calders, T. Datenvorverarbeitungstechniken für eine Klassifizierung ohne Diskriminierung.** Knowl. Inf. Syst. **2012,** 33, **1–33.** [Querverweis]
7. **Calmon, F.; Wei, D.; Vinzamuri, B.; Ramamurthy, KN; Varshney, KR Optimierte Vorverarbeitung zur Diskriminierungsprävention. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4.–9. Dezember 2017.**
8. **Shahbazi, N.; Lin, Y.; Asudeh, A.; Jagadish, HV Darstellungsverzerrung in Daten: Eine Untersuchung zu Identifizierungs- und Lösungstechniken.** ACM Comput. Surv. **2023,** 55, **1–39.** [CrossRef]
9. **Chen, Z.; Zhang, JM; Sarro, F.; Harman, M. Eine umfassende empirische Studie zu Methoden zur Bias-Minderung beim maschinellen Lernen Klassifikatoren.** ACM Trans. Softw. Eng. Methodol. **2023,** 32, **1–30.** [CrossRef]
10. **Perzynski, A.; Berg, KA; Thomas, C.; Cemballi, A.; Smith, T.; Shick, S.; Gunzler, D.; Sehgal, AR Rassistische Diskriminierung und wirtschaftliche Faktoren bei der Redlining-Politik in Ohio.** Bois Rev. Soc. Sci. Res. Race **2023,** 20, **293–309.** [CrossRef]
11. **Steil, JP; Albright, L.; Rugh, JS; Massey, DS Die soziale Struktur der Hypothekendiskriminierung.** Hous. Stud. **2018,** 33, **759–776.** [Querverweis]
12. **Salgado, JF; Moscoso, S.; García-Izquierdo, AL; Anderson, NR** Gestaltung integrativer Arbeitsplätze durch sozialen Dialog; **Springer: Berlin/Heidelberg, Deutschland, 2017.**
13. **Leavy, S.; Meaney, G.; Wade, K.; Greene, D. Abschwächung geschlechtsspezifischer Voreingenommenheit in Datensätzen des maschinellen Lernens. In Proceedings of the Bias and Social Aspects in Search and Recommendation: Erster internationaler Workshop, BIAS 2020, Lissabon, Portugal, 14. April 2020; S. 12–26.**
14. **Hort, M.; Chen, Z.; Zhang, JM; Harman, M.; Sarro, F. Bias-Minderung für maschinelle Lernklassifikatoren: Eine umfassende Untersuchung.** ACM J. Responsible Comput. **2023, 1, 1–52.** [CrossRef]
15. **Fahse, T.; Huber, V.; van Giffen, B. Umgang mit Voreingenommenheit in Projekten zum maschinellen Lernen. In** Innovation Through Information Systems: Band II: Eine Sammlung neuester Forschung zu Technologiefragen; **Springer: Berlin/Heidelberg, Deutschland, 2021; S. 94–109.**
16. **Zhang, Z.; Neill, DB Identifizierung signifikanter prädiktiver Verzerrungen in Klassifikatoren.** arXiv **2017, arXiv:1611.08292.** [CrossRef]
17. **Zheng, Z.; Cai, Y.; Li, Y. Oversampling-Methode für unausgewogene Klassifizierung.** Comput. Inform. **2015,** 34, **1017–1037.**
18. **Chawla, NV; Bowyer, KW; Hall, LO; Kegelmeyer, WP SMOTE: Synthetische Minderheiten-Oversampling-Technik.** J. Artif. Intell. Res. **2002,** 16, **321–357.** [CrossRef]
19. **He, H.; Bai, Y.; Garcia, EA; Li, S. ADASYN: Adaptiver synthetischer Sampling-Ansatz für unausgewogenes Lernen. In Proceedings der 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hongkong, China, 1.–8. Juni 2008; S. 1322–1328.**
20. **Janssen, P.; Sadowski, BM Bias in Algorithms: Über den Kompromiss zwischen Genauigkeit und Fairness. In Proceedings der 23. Zweijahreskonferenz der International Telecommunications Society, Göteborg, Schweden, 21.–23. Juni 2021.**
21. **Hardt, M.; Price, E.; Srebro, N. Chancengleichheit beim überwachten Lernen. In Proceedings of the Advances in Neural Informationsverarbeitungssysteme, Barcelona, Spanien, 5.–10. Dezember 2016.**
22. **Kamishima, T.; Akaho, S.; Asoh, H.; Sakuma, J. Fairnessbewusster Klassifikator mit Vorurteilsbeseitiger-Regularisierer. In Proceedings of the Machine Learning and Knowledge Discovery in Databases: Europäische Konferenz, ECML PKDD 2012, Bristol, Großbritannien, 24.–28. September 2012; S. 35–50.**
23. **Calders, T.; Žliobaitė, I. Warum unvoreingenommene Rechenprozesse zu diskriminierenden Entscheidungsverfahren führen können. In** Diskriminierung und Datenschutz in der Informationsgesellschaft: Data Mining und Profiling in großen Datenbanken; **Springer: Berlin/Heidelberg, Deutschland, 2013; S. 23–33.**
24. **Burt, A. Wie man Diskriminierung in der KI bekämpft. Harvard Business Review. 2020 Online verfügbar: https://hbr.org/2020/08/how-to- fight-discrimination-in-ai (abgerufen am 12. Juli 2024).**
25. **Siegler, A.; Admussen, W. Entdeckung rassistischer Diskriminierung durch die Polizei.** Northwestern Univ. Law Rev. **2021,** 115, **987–1054.**
26. **Grabowicz, P.; Perello, N.; Mishra, A. Wie trainiert man Modelle, die keine Diskriminierung propagieren? Equate and Machine Learning Blog, University of Massachusets, Amherst. 2022. Online verfügbar: https://groups.cs.umass.edu/equate-ml/2022/04/07/how- to-train-models-that-do-not-propagate-discrimination/ (abgerufen am 12. Juli 2024).**
27. **Khani, F.; Liang, P. Von Diskriminierung im maschinellen Lernen zu Diskriminierung im Recht, Teil 1: Ungleiche Behandlung. Stanford AI Lab Blog. 2022. Online verfügbar: https://ai.stanford.edu/blog/discrimination_in_ML_and_law/ (abgerufen am 12. Juli 2024).**
28. **Shahriari, K.; Shahriari, M. IEEE-Standardüberprüfung – Ethisch ausgerichtetes Design: Eine Vision zur Priorisierung des menschlichen Wohlbefindens mit künstlicher Intelligenz und autonomen Systemen. In Proceedings der IEEE Canada International Humanitarian Technology Conference (IHTC), Toronto, ON, Kanada, 21.–22. Juli 2017; S. 197–201.**

Machine Translated by Google

29. Europäische Kommission; Generaldirektion Kommunikationsnetze, Inhalte und Technologien. Ethische Leitlinien für vertrauenswürdige KI. Amt für Veröffentlichungen. 2019. Online verfügbar: https://op.europa.eu/en/publication-detail/-/publication/d39885 69-0434-11ea-8c1f-01aa75ed71a1/Sprache-de/Format-PDF/Quelle-337437547 (abgerufen am 10. Juni 2024)

30. Ebers, M.; Hoch, VRS; Rosenkranz, F.; Ruschemeier, H.; Steinrötter, B. Der Vorschlag der Europäischen Kommission für ein Gesetz zur künstlichen Intelligenz – Eine kritische Bewertung durch Mitglieder der Robotics and AI Law Society (RAILS). J 2021, 4, 589–603. [CrossRef]

31. Hajian, S.; Domingo-Ferrer, J. Eine Methodologie zur direkten und indirekten Diskriminierungsprävention im Data Mining. IEEE Trans. Knowl. Data Eng. **2012,** 25, **1445–1459. [CrossRef]**

32. Fish, B.; Kun, J.; Lelkes, Á.D. Ein auf Vertrauen basierender Ansatz zum Ausgleich von Fairness und Genauigkeit. In Proceedings of the 2016 SIAM Internationale Konferenz zum Data Mining, Miami, FL, USA, 5.–7. Mai 2016; S. 144–152.

33. Hajian, S. Gleichzeitige Diskriminierungsprävention und Datenschutz bei der Veröffentlichung und dem Mining von Daten. arXiv **2013, arXiv:1306.6805. [CrossRef]**

34. Sondeck, LP; Laurent, M.; Frey, V. Die semantische Diskriminationsratenmetrik für Datenschutzmessungen, die den Nutzen von t-Nähe gegenüber l-Diversität in Frage stellt. In Proceedings der 14. Internationalen Konferenz für Sicherheit und Kryptographie, Madrid, Spanien, 24.–26. Juli 2017; Band 6, S. 285–294.

35. Ruggieri, S. Verwendung von T-Closeness-Anonymität zur Kontrolle von Nichtdiskriminierung. Trans. Data Priv. **2014, 2, 99–129.**

36. Romano, Y.; Bates, S.; Candes, E. Erreichen ausgeglichener Chancen durch erneutes Sampling sensibler Attribute. In Proceedings of the Advances in Neural Information Processing Systems, Online, 6.–12. Dezember 2020; S. 361–371.

37. Peng, K.; Chakraborty, J.; Menzies, T. Fairmask: Bessere Fairness durch modellbasierte Neugewichtung geschützter Attribute. IEEE Trans. Softw. Ing. **2022,** 49, **2426–2439. [CrossRef]**

38. Dhar, P.; Gleason, J.; Roy, A.; Castillo, CD; Chellappa, R. PASS: Geschütztes Attributunterdrückungssystem zur Minderung von Verzerrungen bei der Gesichtserkennung. In Proceedings der IEEE/CVF International Conference on Computer Vision, Montreal, QC, Kanada, 10.–17. Oktober 2021; S. 15087–15096.

39. Krasanakis, E.; Spyromitros-Xioufis, E.; Papadopoulos, S.; Kompatsiaris, Y. Adaptive sensible Neugewichtung zur Minderung von Verzerrungen bei fairnessbewusster Klassifizierung. In Proceedings der World Wide Web Conference, Lyon, Frankreich, 23.–27. April 2018; S. 853–8

40. Ranÿci´c, S.; Radovanovi´c, S.; Delibaši´c, B. Untersuchung von Oversampling-Techniken für faire Modelle des maschinellen Lernens. In Proceedings of the Decision Support Systems XI: Decision Support Systems, Analytics and Technologies in Response to Global Crisis Management: 7. Internationale Konferenz zu Decision Support System Technology, ICDSST 2021, Loughborough, Großbritannien, 26.–28. Mai 2021; Springer: Berlin/Heidelberg, Deutschland, 2021; S. 110–123.

41. Yan, S.; te Kao, H.; Ferrara, E. Fairer Klassenausgleich: Verbesserung der Modellfairness ohne Berücksichtigung sensibler Attribute. In Proceedings der 29. ACM International Conference on Information & Knowledge Management, virtuelle Veranstaltung, Irland, 19.–23. Oktober 2010; S. 1715–1724.

42. HuZhang, B.; Lemoine, B.; Mitchell, M. Unerwünschte Vorurteile durch kontroverses Lernen abmildern. In Proceedings der AAAI/ACM-Konferenz zu KI, Ethik und Gesellschaft, New Orleans, LA, USA, 2.–3. Februar 2018; S. 335–340.

43. Celis, LE; Huang, L.; Keswani, V.; Vishno, NK Klassifizierung mit Fairnessbeschränkungen: Ein Metaalgorithmus mit nachweisbaren Garantien. In Proceedings of the Conference on Fairness, Accountability, and Transparency, Atlanta, GA, USA, 29.–31. Januar 2019; S. 319–328.

44. Zafar, MB; Valera, I.; Rodriguez, MG; Gummadi, KP Fairness jenseits von ungleicher Behandlung und ungleicher Auswirkung: Klassifizierung lernen ohne ungleiche Misshandlung. In Proceedings der 26. International Conference on World Wide Web, Perth, Australien, 3.–7. April 2017; S. 1171–1180.

45. Agarwal, A.; Beygelzimer, A.; Dudík, M.; Langford, J.; Wallach, H. Ein Reduktionsansatz zur fairen Klassifizierung. In Proceedings of die Internationale Konferenz für maschinelles Lernen, Stockholm, Schweden, 10.–15. Juli 2018; S. 60–69.

46. Lowy, A.; Baharlouei, S.; Pavan, R.; Razaviyayn, M.; Beirami, A. Ein stochastisches Optimierungsframework zur fairen Risikominimierung. Trans. Mach. Learn. Res. **2022. [CrossRef]**

47. Spinelli, I.; Scardapane, S.; Hussain, A.; Uncini, A. Fairdrop: Biased Edge Dropout zur Verbesserung der Fairness beim Lernen von Graphdarstellungen . IEEE Trans. Artif. Intell. **2021, 3, 344–354. [CrossRef]**

48. Hort, M.; Zhang, JM; Sarro, F.; Harman, M. Fairea: Ein Modellverhaltensmutationsansatz zum Benchmarking von Methoden zur Minderung von Verzerrungen. In Proceedings des 29. gemeinsamen Treffens der ACM zur European Software Engineering Conference und zum Symposium über die Grundlagen des Software Engineering, Athen, Griechenland, 23.–28. August 2012; S. 994–1006.

49. Bhaskaruni, D.; Hu, H.; Lan, C. Verbesserung der Vorhersagefairness durch Modellensemble. In Proceedings der 31. IEEE-Internationalkonferenz zu Tools mit künstlicher Intelligenz (ICTAI), Portland, OR, USA, 4.–6. November 2019; S. 1810–1814.

50. Iosifidis, V.; Fetahu, B.; Ntoutsi, E. Fae: Ein fairnessbewusstes Ensemble-Framework. In Proceedings der IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9.–12. Dezember 2019; S. 1375–1380.

51. Dwork, C.; Hardt, M.; Pitassi, T.; Reingold, O.; Zemel, R. Fairness durch Bewusstsein. In Proceedings der 3. Innovations in Konferenz für Theoretische Informatik, Cambridge, MA, USA, 8.–10. Januar 2012; S. 214–226.

52. Pleiss, G.; Raghavan, M.; Wu, F.; Kleinberg, J.; Weinberger, KQ Über Fairness und Kalibrierung. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4.–8. Dezember 2017.

53. Karimi-Haghighi, M.; Castillo, C. Verbesserung eines Tools zur Rückfallvorhersage mit maschinellem Lernen: Wirksamkeit und algorithmische Fairness. In Proceedings der 18. Internationalen Konferenz zu Künstlicher Intelligenz und Recht, São Paulo, Brasilien, 21.–25. Juni 2017; S. 210–214.

Machine Translated by Google

54. Friedler, SA; Scheidegger, C.; Venkatasubramanian, S. Die (Un-)Möglichkeit von Fairness: Unterschiedliche Wertesysteme erfordern unterschiedliche Mechanismen für faire Entscheidungsfindung. Commun. ACM **2021,** 64, **136–143. [CrossRef]**

55. García, V.; Sánchez, JS; Mollineda, RA Über die Wirksamkeit von Vorverarbeitungsmethoden beim Umgang mit unterschiedlichen Klassenungleichgewichten . Knowl.-Based Syst. **2012,** 25, **13–21. [CrossRef]**

56. Breiman, L. Klassifikations- und Regressionsbäume; **Routledge: London, Großbritannien, 2017.**

57. Pedregosa, F.; Varoquaux, G.; Grafort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Maschinelles Lernen in Python. J. Mach. Learn. Res. **2011,** 12, **2825–2830.**

58. Kelly, M.; Longjohn, R.; Nottingham, K. The UCI Machine Learning Repository. 2024. Online verfügbar: https://archive.ics.uci.edu (abgerufen am 10. Juni 2024).

59. Becker, B.; Kohavi, R. Adult. UCI Machine Learning Repository. 1996. Online verfügbar: https://archive.ics.uci.edu/dataset/2/Erwachsene (abgerufen am 10. Juni 2024). [CrossRef]

60. Hofmann, H.; Statlog (Deutsche Kreditdaten). UCI Machine Learning Repository. 1994. Online verfügbar: https://archive.ics.uci.edu/dataset/144/statlog+german+credit+data (abgerufen am 10. Juni 2024). [CrossRef]

61. Dressel, J.; Farid, H. Genauigkeit, Fairness und Grenzen der Rückfallvorhersage. Sci. Adv. **2018, 4, eaao5580. [CrossRef]**

62. Huang, Q.; Mao, J.; Liu, Y. Ein verbesserter Rastersuchalgorithmus zur Optimierung von SVR-Parametern. In Proceedings der 2012 IEEE 14th International Conference on Communication Technology, Chengdu, China, 9.–11. November 2012; S. 1022–1026.

63. Caton, S.; Haas, C. Fairness im maschinellen Lernen: Eine Untersuchung. ACM Comput. Surv. **2023,** 56, **1–38. [CrossRef]**

64. Alabdulmohsin, I.; Lucic, M. Ein nahezu optimaler Algorithmus zur Entzerrung trainierter Machine-Learning-Modelle. In Proceedings of the 35. Konferenz zu neuronalen Informationsverarbeitungssystemen, online, 6.–14. Dezember 2021.

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultieren.

Machine Translated by Google

Elektronik

MDPI

Artikel

Sichere Verarbeitung und Verteilung von Daten, die auf privaten InterPlanetary File System mit Zero-Knowledge-Beweisen

Kyohei Shibano 1,*, Kensuke Ito [1], Changhee Han[2], Tsz Tat Chu[2], Wataru Ozaki [2] und Gento Mogi [1]

[1] Abteilung für Technologiemanagement für Innovation, Fakultät für Ingenieurwissenschaften, Universität Tokio,
[2] Tokio 113-8656, Japan Callisto
Inc., Tokio 171-0022, Japan * Korrespondenz: shibano@tmi.tu-tokyo.ac.jp

Zusammenfassung: In dieser Studie wird eine neue Methode zum Teilen von Daten vorgeschlagen, die ein privates InterPlanetary File System – ein dezentrales Speichersystem, das innerhalb eines geschlossenen Netzwerks betrieben wird – verwendet, um Daten an externe Stellen zu verteilen und gleichzeitig deren Authentizität überprüfbar zu machen. Unter den beiden Betriebsmodi von IPFS, öffentlich und privat, konzentriert sich diese Studie auf die Methode zur Verwendung von privatem IPFS. Privates IPFS ist nicht für die breite Öffentlichkeit zugänglich; obwohl es bei der Verteilung von Daten an externe Parteien ein Risiko der Datenmanipulation birgt , stellt die vorgeschlagene Methode die Authentizität der empfangenen Daten sicher. Insbesondere wendet diese Methode eine Art Zero-Knowledge-Beweis an, nämlich das Groth16-Protokoll von zk-SNARKs, um sicherzustellen, dass die Daten dem Inhaltsbezeichner in einem privaten IPFS entsprechen. Darüber hinaus wird der Name des Empfängers in die verteilten Daten eingebettet, um eine unbefugte Sekundärverteilung zu verhindern. Experimente bestätigte In zukünftigen Studien wird die vorgeschlagene Methode auf größere und vielfältigere Datentypen angewendet.

Schlüsselwörter: IPFS; Zero-Knowledge-Beweis; circom; zk-SNARKs; privates IPFS; Datenverteilung; Datenverarbeitung ; Datensicherheit

check for updates

Zitat: Shibano, K.; Ito, K.; Han, C.; Chu, TT; Ozaki, W.; Mogi, G. Sichere Verarbeitung und Verteilung von Daten, die auf einem privaten interplanetaren Dateisystem verwaltet werden, unter Verwendung von Zero-Knowledge-Beweise. Electronics **2024,** 13, **3025.**

https://doi.org/10.3390/Elektronik13153025

Wissenschaftliche Redakteurin: Aryya Gangopadhyay

Empfangen: 13. Juni 2024
Überarbeitet: 14. Juli 2024
Akzeptiert: 20. Juli 2024
Veröffentlicht: 31. Juli 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Einleitung

Dezentrale Systeme sind robust, da sie keinen einzelnen Ausfallpunkt aufweisen. Daher werden sie in vielen Unternehmensbereichen eingesetzt, darunter Kryptowährungen, Lieferkettenmanagement , Finanzdienstleistungen und digitale Identität. Um große Datenmengen wie Bilder zu speichern , benötigen diese Systeme Speicherfunktionen, die von Natur aus dezentralisiert sind. Blockchain, das häufig in Verbindung damit verwendet wird, verarbeitet normalerweise kleinere Datenmengen wie Transaktionshistorien und fungiert als Ledger-Datenbank. Das InterPlanetary File System (IPFS) ist ein bekanntes dezentrales Speichersystem, das Daten über mehrere Knoten hinweg speichert, um die Datenverfügbarkeit zu verbessern. Das IPFS gibt es in zwei Varianten: öffentliches IPFS, bei dem die Daten von jedem Benutzer mit uneingeschränktem Zugriff gespeichert werden können , und privates IPFS, bei dem ein geschlossenes Netzwerk eingerichtet wird, auf das nur bestimmte Organisationen oder Gruppen zugreifen können , was mehr Privatsphäre **und Sicherheit bietet .**

Beim Speichern von Daten in IPFS ist es wichtig, die Unterschiede zwischen öffentlichem und privatem IPFS zu verstehen. Öffentliches IPFS ermöglicht jedem den Zugriff auf Daten, während privates IPFS nur innerhalb bestimmter Organisationen oder Gruppen zugänglich ist, was Datenschutz und Sicherheit verbes Beim Speichern sensibler Informationen, wie vertraulicher Daten, in öffentlichen IPFS ist die Anwendung eines geeigneten Verschlüsselungsschemas unerlässlich, um den Datenschutz zu gewährleisten. Im Gegensatz dazu bietet privates IPFS eine höhere Sicherheit für die Datenspeicherung, da es nur innerhalb eines geschlossenen Netzwerks zugänglich ist.

Insbesondere für Organisationen wie Unternehmen oder Gesundheitseinrichtungen birgt die Speicherung von Daten in öffentlichen IPFS trotz der Verwendung starker Verschlüsselungstechnologien inhärente Darüber hinaus besteht in solchen Fällen ständig die Gefahr von Datenlecks aufgrund von Betriebsfehlern . Die Daten sind zwar verschlüsselt, aber der Welt zugänglich und damit anfällig für Brute-Force-Angriffe und andere Sicherheitsbedrohungen.

Machine Translated by Google

In Bezug auf die Zugänglichkeit ermöglicht öffentliches IPFS allgemeinen Benutzern den direkten Zugriff und Abruf Daten. In privaten IPFS müssen Daten jedoch von Mitgliedern der Organisation oder Gruppe, die das Netzwerk bildet. Wenn während dieses Prozesses Daten manipuliert werden, können Benutzer nicht in der Lage, es zu erkennen. Daher ist es wichtig, den für den Umgang mit Daten verantwortlichen Vermittlern zu vertrauen Übertragung in solchen Fällen wird obligatorisch. Um das oben erwähnte Vertrauensproblem anzugehen, Hier wird eine Methode zur Verteilung von in einem privaten IPFS gespeicherten Daten an externe Einheiten vorgeschlagen. und gleichzeitig seine Authentizität überprüfbar zu machen. Das Groth16 [1] Protokoll von zk-SNARKs, einer Art von ZKP wird auf in einem privaten IPFS gespeicherte Daten angewendet, um die Authentizität der Daten sicherzustellen. Darüber hinaus werden die Informationen des Empfängers in die verteilten Daten eingebettet, um zu verhindern, unerlaubte Sekundärverteilung. Die vorgeschlagene Methode der Datenfreigabe ist wichtig weil es auf den Fall privater IPFS zugeschnitten ist.

Die Unterschiede in mehreren Aspekten, einschließlich Sicherheit und Zugänglichkeit, bei der Speicherung Daten in öffentlichen IPFS und privaten IPFS innerhalb der Unternehmensdomäne werden zusammengefasst in Tabelle 1. Diese Studie schlägt Lösungen für die mit privaten IPFS verbundenen Bedrohungen vor.

Tabelle 1. Vergleich zwischen öffentlichen und privaten IPFS im Unternehmensbereich.

	Öffentliches IPFS	Privates IPFS
Vertrauensmodell	**Vertrauenslos**	**Erfordert Vertrauen in die operative Gruppe**
Zugangsbeschränkungen	**Für jeden zugänglich**	**Nur innerhalb der Betriebsgruppe zugänglich**
Risiko von Datenlecks	**Ständiges Risiko von Datenlecks aufgrund von Benutzerfehlern.**	**Geringes Risiko von Datenlecks innerhalb eines geschlossenen Netzwerks.**
Umgang mit vertraulichen Information	**Erfordert eine ordnungsgemäße Verschlüsselung**	**Daten, die in IPFS gespeichert sind, erfordern selbst keine hochgradige Verschlüsselung; es gibt einen Vertrauenspunkt bei der Weitergabe von Daten an Benutzer**
Das Risiko eines Brute-Force-Angriffs ist immer vorhanden		**Niedrig**
Datenabrufmethode	**Direkter Zugriff durch Benutzer**	**Von Mitgliedern der Organisation oder Gruppe** erhaltene Daten
Bedrohungen	**Erfordert eine Verschlüsselung, die eine Entschlüsselung durch unbefugte Benutzer** verhindert	Bei der Übermittlung **von Daten an Benutzer besteht das Risiko einer Manipulation.**

Der Rest dieses Dokuments ist wie folgt gegliedert. Abschnitt 2 präsentiert verwandte frühere Forschung. Abschnitt 3 beschreibt die grundlegenden Technologien, d. h. ZKP und zk-SNARKs. Abschnitt 4 beschreibt die Struktur der vorgeschlagenen Methode, während Abschnitt 5 die mögliche Anwendungen dieser Methode. Abschnitt 6 stellt die Implementierung dieser Methode vor, Abschnitt 7 befasst sich mit den Experimenten, die zur Überprüfung der Wirksamkeit der Implementierung. Abschnitt 8 enthält eine Diskussion der experimentellen Ergebnisse, während Abschnitt 9 stellt die Schlussfolgerungen des Papiers und einen Überblick über künftige Herausforderungen dar.

2. Verwandte Studien

Bestehende dezentrale Systeme nutzen IPFS, insbesondere in Kombination mit Blockchain Technologie. Kumar et al. [2] schlugen eine Methode zur sicheren Verwaltung medizinischer Daten vor, indem Integration von IPFS in eine Blockchain. Azbeg et al. [3] schlugen speziell ein System vor, das Verwaltung und Speicherung medizinischer Daten mithilfe von privatem IPFS und einer autorisierten Blockchain durch Verwendung einer Proxy-Neuverschlüsselung, um eine sichere Entschlüsselung durch benannte Ärzte zu gewährleisten. Wenn ein Wenn ein Arzt einige Patientendaten erhält, bekommt er die neu verschlüsselten Daten über ein Krankenhaus. Hossan et al. [4] schlugen auch ein System zur sicheren Aufzeichnung von Informationen für Mitfahrgelegenheiten vor Dienste mit IPFS und einer privaten Blockchain.

Mit dem Schwerpunkt auf der Kontrolle der Verteilung der von IPFS verwalteten Daten, Lin et al. [5] schlug ein System zum Schutz privater Daten vor, das ein verbessertes IPFS kombiniert mit einem Blockchain. Dieses System zeichnete Dateimetadaten und Zugriffsberechtigungen auf der Blockchain auf, Damit können Benutzer die Dateifreigabe steuern. Darüber hinaus implementierte das System effiziente Verwaltungsfunktionen mithilfe von Smart Contracts und verbesserte so die Datensicherheit und -verwaltung.

Machine Translated by Google

Flexibilität. Battah et al. [6] entwickelten ein System, das Mehrparteienauthentifizierung (MPA), Proxy-Neuverschlüsselung und Smart Contracts auf einer Blockchain für die dezentrale Zugriffskontrolle von verschlüsselten, in IPFS gespeicherten Daten nutzte. Huang et al. [7] führten einen vertrauenswürdigen IPFS-Proxy ein, um Zugriffskontrolle und Gruppenschlüsselverwaltung für verschlüsselte, in IPFS gespeicherte Daten zu realisieren. Sun et al. [8] schlugen ein System vor, das nur Personen mit entsprechenden Attributen erlaubte, in IPFS gespeicherte verschlüsselte Daten unter Verwendung eines auf Chiffretextrichtlinienattributen basierenden Verschlüsselungssystems zu entschlüsseln, wodurch ein effizientes medizinisches Informationsmanagement erleichtert wurde. Kang et al. [9] ermöglichten die Verteilung von mit privatem IPFS und einer privaten Blockchain verwalteten Daten an externe Benutzer unter Verwendung eines Named Data Network (NDN). Zudem schlugen Uddin et al. [10] ein File-Sharing-System vor , das IPFS und Public-Key-Infrastruktur (PKI)-Technologie nutzte, ohne dass eine vertrauenswürdige Dritt

Mehrere Studien haben ZKP zur Datenverteilung verwendet. Beispielsweise schlugen Li et al. [11] ein datenschutzfreundliches Verkehrsmanagementsystem vor, das nicht interaktive Zero-Knowledge-Bereichsnachweise mit einer Blockchain kombinierte. Ein Prototyp mit Hyperledger Fabric und Hyperledger Ursa erfüllte die Datenschutzanforderungen für Echtzeit-Verkehrsmanagement.

Diese Studie schlägt eine Methode zur angemessenen Verarbeitung und Verteilung von Daten vor, die innerhalb eines privaten IPFS verwaltet werden, und zwar an Benutzer außerhalb des Netzwerks. Damit bietet sie einen anderen Ansatz als die in früheren Studien vorgeschlagenen. Einige Studien haben die Proxy-Neuverschlüsselung als geeignete Methode zur Datenspeicherung und -verteilung in IPFS übernommen **[3,6]. Mit dieser Methode können verteilte Daten erneut verschlüsselt werden, um mit dem privaten Schlüssel des Empfängers entschlüsselt zu werden. Darüber hinaus ermöglicht es beim Speichern von Daten in IPFS den Empfängern, die Richtigkeit ihrer empfangenen Daten nach der Entschlüsselung zu überprüfen, indem der Hashwert der vorverschlüsselten Daten in der Blockchain** aufgezeichnet **wird. Diese Methode kann jedoch keine Fälle verarbeiten, in denen Daten verarbeitet werden, wie z. B. das Einbetten des Namens des Empfänge**

3. Zero-Knowledge-Proof ZKP

ist ein kryptografisches Protokoll, das es einem Beweiser ermöglicht, einem Prüfer die Gültigkeit einer Aussage zu beweisen, ohne weitere Informationen außer der Gültigkeit der Aussage preiszugeben. Die Annahme dieser Studie ist, dass die einer externen Entität bereitgestellten Daten auf der Grundlage einer bestimmten CID generiert werden. Unser Ziel ist es, einem Mitglied eines privaten IPFS (Beweiser) zu ermöglichen, diese Aussage einer externen Entität (Prüfer als Empfänger der Daten) zu beweisen, ohne weitere wichtige Informationen (wie IPFS-Zugriffsrechte und Verschlüsselungsschlüssel) preiszugeben.

ZKP, insbesondere das in dieser Studie verwendete Groth16-Protokoll von zk-SNARKs, beginnt mit einem vertrauenswürdigen Setup, bei dem beide Parteien öffentliche Parameter festlegen, die für die sichere Generierung und Überprüfung von Beweisen von entscheidender Bedeutung sind. Im ZKP-Schema generieren wir zunächst einen Schaltkreis, der den Prozess beschreibt, für den ein Beweis vorgesehen ist. Der Schaltkreis enthält die zu überprüfenden Bedingungen, wie z. B. das Vorhandensein einer CID. Durch das ZKP-Schema werden kryptografische Schlüssel – insbesondere ein Beweisschlüssel und ein Überprüfungsschlüssel – erstellt. Diese Schlüssel sind für die Erstellung eines Beweises für den Schaltkreis und dessen Überprüfung von entscheidender Bedeutung. Unter Verwendung des Beweisschlüssels und der Eingabedaten generiert der Beweiser einen Beweis, der die Gültigkeit der Ausgabedaten anhand der im Schaltkreis angegebenen Bedingungen aufzeigt. Der Prüfer verwendet dann den Überprüfungsschlüssel, um den Beweis und die Ausgabedaten zu überprüfen. Wenn der Beweis gültig ist, bestätigt dies die Integrität der Daten, ohne zugrunde liegend

In dieser Studie wird die Groth16-Verarbeitung mit circom [12,13] und snarkjs [14] durchgeführt. Der Prozessablauf ist in Abbildung 1 zusammengefasst.

zk-SNARKs wird aufgrund seiner nicht-interaktiven Natur und Effizienz eingesetzt, die besonders für Systeme mit Smart Contracts von Vorteil sind, da der Rechenaufwand zur Überprüfung des Nachweises gering ist. Darüber hinaus ist zk-SNARKs für seinen hohen Rechenaufwand und die Notwendigkeit fortgeschrittener PC-Spezifikationen bekannt. Beispielsweise dauert in einer der repräsentativen Anwendungen von zk-SNARKs, Zcash [15], der Prozess **der Nachweisgenerierung für eine einzelne anonyme Transaktion über eine halbe Minute [16].**

Machine Translated by Google

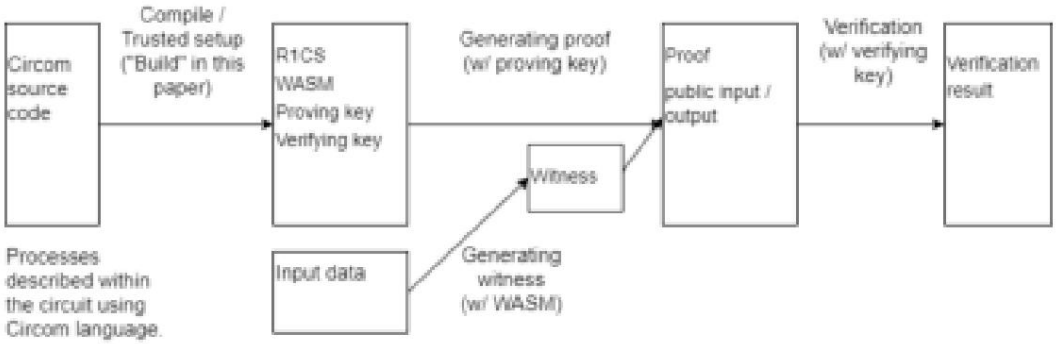


Abbildung 1. Zero-Knowledge-Beweis mit Circom.

4. Vorgeschlagenes

System Abbildung 2 bietet einen Überblick über das vorgeschlagene System.

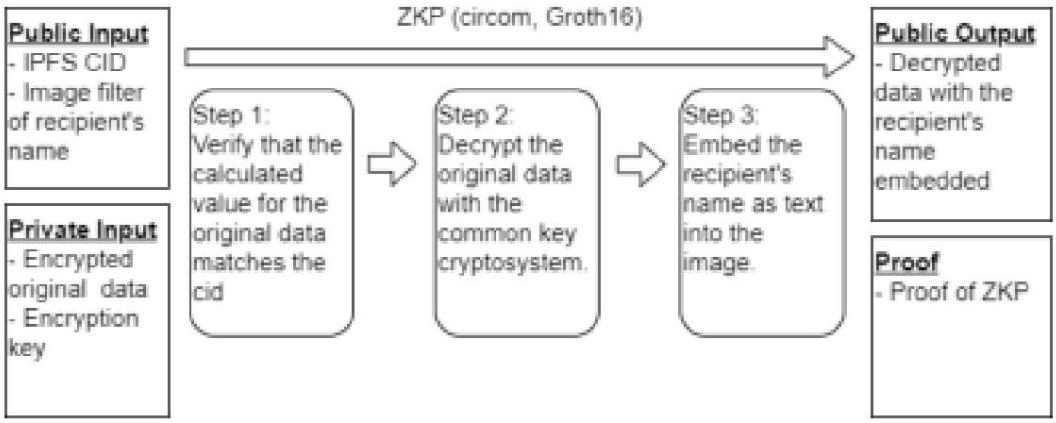


Abbildung 2. Prozessablauf von ZKP: ein Beispiel für die Bilddatenverarbeitung.

Dabei gehen wir von folgenden Annahmen aus: •

Private IPFS wird von einer begrenzten Anzahl von Mitgliedern betrieben. •
Daten werden in IPFS in einem verschlüsselten Format unter Verwendung symmetrischer Kryptografie gespeichert. • Der Verschlüsselungsschlüssel wird ausschließlich von einem einzelnen Mitglied gehalten, d. h. einem Administrator.

Das vorgeschlagene System erleichtert die Erstellung eines ZKP-Nachweises durch den Schaltkreis durch das Mitglied, das den Verschlüsselungsschlüssel besitzt (entspricht einem Administrator). Die Ein- und Ausgaben (außer Beweisen) dieses Prozesses sind wie

folgt: • Öffentliche

Eingabe: – CID der ursprünglich verschlüsselten
Daten; – Ein Filter zum Einbetten des Namens des Empfängers in die

Daten; • Private

Eingabe: – Verschlüsselte
Originaldaten; –

Verschlüsselungsschlüssel; • Ausgabe:

– Entschlüsselte Daten mit eingebettetem Namen des Empfängers.

Beachten Sie, dass die ZK-optimierte Implementierung (Abschnitt 6.1) dem öffentlichen Input mehr Informationen hinzufügt. Insbesondere umfasst der interne Prozess von ZKP die folgenden Schritte:

1. Berechnen der CID der ursprünglich verschlüsselten Daten, um eine Übereinstimmung mit den verschlüsselten Da
registrierte CID.

Machine Translated by Google

2. Entschlüsseln der Originaldaten mithilfe einer symmetrischen Kryptografie.
3. Einbettung des Empfängernamens in die entschlüsselten Bildrohdaten basierend auf dem Filter im öffentlichen Input bereitgestellt.

Im vorgeschlagenen System konzentrieren wir uns nicht nur auf Bilddaten, sondern verwenden sie als Beispieldaten, um die Anwendbarkeit des vorgeschlagenen Schemas zu überprüfen. Filter werden verwendet, um die Effizienz der Verarbeitung innerhalb des Schaltkreises zu verbessern. Da das Einbetten von Namensdaten in den Schaltkreis rechenintensiv ist, wird ein erheblicher Teil der Bildverarbeitung im Voraus außerhalb des Schaltkreises durchgeführt und ein Filter erstellt. Mithilfe öffentlicher Eingaben und Beweise kann der Empfänger überprüfen, ob die entschlüsselten Daten (d. h. die Ausgabe) mit ihrem eingebetteten Namen **aus den Originaldaten (die in der privaten Eingabe enthalten sind) generiert wurden, die** mit der CID verwaltet werden. Der Empfänger **kann bei mindestens einem Mitglied des Netzwerks nachfragen, um die Existenz der CID im privaten IPFS zu bestätigen.**

Zusammenfassend ermöglicht der oben genannte Prozess dem Empfänger die Überprüfung der empfangenen Daten, indem Sie die folgenden Aufgaben ausführen:

- **Überprüfen Sie, ob die empfangenen Daten aus den mit der CID von verwalteten Daten generiert wurden privates IPFS,**
- **bestätigen Sie anhand des Nachweises, dass der gesamte Vorgang korrekt durchgeführt wurde, ohne die verschlüsselten Daten oder den Verschlüsselungsschlüssel direkt zu kennen, und • überprüfen Sie durch Nachfrage bei mindestens einem Netzwerkmitglied, ob die CID konkret innerhalb des privaten IPFS existiert.**

Das Neue an dem vorgeschlagenen System ist, dass es die Überprüfung der Datenauthentizität ermöglicht, indem mindestens einem Mitglied des Netzwerks vertraut wird, selbst wenn der Empfänger nicht über den Verschlüsselungsschlüsse (Für den Empfänger ist es selbstverständlich, mindestens einem Mitglied eines bestimmten Mehrmitgliedersystems zu vertrauen . Wenn keinem der Mitglieder vertraut werden kann, besteht nur ein geringer Anreiz, von diesem Netzwerk verwaltete Daten zu empfangen.)

5. Potentielle Anwendungen In

der Medizinbranche werden Patientendiagnosedaten über mehrere medizinische Einrichtungen hinweg verwaltet. Mit dem vorgeschlagenen System können Patienten überprüfen, ob die Daten, die sie erhalten, tatsächlich in privaten IPFS verwaltet werden, um die Authentizität sicherzustellen. Außerdem wird hier eine All-in-One-Plattform für die Forschung und Entwicklung des maschinellen Lernens mit medizinischen Bildern vorgeschlagen [17]. Auf dieser Plattform werden anonymisierte medizinische Bilder in privaten IPFS verwaltet, die von einer Gruppe medizinischer Einrichtungen betrieben werden. Das System ermöglicht es Forschern im Bereich des maschinellen Lernens, die sich außerhalb des Netzwerks befinden, zu überprüfen, ob die Bilddaten tatsächlich im privaten IPFS verwaltet werden. Darüber hinaus können medizinische Einrichtungen durch die Einbettung der Informationen über die Forscher im Bereich des maschinellen Lernens in die Bilddaten das Risiko einer Sekundärverteilung verringern.

Wenn die Anwendung nicht auf die Einbettung des Empfängernamens beschränkt ist, können die potenziellen Anwendungen des vorgeschlagenen Systems noch weiter erweitert werden. Stellen Sie sich beispielsweise ein Szenario vor, in dem ein bestimmtes Unternehmen ein privates IPFS einrichtet, um vertrauliche Dokumente zwischen seinen Konzernunternehmen auszutauschen. Wenn Mitarbeiterdaten enthalten sind, können diese durch das Verbergen privater Daten und deren Weitergabe an externe Stellen die Zugehörigkeit der Mitarbeiter zum Unternehmen bestätigen und gleichzeitig Nehmen wir außerdem an, dass eine Universität ein privates IPFS eingerichtet hat, um nur dem akademischen Personal Zugriff auf die Leistungsdaten der Studenten zu gewähren. In diesem Fall können die Studenten überprüfen, ob ihre erhaltenen Leistungsdaten tatsächlich im privaten IPFS verwaltet werden.

Das vorgeschlagene System unterstützt also einen Hybridfall – die Verteilung interner Daten an spezifische externe Entitäten nach Bedarf – in realen Umgebungen weit verbreitet.

6. Implementierung Das

vorgeschlagene System wurde zur Verarbeitung von Bilddaten mit circom implementiert, einem renommierten Tool, das auf die Konstruktion von zk-SNARK-Schaltkreisen spezialisiert ist. Circom ermöglicht die Beschreibung von Rechenprozessen innerhalb eines Schaltkreises mithilfe seiner einzigartigen Sprache, und die

Machine Translated by Google

Die nach der Kompilierung generierte ausführbare Datei kann über die JavaScript-Bibliothek snarkjs aufgerufen werden. Diese Anordnung ermöglicht die Beschreibung von Schaltungsprozessen in circom, und die externe Verarbeitung und die Prüfung der Schaltungskorrektheit werden mit JavaScript durchgeführt. Für die zk-SNARKs Schema, Groth16 wurde verwendet; es ist bekannt für seine relativ schnellere Ausführungsgeschwindigkeit als andere zk-SNARKs-Schema.

Im Einzelnen haben wir an zwei Implementierungsarten für Bilddaten gearbeitet: einem Standard Implementierung mit allgemeinen kryptographischen Techniken und eine ZK-optimierte Implementierung mit ZK-freundlichen kryptographischen Techniken zur Reduzierung der Rechenz Schaltung. Diese Implementierungen wurden zum Vergleich der benötigten Rechenzeiten verwendet. ZKP-Schaltungen erfordern beträchtliche Rechenzeit, selbst für Berechnungen, die leicht durch Computer-Software gehandhabt werden (dies ist besonders spürbar bei der Arbeit mit Bilddaten). Daher ist die Rechenleistung für den praktischen Einsatz von entscheidender Bedeutung.

6.1. Standardimplementierung

Abschnitt 4 beschreibt die Dateneingabe in die Schaltung. Zur Vereinfachung der In-Circuit-Die ursprünglichen verschlüsselten Daten wurden als Bitmap-Bilddaten formatiert, die kompatibel sind mit mit OS/2-Standards. Die ersten 54 Bytes der Bilddaten enthalten Informationen wie die Breite, Höhe und Farbtiefe von Bildern [18]. Die Farbtiefe beträgt 8 Bit und jede Farbe Jeder Komponente in RGB wird ein Byte zugewiesen, was zu einer Darstellung von 3 Bytes pro Pixel führt.

Zunächst prüft das System, ob die verschlüsselten Daten, die als private Eingabe eingegeben wurden, stimmt mit der als öffentliche Eingabe angegebenen CID überein. Wenn sie nicht übereinstimmen, gibt das System an, ein Fehler und die als Ausgabe ausgegebenen Bilddaten sind eine Bytefolge, in der alle Werte sind 0x00. CID dient als entscheidender Mechanismus zur eindeutigen Identifizierung von Dateien und effizient Daten aus IPFS abrufen. CID gibt es in zwei Versionen: V0 und V1 [19]. Dabei ist die flexiblere Version CID V1 wurde verwendet. CID enthält einen Hash der jeweiligen Daten, wodurch unterschiedliche Daten haben unterschiedliche CIDs. Normalerweise wird CID V1 mithilfe des SHA256-Hashes berechnet Funktion, und die Standardimplementierung verwendet SHA256 zur Berechnung der CID.

Die Datenstruktur von CID V1 ist in Tabelle 2 dargestellt.

Tabelle 2. CID V1-Datenstruktur.

Byte Position	Beschreibung	Mehrwert bei der Umsetzung
Erstes Byte	CID-Version	Nummer:
Zweites Byte	Multibase-Präfix	0x55: Rohdaten
Drittes Byte	Hash-Funktionskennung	0x12: SHA-256
Viertes Byte	Hash-Länge	0x20: 32 Byte
Ab dem fünften Byte	Hashwert	SHA-256-Hashwert (32 Bytes)

Die Kodierung für CID erfolgt mit Base32. Base32 kodiert eine Folge von Bytes Basierend auf dieser Struktur erstellt, um CID zu generieren.

Innerhalb der Schaltung wird der eingegebene CID-Wert aus Base32 dekodiert und das System prüft, ob der extrahierte Hashwert mit dem SHA256-Hash übereinstimmt, der aus dem verschlüsselte Daten.

Anschließend werden die verschlüsselten Daten entschlüsselt. Als Verschlüsselungsverfahren wird AES-CTR verwendet. Algorithmus, der eine Art symmetrische Kryptographie ist. Die AES-CTR-Verschlüsselung und Entschlüsselung in circom-chacha20 [20] wurde verwendet. Zur Entschlüsselung wurde AES-CTR-Verschlüsselung verwendet, der Verschlüsselungsschlüssel und die Nonce, die bei der Verschlüsselung verwendet werden, sind erforderlich. Sie werden in die Schaltung als 256-Bit-Schlüssel bzw. 128-Bit-Nonce als private Eingänge. Außerdem AES-CTR verarbeitet Datenmengen in Vielfachen von 16. Wenn also die Länge des Bildes Daten vor der Verschlüsselung sind kein Vielfaches von 16, Nullen (0x00) werden an das Ende der Daten angehängt um es dieser Anforderung anzupassen.

Schließlich wird ein Filter auf die entschlüsselten Daten angewendet, um den Namen des Empfängers einzubetten. Die Implementierung der Texteinbettung direkt in der Schaltung kann die Rechenlast erheblich erhöhen ; daher wird außerhalb der Schaltung ein Filter erstellt, der eine erhebliche

Machine Translated by Google

Teil der Bildverarbeitung im Voraus. Die für den Text, der den Namen des Empfängers darstellt, verwendete Schriftart ist die Schriftart Misaki [21]. Der Filter wird dann verwendet, um die Verarbeitung innerhalb der Schaltung zu optimieren. Der Filter ist eine Liste von Zahlen, wobei Werte von 0 bis 255 verwendet werden, um die Farbe jedes Pixels zu ändern, falls sie von der Farbe der Pixel im Originalbild abweicht; außerdem gibt ein Wert von 300 an, dass die Farbe wie im Originalbild bleiben soll. Dieser Filter stellt die Position im Bild dar, an der der Name des Empfängers eingefügt werden soll.

Innerhalb der Schaltung werden die angegebenen Pixelfarben in den entschlüsselten BMP-Daten basierend auf diesem Filter geändert.

6.2. ZK-optimierte Implementierung

Die ZK-optimierte Implementierung ändert die Hash-Funktion, die Verschlüsselungstechnologie und die In-Circuit-Verarbeitung zur standardmäßigen ZK-freundlichen Verschlüsselungsimplementierung. Diese Implementierung verbessert die Rechenleistung und bewertet den Unterschied in der Rechengeschwindigkeit zwischen ZK-freundlicher Verschlüsselung und allgemeiner Verschlüsselung nicht. Daher wurde auch die In-Circuit-Verarbeitung geändert.

Als Hash-Funktion zur Berechnung der CID wurde **Poseidon-Hash [22] verwendet. Die Verwendung des Poseidon-Hash für CIDs wird nicht offiziell unterstützt; daher wurde er speziell für diese Studie entwickelt. Obwohl SHA256 häufig in allgemeinen Berechnungen verwendet wird, erfordert es in ZKP-Schaltungen beträchtliche Rechenzeit. Der Poseidon-Hash ist in circom und JavaScript implementiert (circomlib [23] bzw. circomlibjs [24]). Er wird über einen endlichen Körper mit Primzahlordnung berechnet und kann bis zu 16 Eingangsvariablen akzeptieren.**

Die verwendete Reihenfolge ist kleiner als das Maximum von 32 Bytes, aber größer als das Maximum von 31 Bytes. Dies bedeutet, dass jeder der 16 Eingänge Daten enthalten darf, die diese Reihenfolge nicht überschreiten. In dieser Implementierung werden die für die Hash-Berechnung vorgesehenen Daten als Eingabewerte in 31-Byte-Segmente unterteilt. Wenn die Unterteilung 16 Segmente überschreitet, wird der Poseidon-Hash für die ersten 16 Segmente berechnet. Dieses Ergebnis wird zu den nächsten 15 Datensegmenten für einen nachfolgenden Poseidon-Hash-Eingang addiert. Der Vorgang wird wiederholt, bis alle Eingabedaten für die Hash-Berechnung verwendet wurden. Wenn der letzte Eingang rechnerisch nicht 16 Segmente vervollständigt, werden die fehlenden Eingänge auf Null gesetzt, u

Beim Generieren von CID aus dem Poseidon-Hashwert sollte die Bytefolge der CID V1-Datenstruktur folgen und Base32-codiert sein. Um die Rechenzeit jedoch weiter zu verkürzen , wird bei dieser Implementierung die Base32-Codierung weggelassen und der Poseidon-Hashwert direkt als Ersatz für CID verwendet. Das Aufteilen der Eingabedaten in 16 Segmente innerhalb des Schaltkreises ist rechenintensiv; daher wird diese Aufteilung außerhalb des Schaltkreises durchgeführt und als Eingabe bereitgestellt. In diesem Fall werden die verschlüsselte Datenbytefolge und die Werteliste zur Berechnung des Poseidon-Hashes als öffentliche Eingaben bereitgestellt, wodurch überprüft werden kann, ob beide Datensätze dieselben Informationen darstellen. Empfänger können bestätigen, dass die für den Poseidon-Hash berechneten Daten und die im Schaltkreis entschlüsselten Daten identisch sind, indem sie diese beiden Werte gegenseitig konvertieren und überprüfen. Da Benutzer in diesem Fall die entschlüsselten Daten erhalten können , besteht die Sorge, dass Passwörter durch Brute-Force-Angriffe oder andere Mittel verloren gehen.

Als Verschlüsselungstechnologie haben wir anstelle der AES- CTR-Verschlüsselung die Poseidon-Verschlüsselung **[25]** übernommen . **Bei der Poseidon-Verschlüsselung, die in circom und TypeScript implementiert ist (poseidon- encryption-circom2 [26]), wird der öffentliche Schlüssel des Empfängers empfangen, ein gemeinsamer Schlüssel generiert und eine sichere Ver- und Entschlüsselung durch beide Parteien sichergestellt. In diesem Fall wird jedoch direkt ein gemeinsamer Schlüssel generiert und zur Ver- und Entschlüsselung verwendet. Der Schaltung werden zwei Werte, die die Koordinaten einer elliptischen Kurve darstellen, und ein Nonce-Wert als private Eingänge für die Verschlüsselung bereitgestellt . Darüber hinaus wird der Filter auf die gleiche Weise wie in der Standardimplementierung implementie**

7. Auswertung

Wir haben ein Beispielprogramm auf Basis der oben genannten Implementierungen erstellt, das circom zur Beschreibung der Schaltung und snarkjs zur Ausführung der Schaltung und zur Überprüfung verwendet

Machine Translated by Google

Beweise. Als kryptographische Bibliotheken stehen circom-chacha20 [20], circomlib [23], circomlibjs [24], und poseidon-encryption-circom2 [26] wurden verwendet.

Für jeden Schaltkreis wurden die Standard- und die ZK-optimierte Implementierung umgesetzt. und ihre Rechenzeiten wurden während der Ausführung verglichen. Weiße Bitmap-Bilder waren die Zielbilder, und die Experimente wurden mit dem Buchstaben „A“ als eingebettetes Zeichen. Da das Einbetten beliebiger Zeichen die Verarbeitung nicht verändert Durch das Einbetten eines einzelnen Zeichens durch den Filter wurde ein Vergleich der Berechnungszeiten ermöglicht. Darüber hinaus haben wir die Bildgrößen variiert, um die Ausführungszeiten für jeden Schaltkreis zu messen. Die verwendeten Größen waren 10 × 10, 15 × 15, 30 × 15, 30 × 30, 60 × 30, 60 × 60, 120 × 60, 120 × 120, und 180 × 120 Pixel. Als Ausführungsumgebung diente Windows 11 mit einem Ryzen 9 3950X CPU und 128 GB RAM, betrieben unter Ubuntu 22.04 in einer WSL2-Umgebung.

Abbildung 3 zeigt ein Beispielbild, das von der Schaltung erzeugt wird, speziell für das 60 × 30 Pixel große Größe mit der ZK-optimierten Implementierung. Die Ergebnisse für jede Bildgröße werden präsentiert in Tabelle 3, wo nichtlineare Beschränkungen die Anzahl der nichtlinearen Beschränkungen in der Schaltung, Build-Zeit ist die Zeit, die benötigt wird, um Circom zu kompilieren und die Schaltung auszugeben, und Beweis gen time ist die Zeit, die benötigt wird, um Beweise mit der Schaltung zu generieren. Da die Standardimplementierung AES-CTR-Verschlüsselung verwendet, werden am Ende Daten mit 0x00 angehängt, um sich Die Eingabegröße ist ein Vielfaches von 16.

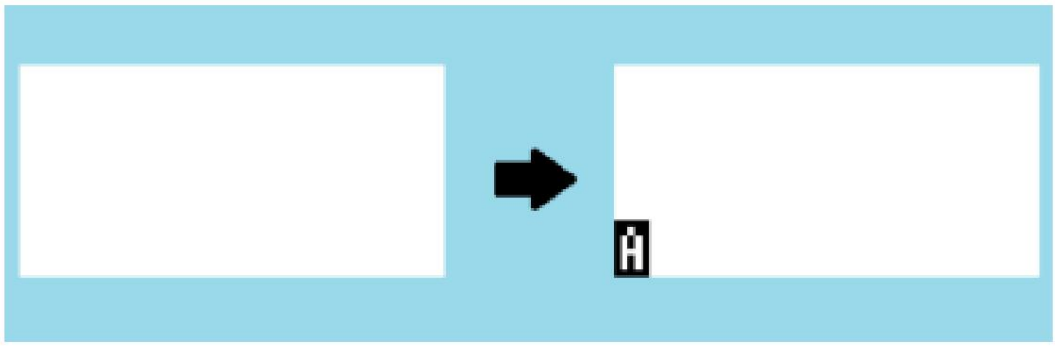


Abbildung 3. Ein vom Schaltkreis für eine Bildgröße von 60 × 30 durch ZK-optimierte Implementierung generiertes Bild.

Tabelle 3. Vergleich der Ausführungszeiten der Schaltung.

	Pixel	Bildgröße [Byte]	Nichtlinear Einschränkungen	Bauzeit [MS]	Beweis Gen Zeit [ms]
Standard					
	10 × 10	384	558.341	668.220	14.377
	15 × 15	784	1.095.292	1.316.666	25.545
	30 × 15	1440	1.980.688	1.696.370	35.161
	30 × 30	2816	3.867.989	3.657.578	65.785
	60 × 30	5456	7.453.300	7.864.583	126.262
ZK-optimiert					
	10 × 10	376	35.407	128.979	3076
	15 × 15	775	72.725	173.210	4032
	30 × 15	1435	134.663	275.630	5900
	30 × 30	2815	263.450	470.872	9733
	60 × 30	5455	509.375	850.612	17.571
	60 × 60	10.855	1.013.483	1.257.418	29.044
	120 × 120	43.255	4.036.913	6.754.928	96.580

In Standard- und ZK-optimierten Implementierungen für 60 × 60 Pixel und 180 × 120 Pixel Bildgrößen, dem System ging der Speicher aus und die Berechnung konnte nicht abgeschlossen werden. In der ZK-optimierten Implementierung wurde die Anzahl der nichtlinearen Einschränkungen reduziert auf ungefähr ein Zehntel der Standardimplementierung bei gleicher Bildgröße.

Machine Translated by Google

Dadurch verkürzte sich die Zeit für den Aufbau und die Proof-Generierung. Die maximal verarbeitbare Bildgröße betrug jedoch immer noch nur bis zu 120 × 120 Pixel, was für praktische Anwendungen erheblich klein ist.

8. Diskussion

Obwohl bei der ZK-optimierten Implementierung ZK-freundliche kryptografische Technologien verwendet und In-Circuit-Prozesse optimiert wurden, betrug die maximal verwaltbare Bildgröße ungefähr 120 × 120 Pixel. Dies beschränkt den praktischen Nutzen auf deutlich kleinere Bildgrößen. Die Forschung zur Verbesserung der Leistung von ZKPs läuft jedoch weiter, und zukünftige technologische Fortschritte könnten die Verarbeitung größerer Bildgrößen ermöglichen.
Beispielsweise erreichten Zhang et al. [27] eine zehnfache Beschleunigung von zk-SNARKs durch den Einsatz von ASICs. Ma et al. [16] verwendeten eine ähnliche Grafikverarbeitungseinheit, um die Beweisgenerierungszeit zu beschleunigen , und erreichten im Vergleich zu herkömmlichen Methoden eine bis zu 48,1-mal schnellere Leistung. Darüber hinaus wurden Methoden zur Vereinfachung von Rechenprozessen vorgeschlagen, wie beispielsweise die „Faltungsmethode“. Diese Methode komprimiert die zu beweisenden Aussagen [28]. Da zunehmend an Geschwindigkeitsverbesserungen geforscht wird, wird wahrscheinlich auch der Speicherverbrauch optimiert. Dies wird in Zukunft möglicherweise die Verarbeitung größerer Bildgrößen ermöglichen.

Darüber hinaus haben wir festgestellt, dass unsere vorgeschlagene Methode Datengrößen von etwa 10 KB verarbeiten kann. Obwohl die direkte Anwendung unseres Vorschlags auf realistische Bilddaten (von mehreren MB bis zu Dutzenden von MB) eine Herausforderung darstellt, könnte die Aufteilung der Daten in Blöcke durch Änderung der Verschlüsselung und eingebetteter Zeichenfolgen die Anwendung machbar mac

Darüber hinaus erfordert unsere Implementierung einen Wert basierend auf der Größe der zu verarbeitenden (verschlüsselnden) Originaldaten als Argument während der Schaltungsgenerierung. Daher muss für alle Daten eine Schaltung generiert werden. Die Schaltungsgenerierungszeit (Build-Zeit) erhöht sich mit der Bilddatengröße erheblich; beispielsweise dauert die Generierung einer Schaltung für eine Bildgröße von 120 × 120 selbst bei einer ZK-optimierten Implementierung mehr als 112 Minuten (6.754.928 ms). Sobald die Schaltung jedoch erstellt ist, ist die Proof-Erstellungszeit unter denselben Bedingungen kurz und beträgt etwa 97 s (96.580 ms). Mit anderen Worten: Sobald eine Schaltung erstellt ist, ist die Proof-Erstellung nicht mehr zeitintensiv. Diese Tatsache stellt keine praktischen Probleme dar, wenn dasselbe Bild an verschiedene Personen verteilt wird.

In ZK-freundlichen Implementierungen werden verschlüsselte Daten als öffentliche Eingabe eingegeben. Der Umgang mit Verschlüsselungsschlüsseln für Bilder erfordert sorgfältige Überlegung. In privaten IPFS verwaltete Daten sind verschlüsselt. Wenn jedoch Verschlüsselungsschlüssel durchsickern, könnten die verschlüsselten Daten entschlüsselt werden. Daher sollten bestimmte Benutzer, die private IPFS verwalten, Administratoren werden, um die Schlüssel sorgfältig zu verwalten, oder es könnte eine Blockchain vom Typ eines Konsortiums im selben Netzwerk eingerichtet werden, um Zugriffsrechte entsprechend festzulegen und

9. Schlussfolgerungen

Hier wurde eine neue Methode vorgeschlagen, um in privatem IPFS gespeicherte Daten an externe Entitäten zu verteilen und gleichzeitig ihre Authentizität überprüfbar zu machen. Die Methode wendet eine Art ZKP, zk- SNARKs, an, um die CID der Daten zu überprüfen und den Namen des Empfängers einzubetten. Mit diesem Ansatz können externe Entitäten überprüfen, ob die empfangenen Daten aus den Originaldaten im privaten IPFS generiert wurden, ohne dass Details wie IPFS-Zugriffsrechte und Verschlüsselungsschlüssel erforde

Um **die Rechenleistung der vorgeschlagenen Methode zu verbessern,** wurden **eine Standardimplementierung mit herkömmlichen kryptografischen Techniken und eine ZK- optimierte Implementierung mit ZK-freundlichen kryptografischen Schemata implementiert .** Experimente mit einem **Beispielprogramm bestätigten die Wirksamkeit der vorgeschlagenen Methode für eine Bilddatengröße von bis zu 120 × 120 Pixeln.**

Diese vorgeschlagene Methode erweitert den nutzbaren Bereich dezentraler Speichersysteme auf einen Hybridfall – die Verteilung interner Daten an bestimmte externe Einheiten nach Bedarf. Diese Studie ebnet einen neuen Weg für den Austausch sensibler Informationen zwischen verschiedenen Sektoren innerhalb und außerhalb einer Gruppe. Für die breite praktische Anwendbarkeit der vorgeschlagenen Methode auf größere und vielfältigere Datentypen wie Bilder und Videos muss die Verarbeitungsgeschwindigkeit jedo

Machine Translated by Google

verbessert werden und es müssen Methoden zur Datenaufteilung eingesetzt werden, die im Rahmen unserer zukünftigen Studien liegen.

Beiträge der Autoren: Konzeptualisierung, KS, CH, TTC und WO; Schreiben – Vorbereitung des Originalentwurfs , KS; Schreiben – Überprüfung und Bearbeitung, KS und KI; Aufsicht, GM; Projektverwaltung, KS. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr

Finanzierung: Diese Forschung wurde als Gemeinschaftsforschungsprojekt zwischen der Universität Tokio und Callisto Inc. durchgeführt und von Callisto Inc. finanziert. Diese Arbeit wurde vom Stiftungslehrstuhl für Blockchain-Innovation und dem Mohammed bin Salman Center for Future Science and Technology for Saudi-Japan Vision 2030 (MbSC2030) an der Universität Tokio unterstützt.

Datenverfügbarkeitserklärung: Der für die Simulationen verwendete Quellcode ist auf GitHub verfügbar. https://github.com/blockchaininnovation/circom_image_processing (abgerufen am 21. März 2024).

Interessenkonflikte: Die Autoren Changhee Han, Tsz Tat Chu und Wataru Ozaki waren bei der Firma Callisto Inc. angestellt. Die übrigen Autoren erklären, dass die Forschung ohne kommerzielle oder finanzielle Beziehungen durchgeführt wurde, die als potenzieller Interessenkonflikt au

Verweise

1. **Groth, J. Zur Größe von paarungsbasierten nicht-interaktiven Argumenten. In Proceedings of the Advances in Cryptology–EUROCRYPT 2016: 35. jährliche internationale Konferenz zur Theorie und Anwendung kryptographischer Techniken, Wien, Österreich, 8.–12. Mai 2016; Proceedings, Teil II 35; Springer: Berlin/Heidelberg, Deutschland, 2016; S. 305–326.**
2. **Kumar, S.; Bharti, AK; Amin, R. Dezentrale sichere Speicherung von Krankenakten mit Blockchain und IPFS: Ein Vergleich Analyse mit zukünftigen Richtungen.** Secur. Priv. **2021, 4, e162. [CrossRef]**
3. **Azbeg, K.; Ouchetto, O.; Andaloussi, SJ BlockMedCare: Ein Gesundheitssystem basierend auf IoT, Blockchain und IPFS für Daten Managementsicherheit.** Ägypten. Inform. J. **2022,** 23, **329–343. [CrossRef]**
4. **Hossan, MS; Khatun, ML; Rahman, S.; Reno, S.; Ahmed, M. Sicherung von Mitfahrdiensten mit IPFS und Hyperledger basierend auf privater Blockchain. In Proceedings der 24. Internationalen Konferenz für Computer- und Informationstechnologie (ICCIT) 2021, Dhaka, Bangladesch, 18.–20. Dezember 2021; IEEE: Piscataway, NJ, USA, 2021; S. 1–6.**
5. **Lin, Y.; Zhang, C. Eine Methode zum Schutz privater Daten in IPFS. In Proceedings der 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Dalian, China, 5.–7. Mai 2021; IEEE: Piscataway, NJ, USA, 2021; S. 404–409.**
6. **Battah, AA; Madine, MM; Alzaabi, H.; Yaqoob, I.; Salah, K.; Jayaraman, R. Blockchain-basierte Mehrparteienautorisierung für den Zugriff auf IPFS-verschlüsselte Daten.** IEEE Access **2020, 8, 196813–196825. [CrossRef]**
7. **Huang, HS; Chang, TS; Wu, JY Ein sicheres Dateifreigabesystem basierend auf IPFS und Blockchain. In Proceedings der 2. International Electronics Communication Conference, Singapur, 8.–10. Juli 2020; S. 96–100.**
8. **Sun, J.; Yao, X.; Wang, S.; Wu, Y. Blockchain-basiertes sicheres Speicher- und Zugriffsschema für elektronische Krankenakten in IPFS.** IEEE Access **2020, 8, 59389–59401. [CrossRef]**
9. **Kang, P.; Yang, W.; Zheng, J. Blockchain-Methode zum Teilen privater Dateispeicher basierend auf IPFS.** Sensors **2022,** 22, **5100. [CrossRef] [PubMed]**
10. **Uddin, MN; Hasnat, AHMA; Nasrin, S.; Alam, MS; Yousuf, MA Sicheres Dateifreigabesystem mit Blockchain-, IPFS- und PKI-Technologien. In Proceedings der 5. Internationalen Konferenz für elektrische Informations- und Kommunikationstechnologie (EICT) 2021, Khulna, Bangladesch, 17.–19. Dezember 2021; IEEE: Piscataway, NJ, USA, 2021; S. 1–5.**
11. **Li, W.; Guo, H.; Nejad, M.; Shen, CC Datenschutzwahrendes Verkehrsmanagement: Ein von Blockchain und Zero-Knowledge-Proof inspirierter Ansatz.** IEEE Access **2020, 8, 181733–181743. [CrossRef]**
12. **Bellés-Muñoz, M.; Isabel, M.; Muñoz-Tapia, JL; Rubio, A.; Baylina, J. Circom: Eine Schaltungsbeschreibungssprache für den Bau Zero-Knowledge-Anwendungen.** IEEE Trans. Dependable Secur. Comput. **2022, 20, 4733–4751. [CrossRef]**
13. **Offizielle Circom-Website. Online verfügbar: https://iden3.io/circom (abgerufen am 24. März 2024).**
14. **Snarkjs Github Repository. Online verfügbar: https://github.com/iden3/snarkjs (abgerufen am 4. Juni 2024).**
15. **ZCash. Online verfügbar: https://z.cash/ (abgerufen am 12. Juli 2024).**
16. **Ma, W.; Xiong, Q.; Shi, X.; Ma, X.; Jin, H.; Kuang, H.; Gao, M.; Zhang, Y.; Shen, H.; Hu, W. Gzkp: Ein GPU-beschleunigtes Zero-Knowledge-Proof-System. In Proceedings der 28. ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Band 2, Vancouver BC Kanada, 25.–29. März 2023; S. 340–353.**
17. **Han, C.; Shibano, K.; Ozaki, W.; Osaki, K.; Haraguchi, T.; Hirahara, D.; Kimura, S.; Kobayashi, Y.; Mogi, G. All-in-one-Plattform für KI-Forschung und -Entwicklung in der medizinischen Bildgebung, die Datenerfassung, -auswahl, -annotation und -vorverarbeitung umfasst.** In Proceedings of the Medical Imaging 2024: Imaging Informatics for Healthcare, Research, and Applications, San Diego, CA, USA, 18.–23. Februar 2024; **SPIE: Bellingham, WA, USA, 2024; Band 12931, S. 311–315.**
18. **Miano, J.** Komprimierte Bilddateiformate: Jpeg, png, gif, xbm, bmp; **Addison-Wesley Professional: Boston, MA, USA, 1999.**

Machine Translated by Google

19. Inhaltskennungen (CIDs). Online verfügbar: https://docs.ipfs.tech/concepts/content-addressing/#cids-are-not-file-hashes (abgerufen am 24. März 2024). 20. circom-chacha20 Github Repository. Online verfügbar: https://github.com/reclaimprotocol/circom-chacha20 (aufgerufen am 24. März 2024).

21. Die 8 × 8 Punkte große japanische Schriftart „Misaki Font“. Online verfügbar: https://littlelimit.net/misaki.htm (abgerufen am 11. Juni 2024 (Auf Japanisch)

22. Grassi, L.; Khovratovich, D.; Rechberger, C.; Roy, A.; Schofnegger, M. Poseidon: Eine neue Hash-Funktion für {Zero-Knowledge} -Beweissysteme. In Proceedings des 30. USENIX Security Symposiums (USENIX Security 21), Vancouver, BC, Kanada, 11.–13. August 2021; S. 519–535.

23. Circomlib Github Repository. Online verfügbar: https://github.com/iden3/circomlib (abgerufen am 21. März 2024).

24. Circomlibjs Github Repository. Online verfügbar: https://github.com/iden3/circomlibjs (abgerufen am 21. März 2024).

25. Khovratovich, D. Verschlüsselung mit Poseidon. 2019. Online verfügbar: https://drive.google.com/file/d/1EVrP3DzoGbmzkRmYnyEDcIQcXVU7GlOd/Ansicht (abgerufen am 19. Juli 2024).

26. Poseidon-Encryption-Circom2 Github Repository. Online verfügbar: https://github.com/Shigoto-dev19/poseidon-encryption-c ircom2 (abgerufen am 21. März 2024).

27. Zhang, Y.; Wang, S.; Zhang, X.; Dong, J.; Mao, X.; Long, F.; Wang, C.; Zhou, D.; Gao, M.; Sun, G. Pipezk: Beschleunigung von Zero- Knowledge-Beweisen mit einer Pipeline-Architektur. In Proceedings des 2021 ACM/IEEE 48th Annual International Symposium on Computer Architecture (ISCA), Valencia, Spanien, 14.–18. Juni 2021; IEEE: Piscataway, NJ, USA, 2021; S. 416–428.

28. Kothapalli, A.; Setty, S.; Tzialla, I. Nova: Rekursive Zero-Knowledge-Argumente aus Faltungsschemata. In Proceedings der Annual International Cryptology Conference, Santa Barbara, CA, USA, 15.–18. August 2022; Springer: Cham, Schweiz, 2022; S. 359–388.

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultieren.

Machine Translated by Google

Elektronik

MDPI

Artikel

Ein autonomer kooperativer Navigationsansatz für Mehrere unbemannte Bodenfahrzeuge in einer variablen Kommunikationsumgebung

Xudong Lin und Mengxing Huang *

Fakultät für Informations- und Kommunikationstechnik, Hainan-Universität, Haikou 570228, China; linx424523604@163.com

*** Korrespondenz: huangmx09@163.com**

Zusammenfassung: Roboter unterstützen Einsatzkräfte, indem sie wichtige Informationen aus der Ferne sammeln. Der Einsatz mehrerer kooperativer unbemannter Bodenfahrzeuge (UGVs) für einen Einsatz kann die Reaktionszeit verkürzen, die Situationswahrnehmung verbessern und die Kosten minimieren. Eine zuverlässige Kommunikation ist für mehrere UGVs für Umwelteinsätze von entscheidender Bedeutung, da mehrere Roboter Informationen für eine kooperative Navigation und Datenerfassung austauschen müssen. In dieser Arbeit untersuchen wir eine Steuerungsrichtlinie für eine optimale Kommunikation zwischen mehreren UGVs und Basisstationen (BSs). Ein Multi-Agent Deep Deterministic Policy Gradient (MADDPG)-Algorithmus wird vorgeschlagen, um die Steuerungsrichtlinie für das maximale Signal-Interferenz-Verhältnis zu aktualisieren. Die UGVs kommunizieren sowohl mit den festen BSs als auch mit einer mobilen BS. Die vorgeschlagene Steuerungsrichtlinie kann die UGVs und die mobilen BS navigieren, um die Kommunikation und Signalstärke zu optimieren. Schließlich wird ein genetischer Algorithmus (GA) vorgeschlagen, um die Hyperparameter des MADDPG-basierten Trainings zu optimieren. Simulationsergebnisse demonstrieren die Rechenleis

Schlüsselwörter: unbemannte Bodenfahrzeuge (UGVs); genetischer Algorithmus (GA); Multi-Agent Deep Deterministic Policy Gradient (MADDPG); autonome Navigation

check for updates

Zitat: Lin, X.; Huang, M. An Autonome kooperative Navigation Ansatz für mehrere unbemannte Bodenfahrzeuge in einer Variable Kommunikationsumgebung. Electronics **2024,** 13, **3028. https:// doi.org/ 10.3390/electronics13153028**

Wissenschaftliche Herausgeber: Eric Matson und Young Im Cho

Empfangen: 18. Juni 2024
Überarbeitet: 25. Juli 2024
Akzeptiert: 30. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

Ein Netzwerk aus verteilten unbemannten Bodenfahrzeugen (UGVs) und einem zentralen Controller wird als Multi-UGV-Steuerungssystem bezeichnet [1]. Dieses System ermöglicht autonome Dominanz, autonome Navigation und autonome Zusammenarbeit. Es kann entweder innerhalb eines begrenzten Bereichs oder als Teil eines größeren Transportsystems betrieben werden. Multi-UGV-Steuerungssysteme bieten einen einzigartigen Ansatz zur Navigation, der äußerst zuverlässig, wirtschaftlicher und energiesparend ist . In den letzten Jahren hat die dringende Nachfrage nach Multi-UGV-Navigationssystemen zu zunehmenden Diskussionen in der Wiss

Die Navigation von UGVs in einer Kommunikationsumgebung war Gegenstand von Forschungen [6], und traditionelle Optimierungsmethoden haben gute Ergebnisse erbracht [7]. Um ein autonomes Navigationssystem zu entwickeln, entwickelten D. Chen et al. **[8]** einen heuristischen Monte-Carlo **-Algorithmus, der auf einer diskreten Hough-Transformation und Monte-Carlo-Lokalisierung basiert, wodurch eine niedrige Komplexität für die Verarbeitung in Echtzeit gewährleistet wird. Anders als bei der Innovation von Algorithmen , um in unbekannten und überfüllten Umgebungen robust zu funktionieren, haben HU Unlu et al. [9] einen robusten Ansatz für visuell unterstützte Trägheitsnavigation entwickelt, der Unsicherheiten standhalten kann . Anders als bei der Verwendung visueller Hilfsmittel ließen sich X. Lyu et al. [10] von einer geometrischen Sichtweise inspirieren und entwarfen ein neues adaptives, Sharing-Faktor-integriertes Navigationsinformationsfusionstechnologieschema, das eine adaptive Navigation bei nichtlinearen Systemen bietet und eine nicht-Gauß-Verteilung verwendet. Diese oben erwähnten traditionellen Optimierungsmethoden sind einfach umzusetzen. Allerdings müssen diese Methoden an bestimmte Voraussetzungen geknüpft werden , weshalb sie nur für statische Umgebur**

Machine Translated by Google

Folglich werden Multi-UGV-Systeme bei der Handhabung komplexer Szenarien mit diesen beiden Herausforderungen konfrontiert sein , und um sie effektiv zu bewältigen, ist die Einbindung von maschinellem Lernen (ML) erforderlich [12–14].

Angesichts der rasanten Fortschritte im Bereich ML gibt es gute Gründe, ML-Techniken bei der UGV-Navigation einzusetzen . Um eine bessere Ortungsleistung zu erreichen, haben H. Lee et al. [15] eine ML-Technik zur Berechnung der Distanz zwischen BS und UGVs entwickelt, die eine Lokalisierung ohne zusätzliche Infrastruktur ermöglicht. Anstatt sich auf die direkte Ortungsmessung zu verlassen, haben HT Nguyen et al. ein Koordinierungssystem zwischen unbemannten Luftfahrzeugen und UGVs entwickelt, das eine effektive kollaborative Navigation ermöglicht [15].
Mit zunehmender Komplexität der Simulationsumgebung nimmt die Wirksamkeit der vorgeschlagenen Lösung jedoch rapide ab. Um diese Herausforderung zu bewältigen, ist der Einsatz von Reinforcement-Learning-Algorithmen (RL) eine vielversprechende Wahl. RL betont, wie Agenten die beste Strategie finden können, um alle Belohnungen bei der Interaktion mit der Umgebung zu maximieren, wodurch es sich gut für die Erkundung und Anpassung an zunehmend komplexere Umgebungen eignet [16].

Die Forschung wurde kürzlich durch Diskussionen über die Verwendung von RL zur Lösung des Problems der kooperativen Navigation mehrerer UGVs vorangetrieben [17]. Um Kollisionen mit Hindernissen zu vermeiden, schlugen X. Huang et al. [18] ein innovatives, auf tiefem RL basierendes lokales Pfadplanungssystem für UGVs vor, das multimodale Wahrnehmung nutzt , um Policy-Learning zu erleichtern und flexible Navigationsaktionen zu generieren. Im Unterschied zur Navigation mit einzelnen UGVs schlugen S. Wu et al. [19] zur Verbesserung der durchschnittlichen spektralen Effizienz eine Technologie zur Flugbahnoptimierung auf Basis eines gemeinsamen tiefen deterministischen Policy-Gradienten mit mehreren Agenten (F-MADDPG) vor, der die Fähigkeit von MADDPG übernimmt, mehrere UGVs kooperativ zu steuern, und gemeinsame Mittelwertbildung nutzt, um Datenisolation zu vermeiden und die Konvergenz zu beschleunigen. Mit diesen auf RL basierenden UGV-Navigationsmethoden wurden erhebliche Fortschritte erzielt . Sie übersehen jedoch die Beschränkungen statischer Kommunikationsumgebungen und Konve
Diese beiden Elemente müssen bei der Planung einer kooperativen Navigation in einer Kommunikationsumgebung unbedingt berücksichtigt werden .

Angesichts der Einschränkungen bei der kooperativen Navigation mit Kommunikationsabdeckung für UGVs müssen drei Hauptherausforderungen bewältigt werden: die Schwierigkeit der gleichzeitigen Steuerung von UGVs, die Schwankungen bei der Kommunikationsabdeckung und die Komplexität der kooperativen Steuerungsumgebung für UGVs. Erstens liefern traditionelle Steuerungsmethoden wie Q-Learning [20], Proportional-Integral-Derivative-Steuerung (PID) [21] und Deep Q-Network [22] häufig eine suboptimale Leistung in Bezug auf die Kommunikationsabdeckung, wenn mehrere Agenten gleichzeitig gesteuert werden müssen. Zweitens kommt es angesichts der Schwankungen in der Kommunikationsumgebung während der Navigation mehrerer UGVs häufig zu Bereichen mit schlechter Kommunikation, was eine effektive Zusammenarbeit zwischen mehreren UGVs b
Eine vielversprechende Lösung zur Bewältigung der Herausforderungen der kooperativen Steuerung mehrerer Agenten bieten jedoch Multi-Agenten-RL-Algorithmen [23]. Diese Algorithmen steuern die Zusammenarbeit mehrerer Agenten durch das Paradigma des zentralisierten Trainings und der dezentralen Ausführ
Zusätzlich führen wir in unserem vorgeschlagenen Ansatz eine bewegliche UGV-BS ein, die in die UGVs integriert ist und dynamische Änderungen an der festen Kommunikationsumgebung ermöglicht. Diese Zusammenarbeit unterstützt effektiv die Navigationsaufgaben der UGVs. Die zunehmende Komplexität der konstruierten Umgebung kann jedoch die Effektivität und Konvergenz des Algorithmus beeinträchtigen. Glücklicherweise mildern wir Konvergenzschwierigkeiten, indem wir dynamische Hyperparameter mithilfe eines genetischen Algorithmus (GA) adaptiv aktualisieren [25]. Erfreulicher ist, dass es einige Forschungsarbeiten zur Integration von GA zur Hyperparameter-Optimier
A. Sehgal et al. verwendeten einen GA, um die Rückschau auf die Erfahrung (HER) zu finden, die in einem tiefen deterministischen Richtliniengradienten (DDPG) in einer Robotermanipulationsaufgabe verwendet wird, um dem Agenten zu helfen, das Lernen zu beschleunigen [26]. Anders als bei der Änderung eines einzelnen Parameters schlugen Chen R et al. für das flexible Job-Shop -Scheduling-Problem (FJSP) eine GA-Parameteranpassungsmethode auf Basis von Q-Learning vor, die mehrere Schlüsselparameter im Q-Learning ändert, um höhere Belohnungswerte zu erhalten [27]. Dieser belohnungsbasierte Ansatz neigt jedoch dazu, in die lokale Optimalität zu fallen. Darüber hinaus sind diese Methoden nicht für Szenarien geeignet, in denen die Anzahl der Agenten zunimmt. Um diese Probl

Machine Translated by Google

eine Multi-Agenten-RL-Heuristik zur Lösung des Problems des Handlungsreisenden. Auf diese Weise kann ein GA mit einem neuartigen Crossover-Operator fungiert als Heuristik zur Reiseverbesserung, während MARL als eine Konstruktionsheuristik [28]. Obwohl dieser Ansatz das Risiko lokaler Optimalität vermeidet, Es verzichtet auf den Lernprozess von MARL und verwendet ihn nur als Heuristik, anstatt GA für das Training, was bedeutet, dass der Algorithmus nicht zu viel Aufmerksamkeit auf die Zusammenarbeit zwischen intelligenten Agenten. Liu et al. verwendeten eine dezentralisierte, teilweise beobachtbare Multi-Agenten-Pfadplanungsmethode basierend auf evolutionärem RL (MAPPER) zum Erlernen effektiver lokale Planungsstrategien in gemischten dynamischen Umgebungen. Basierend auf einem Multi-Agent-Reinforcement- Learning-Training verwendeten sie GA, um den ursprünglich trainierten Algorithmus iterativ zu er zu einem komplexeren Modell. Obwohl diese Methode Leistungseinbußen bei langfristigen Aufgaben vermeidet, ist die iterative GA möglicherweise nicht gut an komplexere Umgebungen anpassbar [29].

In unserer Forschung kombinieren wir die Vorteile der oben genannten GA-Papiere und übernehmen das CTDE-Paradigma zur Durchführung von Forschung in einem Multi-Agenten-RL-Framework. Die GA weist unterschiedliche Gewichtungen für Algorithmus-Updates basierend auf dem Beitrag des Übergangs, was bedeutet dass wir den Hyperparametern mehr Aufmerksamkeit schenken, die mehr zur Modellaktualisierung beitragen als diejenigen, die höhere Belohnungswerte erzielen. Dadurch vermeiden wir, in lokale Optimalität bei gleichzeitiger Erhöhung der Agentenanzahl.

Um diese drei Herausforderungen zu bewältigen und eine kooperative Navigation in komplexen Umgebungen wird eine neue Navigationsmethode für die Kommunikationsabdeckung mehrerer UGVs vorgeschlagen , die auf einem tiefen deterministischen Multi-Agenten-Richtliniengradienten mit GA (GA-MADDPG) basiert . Im Folgenden werden die wichtigsten Beiträge der **Navigationsmethode für die Kommunikationsabdeckung mehrerer UGVs zusammengefasst :**

- **Ein umfassendes Multi-Agenten-Muster wird in die Multi-UGV-Kollaboration integriert Navigationssystem und die optimale Koordination mehrerer UGVs innerhalb des Kommunikationsabdeckungsbereichs werden als Echtzeit-Multiagenten-Markov-Entscheidungsprozess formuliert. (MDP)-Modell. Alle UGVs sind als unabhängige Agenten mit Selbststeuerungsfähigkeiten eingerichtet.**
- Es wird **eine Methode zur kollaborativen Navigation mehrerer Agenten mit verbesserter Kommunikationsabdeckung vorgeschlagen. Durch die Einführung einer mobilen Basisstation wird die Komm Umgebung wird dynamisch verändert. Simulationsergebnisse zeigen, dass diese Methode verbessert effektiv die Kommunikationsqualität während der Navigation.**
- **Ein GA-basierter hyperparameteradaptiver Ansatz wird zur Optimierung von UGV vorgestellt Kommunikationsabdeckung und Navigation. Es weist Hyperparametern Gewichte zu nach dem Grad der Algorithmus-Aktualisierung und trifft eine Auswahl auf der Grundlage der Größe des Gewichts bei der nächsten Auswahl, die sich von der herkömmlichen Strategie mit festen Hyperparametern unterscheidet und lokalen Optima entgehen kann.**

Der Aufsatz ist für die verbleibenden Teile wie folgt gegliedert. Die Modellierung von Multi- UGV-Kommunikations- und Navigationssystemen wird in Abschnitt 2 ausführlich erläutert. **Einzelheiten zu der von uns vorgestellten RL-Methode werden in Abschnitt 3 erläutert. Mehrere experimentelle Vergleiche in Abschnitt 4 dienen dazu, die Wirksamkeit unseres Ansatzes zu überprüfen. Abschließend diskutieren wir zuk Forschungsrichtungen und fassen Sie die wichtigsten Punkte des Artikels im Fazitabschnitt 5 zusammen.**

2. MDP für Navigation und Kommunikationsabdeckung für Multi-UGVs in Umgebungen

Um die Entscheidungsfindung von Multi-UGVs in realen Systemen zu emulieren, übernehmen wir ein MDP-Modell. Mit der schnellen Weiterentwicklung des Multi-Agenten-RL hat sich MDP zu ein vertrauenswürdiges Entscheidungsmodell [30]. In dieser Studie konstruieren wir eine komplexe Umgebung bei dem drei UGVs und eine mobile BS zusammenarbeiten und der verschiedene Hindernisse beinhaltet. Darüber hinaus führen wir ein Kommunikationskonzept ein, bei dem die Kommunikation Die Abdeckung wird insgesamt durch vier feste BSs und eine mobile BS bestimmt.

2.1. Problembeschreibung

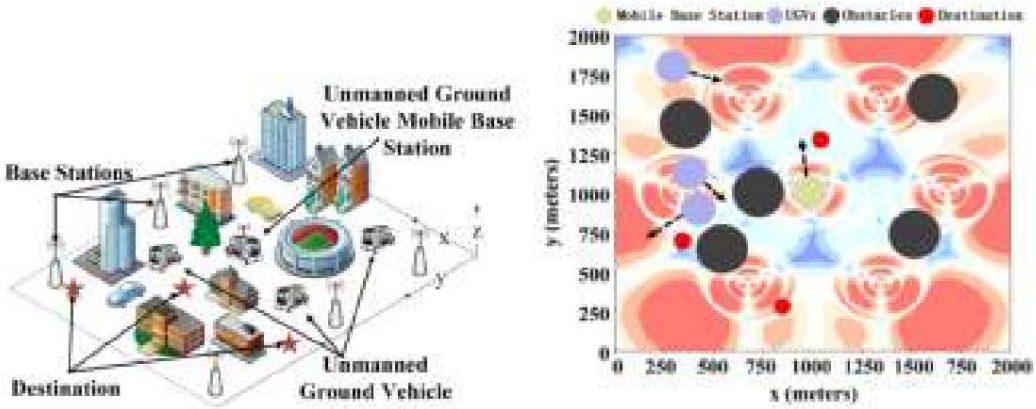
Das Hauptziel unseres Artikels ist die Durchführung von Multi-Agent-Navigationsaufgaben in einem breites Spektrum großflächiger, unbekannter und komplexer Umgebungen so schnell wie möglich. Die Navigationsaufgabe erfordert, dass die UGVs entsprechend unterschiedlicher Umgebungsmerkmale zusammenarbeiten können und in der Lage sind, externe Umgebungsinformationen zu überwinden.

Machine Translated by Google

Störungen und mit der Fähigkeit, Ziele effizient und autonom in Echtzeit zu verfolgen.
Genauer gesagt wird das Ausmaß der Kommunikationsabdeckung gemeinsam von den stationären BSs und der mobilen BS festgelegt. Innerhalb dieses Kommunikationsabdeckungsbereichs navigieren die mobile BS und die UGVs kooperativ. Wir konstruieren ein Aufgabenszenario mit mehreren Optimierungszielen. Das Ziel der UGVs besteht darin, das Ziel erfolgreich zu erreichen, während die mobile BS die Aufgabe hat, die Kommunikationsabdeckung dynamisch in Echtzeit anzupassen, um die Kommunikationsqualität für die UGVs zu optimieren. Die UGVs und die mobile BS führen eine global optimale kooperative Navigation durch, um ihre jeweiligen und gemeinsamen Ziele zu erreichen .

2.2. Modellierung der Umgebung

Unsere Arbeit beinhaltet die Simulation einer realen Umgebung, in der mehrere UGVs zusammenarbeiten, um einen Zielpunkt zu erreichen. Zusätzlich enthält diese Umgebung Hindernisse, die die Bewegung der UGVs behindern und reale Szenarien nachbilden. Wir verwenden eine Multi-Agent-Particle- Umgebung (MPE) [24] als Basisumgebung für unsere sekundäre Entwicklung, wie in Abbildung 1 dargestellt. In dieser Umgebung verwenden wir M **UGVs (wobei** M **als drei definiert ist),** W **mobile BSs (wobei** W **als eins definiert ist) und eine bestimmte Anzahl von Hindernissen. Das Ziel der UGVs besteht darin, gemeinsam Kollisionen zu vermeiden und ihre jeweiligen optimalen Zielpunkte zu erreichen, wobei die Kommunikation im glob**
UGVs wählen eine Hindernisvermeidungsroute mit besserer Kommunikation, um ihre
Bewegung zum Zielpunkt (das Kommunikationsmodell wird in Abschnitt 2.3 näher erläutert).
Die Aufgabe der mobilen BS besteht darin, die Kommunikation für die drei beweglichen Einheiten zu verbessern, indem die Kommunikationsabdeckung im globalen Zustand angepasst wird, der in Abbildung 1b dargestellt ist.



(A) **(B)**

Abbildung 1. Schematische Darstellung der Zusammenarbeit eines Schwarms von UGVs in einer kommunikationsfähigen Umgebung. (a) 3D-städtische Umgebung. (b) Draufsicht der visualisierten Kommunikation

2.3. Modellierung der Kommunikationsabdeckung

In unserer Simulation haben wir die Kommunikation in die MPE-Umgebung integriert und sie als Kriterium zur Bewertung der Aufgabenerledigung verwendet. In diesem Unterabschnitt stellen wir das von uns übernommene Kommunikationskanalmodell vor, zusammen mit dem Kommunikationsmodell, das von der Bewegung der mobilen BS beeinflusst wird, wie in Abbildung 1b dargestellt. Der Kommunikationsbereich innerhalb des mittleren roten Kreises variiert mit dem Standort der be
Beachten Sie, dass Abbildung 2a–l Diagramme darstellen, die zeigen, wie sich die Kommunikationsumgebung mit der Bewegung der mobilen BS in Schritt t ändert . Die mobile BS ist **in Abbildung** 2a **zunächst in der Mitte positioniert und bewegt sich allmählich in Richtung der unteren rechten Ecke, wie in Abbildung 2l dargestellt . Diese Verlagerung der mobilen BS wird durch ihre Beobachtung des Bewegungsmusters der UGVs veranlasst. Folglich wird die mobile BS von der Mitte in die untere rechte Ecke verlagert, um die Kommunikationsqualität in diesem Bereich zu verbessern und so die rote Abdeckungszone zu erweit**

Machine Translated by Google

Mobile BS führt zu einer Reduzierung des Abdeckungsbereichs mit besserer Kommunikationsqualität im oberen linken Quadranten. Um die sich entwickelnde Kommunikationsqualität hervorzuheben und die Kommunikationsänderungen deutlicher zu machen, haben wir außerdem die Unterschiede zwischen jedem Diagramm und seinem vorhergehenden Gegenstück dargestellt.

Wir haben insgesamt M BSs in der Umgebung aufgebaut, wobei M als sieben definiert ist und eine mobile BS umfasst. Der von den UGVs von BS m (m **ÿ 7) erhaltene Signalleistungsgewinn** wird als p definiert , **das** $\overset{M}{t}$. **Anschließend wird das Signal-Interferenz-Verhältnis (SIR) verwendet als primäre Kriterium zur Bewertung der Kommunikation der UGVs. Dieses Kriterium kann wie folgt ausgedrückt werden:**

$$\textbf{SIRt}\ \ddot{y}\ p\ \frac{Es}{\ddot{y}m\ddot{y}=Es\ \overset{M}{p}} \quad \textbf{(1)}$$

wobei lt **ÿ {1, · · · , M} die BSs darstellt, die bei Schritt t nicht mit den UGVs verbunden sind. Es ist erwähnenswert, dass wir der Einfachheit halber die Auswirkungen von Rauschen weggelassen haben, da bekannt ist, dass die Leistung der BS-UGV-Kommunikation häufig durch Interferenzen beeinträchtigt wird. Darüber hinaus haben wir mit der globalen Frequenzwiederverwendung den schlimmsten Fall berücksichtigt, in dem alle diese nicht verbundenen BSs zum Interferenzterm in Gleichung (1) beitragen.**

In unserer Studie war die Signalleistung, die die UGVs bei Schritt t empfingen, **hauptsächlich von ihrer relativen Position zu den BSs abhängig, und** pt **kann wie folgt geschrieben werden:**

$$pt = P^{-}\ddot{y}(qt)G(qt)\tilde{}ht \quad \textbf{(2)}$$

wobei P¯die Sendeleistung der BSs darstellt, während ÿ(qt) **die großflächige Kanalverstärkung darstellt; die großflächige Kanalverstärkung berücksichtigt die Auswirkungen von Pfadverlust und Schattenschwund. Sie kann wie folgt ausgedrückt werden:**

$$\ddot{y}(qt) = \textbf{ÿ0}\ \ \frac{}{\textbf{d0}\ d(qt)}^{\ddot{y}} \quad \textbf{(3)}$$

wobei ÿ0 der Pfadverlust bei der Referenzentfernung d0, d(qt) **die Entfernung zwischen UGV und BS und** *ÿ* **der Pfadverlustexponent ist. Und** G(qt) **bezeichnet den Antennengewinn der BS; der Antennengewinn der BS berücksichtigt den Richtungsgewinn des UGV relativ zur BS-Antenne. Er kann durch das Antennenstrahlungsmuster dargestellt werden, das typischerweise wie folgt ausgedrü**

$$G(qt) = \textbf{Gmax} \cdot A(\ddot{y}t\ ,\ \ddot{y}t) \quad \textbf{(4)}$$

wobei Gmax der maximale Antennengewinn und *A(ÿt ,* ÿt) **die Gewinnfunktion der Position des UGV relativ zur Hauptkeulenrichtung der Antenne ist. Diese Parameter hängen normalerweise vom Standort** qt **des UGV ab. Zusätzlich wird die Zufallsvariable** ˜ht **verwendet, um die Auswirkungen des Abklingens zu berücksichtigen .** Es ist wichtig zu beachten, dass jedes UGV bei jedem **Schritt** t **ein unabhängiges SIR hat , das verwendet wird, um die Kommunikationsleistung der UGVs zu diesem bestimmten Zeitpunkt zu bewerten. Es sollte auch beachtet werden, dass während der Initialisierung des Szenarios die Anfangspositionen aller Basisstationen, einschließlich der beweglichen Basisstation, festgelegt sind, d. h. sie befinden sich alle an einer festen Position, und dann werden die drei UGVs und die bewegliche Basisstation trainiert, durch die Strategie unterschiedliche Aktionen auszuführen, wobei zu diesem Zeitpunkt, basierend auf der ausgewählten Aktion, die nächste Position der beweglichen Basisstation durch die ausgewählte Aktion sowie die ursprüngliche Position zusammen bestimmt wird. Der Wert von** qt **ist an diesem Punkt festgelegt, da** qt **nur mit der Positionsvariable** (x, y) zusammenhängt. **Es ist ersichtlich, dass die Anfangsposition der mobilen BS voreingestellt ist, während das nachfolgende** qt **die Entscheidungsvariable ist und durch die Aktion der mobilen BS bestimmt wird, die darauf abzielt, den verbleibenden UGVs eine bessere Kommunikationsumgebung bereitzustellen .**

Machine Translated by Google

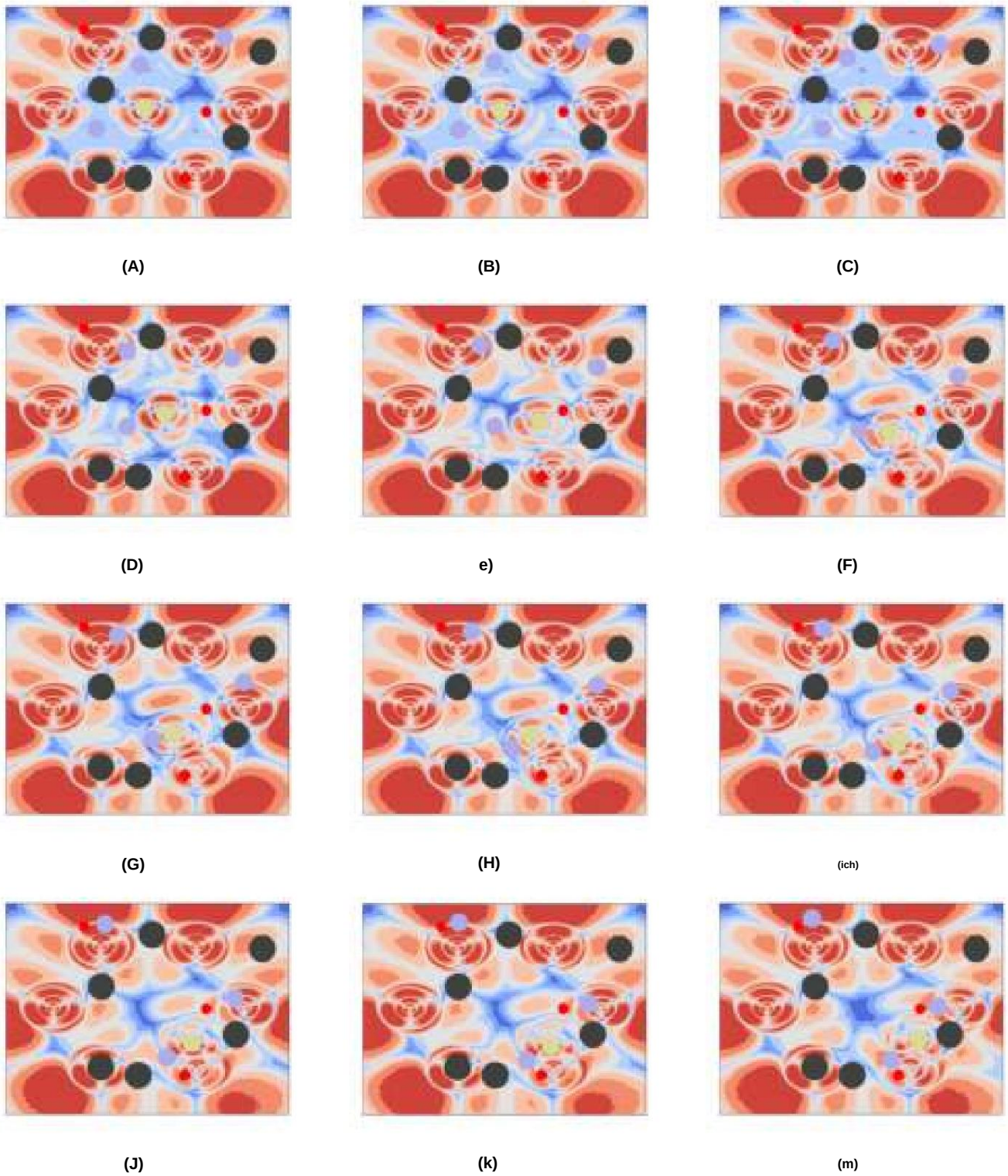


Abbildung 2. Die Auswirkungen der sich ändernden Kommunikation in der Umgebung, wenn sich die mobile BS bewegt. Rote Bereiche zeigen eine bessere Kommunikation an, während blaue Bereiche eine schlechtere Kommunikation anzeigen. Braune Kreise stellen die mobile BS dar, blaue Kreise stellen UGVs dar, schwarze Kreise stellen Hindernisse dar, und rote Kreise stellen Zielpunkte dar. Diese Bewegung zielt darauf ab, die Kommunikationsqualität für die UGV.

Machine Translated by Google

2.4. Zustand und Aktion der UGVs **Der Zustand der**

UGVs wird als s = **(s1,s2, . . . ,sN)** bezeichnet. **Für jedes UGV u wird der Zustand als** su = (sPu,sEu) definiert, **wobei** sPu = xu, yu, vxu, vyu, SIRu **eine Kombination aus Position** (xu, yu), **Geschwindigkeit** (vxu, vyu) **und** SIRu ist. **Zusätzlich** stellt sEu = xug, yug, **x0, y0,** vox, voy **die Daten dar, dass die UGVs andere UGVs oder Hindernisse beobachten. Der Begriff** su **stellt die Positionen des Agenten in einem Koordinatensystem dar. In vielen tatsächlichen Situationen ist es jedoch möglicherweise nicht möglich, absolute Standorte zu ermitteln. Daher können der Agent und die Barrieren in einem Polarkoordinatensystem für die Bewegung modelliert werden. In unserer ursprünglichen Formulierung soll** sEu = xug, yug, **x0, y0,** vox, voy die beobachteten Daten für jedes UGV **u darstellen. Zur Verdeutlichung:** (xug, yug) **stellt die Entfernung von der** g- **Einheit (einschließlich der UGVs und aller Hindernisse) zu UGV u dar. Und (x0, y0) stellt die globale Koordinatenposition von UGV u dar.**

Durch eine Reihe von Transformationsberechnungen können wir auch die globalen Positionen anderer Objekte ermitteln, die vom UGV u beobachtet werden. Die Kombination dieser Komponenten ermöglicht es jedem UGV, zu seinem Ziel zu navigieren und dabei die Anwesenheit und Bewegung von Hindernissen oder anderen UGVs zu berücksichtigen.

Die Aktion von UGVs wird als a = **(a1, a2, . . . ,** aN) bezeichnet **und ist als eine Sammlung einzelner Aktionen für jedes UGV in einem Multiagentensystem definiert. In diesem speziellen Dokument wird die Bewegung von UGVs vereinfacht, indem eine Anfangsgeschwindigkeit von 0 und eine konstante Beschleunigung angenommen werden, die durch eine Formulierung dargestellt wird:** vt = **v0** + at, **die als 2-dimensiona**

2.5. Belohnungsfunktion

Das Hauptziel besteht darin, die optimale kollaborative Strategie für einen bestimmten Zustand zu finden, um während Schritt t **und dem nächsten Schritt** t + **1 mit verbesserter Kommunikation kollaborativ zu navigieren. Bei Schritt t wird der bestimmte Zustand als** st bezeichnet . **Die Belohnung für die Durchführung einer Aktion** bei kann durch r(st , at) **dargestellt werden .** Folglich **kann die Gesamtbelohnung für die Umsetzung der Strategie** *ÿ* **wie folgt ausgedrückt werden:**

$$R(ÿ) = \mathrm{L}\, ÿÿ^{T} \; r(st , ÿ(st)) \mid \mathbf{s0} = s \qquad \textbf{(5)}$$

Unser Ziel ist es, die optimale Strategie zu bestimmen, bezeichnet als *ÿ* ÿ **, das maximiert die Gesamtbelohnung unter Einhaltung aller gegebenen Einschränkungen. Der Hauptschwerpunkt des Artikels liegt darauf,** die Richtlinie zu finden, die **unter allen möglichen Richtlinien** ÿ **die höchstmögliche Belohnung (bezeichnet als** *R(ÿ))* abwirft .

Es ist zu beachten, dass die Navigationsprinzipien für die mobile BS denen der UGVs ähneln . Sowohl die UGVs als auch die mobile BS können **basierend auf ihren Zielpositionen und aktuellen Umgebungsdaten ähnliche Prinzipien für die Pfadplanung und Hindernisvermeidung verwenden . Die Aufgaben von UGVs sind dreifach: Erstens erreichen UGVs ihr Ziel durch kollaborative Navigation. Zweitens sollten UGVs versuchen, Kollisionen zu vermeiden. Drittens sollten UGVs in einer Kommunikationsumgebung mit hoher Qualität unterwegs sein. Die mobile BS hat nur zwei Aufgaben: Eine besteht darin, mit den UGVs zusammenzuarbeiten, um die Kommunikationsabdeckung durch Anpassung der Position anzupassen und so sicherzustellen, dass sich die UGVs in einer Kommunikationsreichweite mit hoher Qualität bewegen. Eine besteht darin, Kollisionen so weit wie möglich zu vermeiden, ähnlich dem Ziel der UGVs. Es ist auch wichtig zu beachten, dass sich die Ausgangsposition der mobilen BS in allen von uns eingerichteten Szenarien genau im Zentrum der Szene befin In diesem Trainingsmodell besteht die Belohnungsfunktion der UGVs hauptsächlich aus drei Teilen, die auf einer theoretischen Grundlage basieren. Erstens hängt sie mit der Entfernung zwischen den UGVs und dem Zielpunkt zusammen. Zweitens hängt sie mit der Anzahl der Kollisionen zusammen, einschließlich Kollisionen zwischen UGVs, Kollisionen zwischen UGVs und der mobilen BS und Kollisionen zwischen UGVs und Hindernissen. Schließlich hängt sie mit dem von den UGVs in Schritt t erreichten SIR zusammen ,** der **als** r(st , at) formuliert **werden kann .**

$$r(st , at) = SIRt \; ÿ \; D(UGVs, Ziel) \; ÿ \; coll \qquad \textbf{(6)}$$

wobei SIRt **die umfassende Kommunikationsqualität darstellt, die von allen UGVs in jedem Schritt t erreicht wird. Die Definition von** SIRt **wurde ausführlich in Gleichung (1) eingeführt.**
D(UGVs, Ziel) **stellt die Summe der Längen zwischen allen UGVs und ihren jeweiligen**

Machine Translated by Google

Ziele bei jedem Schritt t; es ist zu beachten, dass kein UGV ein festes Ziel zu erreichen hat, was bedeutet, dass alle UGVs das zu erreichende Ziel basierend auf ihren Strategien und Beobachtungen autonom zuweisen . Der Begriff coll **stellt die Anzahl der Kollisionen dar, die bei jedem Schritt t zwischen allen UGVs aufgetreten sind.**

Die Berechnungsformel von SIRt **in Gleichung (6) wurde in Abschnitt 2.3 vorgestellt.**
Die Kommunikation wirkt sich direkt auf die Belohnungsfunktion der UGVs aus, denn eine bessere Kommunikation führt zu einer größeren Belohnung. Die UGVs werden daher dazu angeregt, Standorte mit besserer Kommunikation zu priorisieren, was sie dazu anregt, sich ausgiebig in diese Gebiete zu bewegen.

D(UGVs, Ziel) **wird wie folgt berechnet:**

$$D(UGVs, Ziel) = (xu \ddot{y} xZiel) \qquad ^{2} + (yu \ddot{y} yZiel)^{2} \qquad (7)$$

wobei (xu, yu) **die Koordinateninformationen aller Zielpunkte enthält, was darauf hinweist, dass der Verlust ebenfalls abnimmt, wenn die Entfernung zwischen den UGVs und dem Ziel kleiner wird.**
Ein geringerer Verlust entspricht folglich einer höheren Belohnung. Im Wesentlichen ist es wahrscheinlicher, dass die UGVs eine höhere Belohnung erhalten, wenn sie sich näher am Zielpunkt befinden.
Der Begriff coll **kann wie folgt bestätigt werden:**

$$kol = \begin{cases} \textbf{0, wenn } D(UGVs, Tupel) > K \\ \ddot{y}\textbf{1, wenn } D(UGVs, Tupel) \ddot{y} K \end{cases} \qquad (8)$$

wobei D(UGVs, Tupel) **die Distanzen zwischen den UGVs und verschiedenen Objekten wie anderen UGVs, der mobilen BS und Hindernissen im gegebenen Szenario darstellt. Zusätzlich wird eine Konstante „K“ verwendet, um die Möglichkeit einer Kollision einzuschätzen. Wenn die Distanz zwischen zwei beliebigen Objekten kleiner als der Wert von K ist, wird eine Kollision registriert. Folglich arbeiten bei diesem Ansatz mehrere Agenten zusammen, um das Auftreten von Kollisionen zu minimie**

3. RL Multi-Agent-Kommunikationsabdeckungsnavigation mit GA In diesem

Abschnitt beschreiben wir eine kurze Zusammenfassung der MDP-Formulierung für die Kommunikationsabdeckungsnavigation mit Zusammenarbeit zwischen der mobilen BS und den UGVs. Als nächstes stellen wir den DDPG-Algorithmus [31] vor, der für den kontinuierlichen Kontrollraum entwickelt wurde. Aufbauend auf diesen Grundlagen entwickeln wir einen innovativen RL-Algorithmus namens GA-MADPPG, um die Herausforderungen bei der Kommunikationsabdeckung und Navigation zu bewältigen. Der GA-MADPPG-Algorithmus besteht aus zwei Hauptkomponenten. Erstens übernehmen wir den MADPPG- Algorithmus, der DDPG nach dem CTDE-Paradigma erweitert. Dadurch können wir **die Vorteile von MADPPG bei der Handhabung von Multi-Agenten-Systemen und kontinuie Zweitens integrieren wir GA in den MADPPG-Algorithmus und ermöglichen so Echtzeit-Hyperparameter- Updates basierend auf der Verlustfunktion während des Trainingsprozesses. Die vorgeschlagene Richtlinie unterstreicht die Fähigkeit des GA-MADPPG-Algorithmus, Hyperparameter basierend auf der Verlustfunktion dynamisch anzupassen . Durch die Kombination dieser beiden Komponenten zielt GA-MADPPG darauf ab, eine effiziente Kommunikationsabdeckung und Navigation in kom**

3.1. MDP-Modell

Das Multi-Agenten-Markov-Spiel, eine bedeutende Erweiterung des MDP in einem Multi-Agenten- Szenario, ist Gegenstand von [32]. In diesem Spiel wird der theoretische Zustand von N Agenten **durch s dargestellt .** Zu jedem Zeitpunkt **t verfolgen die Agenten den aktuellen Zustand** st **und wählen eine Aktion** bei . **Anschließend wechselt der Zustand in den nächsten Zustand** st + **1, und alle Agenten erhalten eine Belohnung,** r(st , at).

Zur Auswertung von Aktions-Wert-Funktionen und Zustands-Aktions-Mapping-Wert-Funktionen, Die Berechnung der Wertfunktion für stochastische Strategien beinhaltet:

$$V\ddot{y}(st) \models E \qquad \sum_{l=0}^{\ddot{y}} \ddot{y}\ddot{y} \qquad r(st+l , at+l) \mid st \qquad (9)$$

Machine Translated by Google

wobei der Diskontierungsfaktor *ÿ* **ÿ [0,** 1] ist. **Und die Aktionswertfunktion wird wie folgt berechnet:**

$$Q\ddot{y}(st , at) = E \sum_{l=0}^{\ddot{y}} \ddot{y}^{\,l}\, \ddot{y}_{\,st\,,\,bei} \quad \textbf{(10)}$$

Erlernen eines idealen $\ddot{y}$ **Ziel ist eine Strategie, die den erwarteten Gesamtertrag optimiert** *ÿ* **aller Agenten.**

$$\ddot{y}^{\,\ddot{y}} = \textbf{arg max}_{\ddot{y}} \; E \sum_{t=0}^{\ddot{y}} \ddot{y}^{\,t}\, r(st , um) \quad \textbf{(11)}$$

3.2. Grundlagen des DDPG-Ansatzes

DDPG ist ein tiefer deterministischer Policy-Gradient-Algorithmus, der zur Lösung kontinuierlicher Aktionssteuerungsprobleme entwickelt wurde. Er basiert auf Policy-Gradienten und passt die Policy-Parameter *ÿ* direkt an , **um die Zielfunktion zu optimieren.**

$$J(\ddot{y}) = \textbf{Es}\ddot{y}\textbf{p}\ddot{y}\textbf{|,a}\ddot{y}\ddot{y}\ddot{y} \quad \textbf{(12)}$$

Dies ist die Kernidee von DDPG, da dabei bei jedem Schritt der Policy-Gradient *ÿÿ J(ÿ)* verwendet wird . **Der Policy-Gradient kann wie folgt ausgedrückt werden:**

$$\ddot{y}\ddot{y}\, J(\ddot{y}) = \textbf{Es}\ddot{y}\textbf{p}\ddot{y}\textbf{,a}\ddot{y}\ddot{y}\ddot{y}\,\textbf{[}\ddot{y}\ddot{y}\,\textbf{log}\,\ddot{y}\ddot{y}\,(a \,\textbf{|}\, s) Q\, \ddot{y}(s, a)] \quad \textbf{(13)}$$

wobei $Q\ddot{y}(s, a) = E\, R \,\textbf{|}\, s^{\,T}_{\,t = \textbf{s,}\,ein} = a$ **ist eine Aktionswertfunktion und** $p^{\,\ddot{y}}$ **ist der Staat -Verteilung.**

Deterministische Strategien können ebenfalls in das Strategiegradientenmodell integriert werden und werden als *µÿ :* **S ÿ A [1] bezeichnet. Insbesondere können wir unter bestimmten Umständen den Gradienten des Ziels** *J(ÿ)* = Es **ÿ** p ***µ*** [R(s, a)] **wie folgt schreiben:**

$$\ddot{y}\ddot{y}\, J(\ddot{y}) = \textbf{Es}\ddot{y}\textbf{D}\, \ddot{y}\ddot{y}\boldsymbol{\mu}\ddot{y}\, (a \,\textbf{|}\, \textbf{s)}\ddot{y}\textbf{aQ}\, \boldsymbol{\mu}\, (s, a)|_{\boldsymbol{a=\mu\ddot{y}}}\, (s) \quad \textbf{(14)}$$

Der Satz erfordert, dass der Aktionsraum a **stetig ist, da er von** ÿaQµ (s, a) abhängt.

Bei der DDPG-Methode, einer Variante des deterministischen Policy-Gradient-Algorithmus, werden tiefe neuronale Netzwerke verwendet, um das Policy- *µ* **und das kritische** *Qµ abzuschätzen.* **Es handelt sich um einen Off-Policy- Ansatz, d. h. er lernt während des Trainings aus Erfahrungen. Zusätzlich zum Online-Netzwerk verwendet DDPG auch ein Zielnetzwerk, um das Training zu stabilisieren. Das Zielnetzwerk wird regelmäßig überarbeitet, um die Auswirkungen von Policy-Schwankungen w**

3.3. Multi-Agenten-tief deterministischer Policy-Gradient

Die DDPG-Richtlinie demonstriert die inhärente Robustheit und Generalisierungsfähigkeiten des Agenten , was zu einer maximalen Leistung führt [31]. Dieser Vorteil macht DDPG besonders gut geeignet für das Lernen unter schwierigen Umständen, in denen Unbekanntes und externe Störungen vorhanden sind. Vor diesem Hintergrund haben wir ein Trainingsparadigma für die Navigation der Kommunikationsabdeckung basierend auf MADDPG übernommen. Der Agent in der Umgebung ist autonom und kann nicht mit anderen Agenten interagieren, ist jedoch wahrnehmbar. Bei jedem Schritt t kann der Agent die aktuellen Mobilitätsschemata anderer Agenten nicht beobachten. Der Vorteil von CTDE besteht darin, dass es die Notwendigkeit beseitigt, die Kompromisse zwischen Agenten zu berücksichtigen, und das Optimierungsziel besteht darin, den Gesamtertrag aller Agenten zu erhöhen [33].

$$G = \ddot{y}\textbf{s}\hat{}\textbf{, a,}\, p,r,\, \textbf{o, u}\ddot{y} \quad \textbf{(15)}$$

wobei u **den Index jedes Agenten darstellt und s^ die globalen Status und lokalen Beobachtungen jedes Agenten speichert . Der Term** a **ist eine Darstellung der Aktivität aller Agenten und die Belohnung jedes Agenten ist Teil des Tensors r. Die Beobachtungsfunktion wird durch o angegeben und** p stellt **die Wahrscheinlichkeit eines Übergangs vom aktuellen Zustand zum nächsten Zustand dar .**

Machine Translated by Google

Genauer gesagt hat das Spiel N **Agenten und Strategien, parametrisiert durch** *ÿ* = **{ÿ1, . . . , ÿN}. Der Term** *ÿ* = **{ÿ1, . . . ,** ÿN} **stellt die Gesamtheit aller Agentenstrategien dar. Für Agent i kann der Gradient der erwarteten Rendite, bezeichnet als** *J(ÿi)* = E[Ri], **daher wie folgt ausgedrückt werden:**

$$\ddot{y}\ddot{y}i\ J(\ddot{y}i) = E_{s\ddot{y}p\ \mu,ai\ddot{y}\ddot{y}i}\left[\ddot{y}\ddot{y}i \log \ddot{y}i(ai \mid oi)\ Q^{\ddot{y}}_{_}(x, a1, \ldots, aN)\right] \quad (16)$$

unserem Szenario werden die Gesamtaktionen a1, . . . , aN in **die** $_{_}$ **(x, a1, . . . ,** aN), **was ein In zentralisierte Aktionswertfunktion** ***Qÿ*** **eingespeist ,** die den Q-Wert für Agent i **zusammen mit einigen Zustandsdaten erzeugt. Im einfachsten Szenario könnten Zustände die Summe der von jedem Agenten gemachten Beobachtungen sein, (o1, . . . ,** oN), **aber wenn möglich, könnten wir auch zusätzliche Zustandsdaten Agenten dürfen beliebige Anreizsysteme haben, auch solche, die in einer feindlichen Umgebung konkurrierende Belohnungen bieten. In diesem Artikel legen wir die Belohnungsfunktion jedoch als Summe der Belohnungen für alle Agenten fest, da sich unsere Forschung auf Situationen konzentriert, in denen alle Agenten zusammenarbeiten, um ein gemeinsames Ziel zu erreichen, was zu kooperativen Belohnungen führt.**

Das erwähnte Konzept kann erweitert werden, um auf deterministische Strategien anzuwenden. Da wir nun N **kontinuierliche Strategien** *µÿi* **parametrisiert durch** ÿi haben , **können wir den Gradienten wie folgt ausdrücken:**

$$\ddot{y}\ddot{y}i\ \boldsymbol{J(\mu i)} = E_{x,a\ddot{y}D[\ddot{y}\ddot{y}i}\left[\ddot{y}aiQ^{\boldsymbol{\mu}}_{_}(x, a1, \ldots, aN)|ai=\mu i\ (oi)\ \right] \quad (17)$$

wobei die Übergänge (x, x) $^{\ddot{y}}$**, a1, . . . ,** aN,r1, . . . ,rN) **werden im Wiedergabepuffer D gespeichert, der alle Agentenerfahrungen sind.**

Damit Gleichung (17) **angewendet werden kann, müssen die Strategien anderer Agenten aktualisiert werden .** Die Kenntnis **der Beobachtungen und Strategien anderer Agenten ist keine besonders einschränkende Annahme, da** diese Informationen normalerweise allen Akteuren zur Verfügung stehen, wenn unser Ziel darin besteht, Agenten zu schulen, **in der Simulation ein anspruchsvolles Kommunikationsverhalten zu zeigen .**

3.4. Der genetische

Algorithmus **GA ist ein Computermodell, das von Darwins biologischer Evolutionstheorie inspiriert ist und zur Suche nach optimalen Lösungen durch Simulation der natürlichen Evolution verwendet wird. Es arbeitet direkt mit Strukturobjekten und vermeidet Differenzierungs- und Funktionskontinuitätsbeschränkungen [34–36]. Mit inhärenter impliziter Parallelität und starker globaler Optimierungsfähigkeit verwendet es probabilistische Optimierungsmethoden, um den Suchraum automatisch ohne strenge Regeln zu erhalten und zu leiten, was adaptive Anpassungen der Suchrichtung ermöglicht. GA zielt auf alle Individuen in einer Population ab und erkundet effizient einen codierten Parameterraum mithilfe von Randomisierungstechniken . Seine genetischen Operationen umfassen Selektion, Crossover und Mutation. Die Kernkomponenten eines GA sind Parametercodierung, anfängliche Populationseinstellung, Fitnessfunktionsdesign, genetisches Operationsdesign und Steuerparametereinstellung. Um die Funktionsweise eines GA zu demonstrieren, betrachten wir ein uneingeschränktes Optimierungsproblem. Das Ziel besteht darin, die folgende Funktion zu ma**

$$\textbf{Maximieren Sie } f(k),\ k^{m}_{N}\ \ddot{y}\ kn\ \ddot{y}\ k^{Sie}_{N,} \qquad n = 1, 2, \ldots, N. \quad (18)$$

Die Variable ki **kann Werte im Bereich von** k ln annehmen . **Obwohl wir ein Maximierungsproblem betrachten, kann ein GA auch für Minimierungsprobleme verwendet werden. Um die ordnungsgemäße Funktion des GA sicherzustellen, werden die folgenden Schritte ausgeführt.**

Die Variablen ki **in Gleichung (18) werden zunächst in spezifischen Zeichenfolgenstrukturen codiert, bevor GAs verwendet werden, um das oben genannte Problem zu lösen. Es ist wichtig zu erwähnen, dass die Codierung der Variablen in dieser Phase nicht immer erforderlich ist. In einigen Studien werden GAs direkt auf die Variablen angewendet, aber um die grundlegenden Ideen eines einfachen GAs zu diskutieren, werden wir diese Ausnahmen außer Acht lassen.**

Machine Translated by Google

Die Fitnessfunktion wird für jedes Individuum in der Ausgangspopulation und anschließend für jede neue Generation nach Anwendung der genetischen Operatoren Selektion, Crossover und Mutation ausgewertet. Da die Fitness jedes Individuums unabhängig von der der anderen ist, ist eine parallele Berechnung möglich.

Solche Übergänge können viele verschiedene Formen annehmen. Nachfolgend finden Sie zwei häufig verwendete Fitness-Mappings.

$$F(k) = \mathbf{1} \frac{1}{+ f(k)} \quad \textbf{(19)}$$

Diese Transformation wandelt ein Minimierungsproblem in ein äquivalentes Maximierungsproblem um, ohne die Position des Minimums zu ändern. Die Zielfunktion kann mithilfe einer anderen Funktion transformiert werden, um den Fitnesswert F(i) zu erhalten, **wie unten gezeigt:**

$$F(i) = V \ddot{y} \ddot{y} \frac{O(i)P}{P} \quad \textbf{(20)}$$

wobei V **ein großer Wert ist, um nicht negative Fitnesswerte sicherzustellen,** P **die Populationsgröße und** O(i) **der Zielfunktionswert des n-ten Individuums. Für diese Studie** wird V **als Maximalwert des zweiten Terms in Gleichung (20) gewählt, was zu einem Fitnesswert von Null führt, der dem Maximalwert der Zielfunktion entspricht. Diese Transformation ändert die Position der Lösung nicht; sie wandelt lediglich ein Minimierungsproblem in ein äquivalentes Maximierungsproblem um. Der Begriff „String-Fitness“ bezieht sich auf den Fitnessfunktionswert eines Strings.**

Auf die Population werden genetische Operatoren wie Selektion, Crossover und Mutation angewendet, wodurch eine neue Generation auf der Grundlage der fitteren Individuen der aktuellen Generation entsteht. Bei der Selektionsoperation werden Individuen mit Vorteilen in der aktuellen Population ausgewählt. Bei der Crossover- oder Rekombinationsoperation werden Nachkommen erzeugt, indem ein Teil der Chromosomen zwischen zwei ausgewählten Individuen ausgetauscht wird, wodurch zwei neue Chromosomen entstehen, die die Nachkommen darstellen. Bei der Mutationsoperation werden zufällig ein oder mehrere Chromosomenwerte (Gene) jedes neu geschaffenen Individuums geändert. Mutationen treten normalerweise mit einer sehr geringen Wahrscheinlichkeit auf.

3.5. GA-MADDPG zur Adressierung von Kommunikationsabdeckung und Navigation in einer eigenen abstrakten Formulierung

In der abstrakten Formulierung in Abschnitt 3.1 kann die Strategie der Zielfunktion als ÿ(st) = at(st) ausgedrückt werden. **In jeder Episode j besteht das Ziel darin, die Zielfunktion zu optimieren,** indem **für jeden Zustand** (s) **die beste Koordination und optimale Aktion** (a) ausgewählt wird . **Verschiedene Agenten werden beauftragt, selbst zu navigieren, um den Zielpunkt zu erreichen, und jeder Agent verfolgt eine unabhängige Strategie. Um Einschränkungen zu berücksichtigen und verschiedene Szenarien zu untersuchen, verwenden wir Off-Policy-Methoden anstelle von On-Policy-Methoden, da Off-Policy leistungsfähiger und allgemeiner ist. Es stellt sicher, dass die Daten umfassend sind und alle Aktionen abgedeckt sind. Sie können sogar aus einer Vielzahl von Quellen stammen – selbst generiert oder extern [37]. Abbildung 3 veranschaulicht die Highlights des vorgeschlagene**

Alle Kritikpunkte werden gleichzeitig aktualisiert, um die kombinierte Regressionsverlustfunktion für Episode j zu reduzieren:

$$L(\ddot{y}i) = \frac{1}{S} \ddot{y}_{J} \; j \ddot{y} Q y \mu \; X^{J}, {}^{J}_{aa\,1,\ldots,N}{}^{J} \quad \textbf{(21)}$$

Der Akteur wird mithilfe des abgetasteten Richtliniengradienten aktualisiert:

$$\mathbf{1}\, \ddot{y}\ddot{y}i \; J \ddot{y} \; \frac{}{S} \; \ddot{y}_{J} \ddot{y}\ddot{y}i \, \mu i \, (o \quad {}^{J}) \qquad \textbf{(22)}$$

$$\ddot{y}aiQ \quad \mu \; (x\, j\, , {}^{ja}1\, , \ldots , ai\, , \ldots , {}^{ja}N\,) | \, j\, ai = \mu i \, (o\,)$$

Machine Translated by Google

Und die zentralisierte Aktionswertfunktion Q μ **wird wie folgt aktualisiert:**

$$L(\ddot{y}i) = Ex,a,r,x\ \ddot{y}[(Q \quad \boldsymbol{\mu} \quad (\mathbf{x}, \mathbf{a1}, \ldots, aN)\ \ddot{\mathbf{y}}\ y) \quad ^{2}],$$
$$y = ri + \ddot{y}Q \quad \boldsymbol{\mu}^{\ddot{y}} \quad (\mathbf{X}^{\ddot{y}}, \quad {}^{\ddot{y}\ \ddot{y}\ aa\ \mathbf{1}, \ldots,}_{a\ j\ (oj\ )\ j\ =\mu^{\ddot{y}}} \tag{23}$$

Wo

$$\boldsymbol{\mu}^{\ddot{y}} = \mu\ddot{y}\ \ddot{\mathbf{y}}, \ldots, \mu\ddot{y}\ \ddot{\mathbf{y}}_{1 \quad N} \tag{24}$$

ist die Sammlung von Zielrichtlinien mit verschobenen Parametern ÿi .

Der Trainingsprozess des GA-MADPPG-Algorithmus ist in Algorithmus 1 zusammengefasst.

Um den Nutzen zu maximieren, verwenden wir DDPG-Training außerhalb der Richtlinien.

Algorithmus 1 GA-MADDPG-Algorithmus

Erforderlich: Eingangszustand s, Diskontierungsfaktor ÿ und Aktion a

Initialisierung: Initialisieren Sie die MPE-Umgebung mit vier Agenten (einschließlich 3 UGVs und 1 mobiler BS). Initialisieren Sie die Hyperparameterpopulation.

Ecount = **0**

für Episode = **1 bis max. Episode do**

Umgebungen zurücksetzen, erste Beobachtungen sammeln oi **von Agenten für** Schritt = **1 bis max.**

Schritt do At **für jeden Agenten**

wählen i **Agenten nehmen** At **und erhalten die nächsten** $\ddot{y}$ **Beobachtungen** o **Gesamtbelohnung in Gleichung (6)** berechnen

Speichern Sie alle Übergänge der Agenten in D und speichern Sie das L der Übergänge in D.

Ecount = Ecount + **1 wenn**

Ecount **ÿ Episode aktualisieren dann**

für g = **1 zu Kritik Updates Schritte tun**

Mustercharge B aus D j

Setze y $^{j}_{+\ \ddot{y}Q = r\ i}$ $\boldsymbol{\mu}^{\ddot{y}}$ $\mathbf{X}^{'J}$, $^{\ddot{y}\ \ddot{y}\ aa\ \mathbf{1}, \ldots, N}$ $^{'A}_{k = \mu\ k}$ $^{\ddot{y}}$ $^{J}_{OK}$

Minimieren Sie den Verlust in Gleichung (21), um den Kritiker zu aktualisieren. Aktualisieren Sie den Akteur mithilfe des abgetasteten Richtliniengradienten gemäß Gleichung (22).

Bewerten Sie die Fitness der Hyperparameterpopulation gemäß Gleichung (19).

Crossover-Hyperparameterpopulation Mutationsoperation

Neue Hyperparameterpopulation

gemäß D festlegen. Ende für Zielparameter aktualisieren: $\ddot{y}$ i

$^{\ddot{y}}$ **ÿ** $\ddot{y}\ddot{y}i$ + **(1 ÿ** ÿ)ÿ $^{\ddot{y}}$

Ecount = **0**

Ende

wenn Ende

für Ende für

Machine Translated by Google

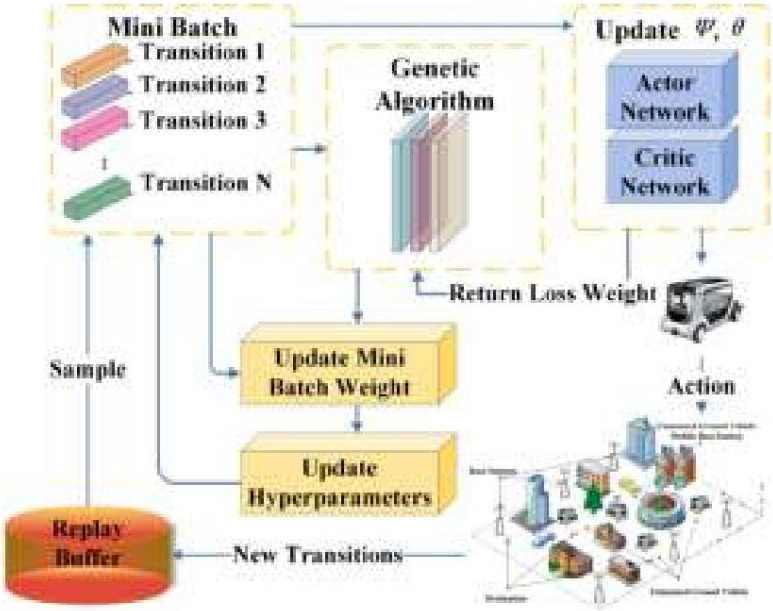


Abbildung 3. Detailliertes Diagramm des GA-MADDPG-Algorithmus.

4. Simulationsergebnisse

In diesem Abschnitt präsentieren wir anschauliche Beispiele, um den Versuchsaufbau dieses Papiers zu veranschaulichen. Basierend auf diesen Beispielen schlagen wir mehrere Metriken vor, um die Wirksamkeit des Algorithmus zu bewerten, und führen eine quantitative Analyse durch, um die Vorteile unseres dargestellten Modellierungsansatzes und unserer Strategie zu verdeutlichen. Anschließend präsentieren wir Ergebnisse numerischer Simulationen, um die Wirksamkeit und Effizienz der Algorithmen zu demonstrieren. Darüber hinaus liefern wir aufschlussreiche Kommentare zu den Ergebnissen.

4.1. Einstellungen der Experimente

In diesem Unterabschnitt stellen wir die genauen experimentellen Koeffizienteneinstellungen vor. Das simulierte Gebiet ist ein dichtes Stadtgebiet von 2 × 2 km² mit sieben zellularen BS-Standorten. Abbildung 4 zeigt eine Draufsicht des Kanalmodells in diesem Dokument, in dem sieben Bodenbasisstationen durch blaue fünfzackige Sterne dargestellt werden und der blaue fünfzackige Stern in der Mitte die bewegliche Basisstation darstellt. Jede Basisstation hat drei Einheitengruppen. Da es insgesamt sieben Basisstationen gibt, beträgt die Anzahl der Einheiten 21. Die Sendeleistung der Einheitszelle wird auf Pm **= 20 dBm eingestellt, der Kommunikationsunterbrechungsschwellenwert auf** ÿth **= 0 dB und die Rauschleistung wird als** $\ddot{y} = \ddot{y}65\ dBm^2$ **definiert. In diesem Dokument wird das in der 3GPP-Spezifikation geforderte Basisstationsantennenmodell übernommen . Der Einfachheit halber nehmen wir an, dass die Betriebshöhe der UGVs auf 0 m eingestellt ist Die spezifischen Werte der in der simulierten Umgebung beteiligten Parameter lauten wie folgt: Die Anzahl der UGVs ist auf vier festgelegt (einschließlich einer beweglichen BS), die Anzahl der Hindernisse ist in den Hauptbereichen auf fünf festgelegt und es gibt drei Zielpunkte. Die Positionen dieser Elemente werden bei jedem Auftreten zufällig bestimmt. Da wir einen dynamischen Aktualisierungsmechanismus für Hyperparameter verwenden, listen wir die gemeinsamen Parameter des Basisalgorithmus und des GA-MADDPG-Algorithmus in Tabelle 1 auf und führen außerdem die anfängliche Hyperparameterpopulation des GA-MADDPG-Algorithmus in Tabelle 2 auf.**

In dieser Studie ist es wichtig zu beachten, dass die Kommunikationsumgebung ausschließlich durch die Positionierung jedes UGV bestimmt wird. Die Qualität der Kommunikation zwischen mehreren UGVs beeinflusst ihre kollaborative Navigation nicht. Dies liegt daran, dass der kollaborative Navigationsprozess ausschließlich auf einem Multi-Agenten-Algorithmus basiert, um die UGVs bei der Erkundung der Umgebung zu koordinieren.

Machine Translated by Google

Tabelle 1. GA-MADDPG-Parametereinstellungen.

Definition	Wert	Definition	Wert
Max. Anzahl Folgen	60.000	Minibatch-Größe	512
Wiedergabepuffer Kapazität	1.000.000	Diskontierungsfaktor	0,99
Schritte pro Update	100	Lernrate	0,0001
Max. Schritte pro Folge	25	Bevölkerungsrate aktualisieren	100
Zeitschrittlänge	1	Versteckte Dimension	64

Tabelle 2. Anfängliche Hyperparameterpopulation des GA-MADDPG-Algorithmus.

Diskontierungsfaktor	Lernrate	Wiedergabepufferkapazität	Minibatchgröße
0,9	0,01	10.000	512
0,95	0,001	100.000	1024
0,99	0,0005	1.000.000	2048

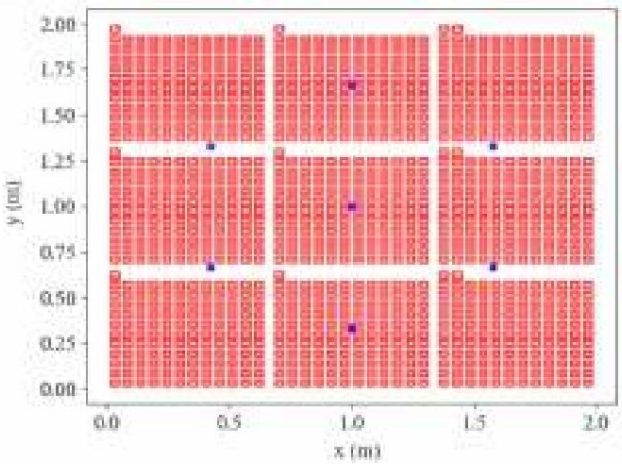


Abbildung 4. Draufsicht der Verteilung der Basisstationsmodelle.

4.2. Bewertungsindikatoren für die UGV-Navigation

Um die Navigationssicherheit, Effektivität, Robustheit und Kommunikationsanbindung von UGVs objektiv zu messen, haben wir spezielle Bewertungsi Wir haben auch den sich ändernden Status der Bewertungsmetriken aufgezeichnet, wie in Abbildung 5.

- **Kommunikationsrendite.Die Kommunikationsrendite ist die durchschnittliche Qualität pro Episode für die UGVs und wird basierend auf Gleichung (1) berechnet. Die Kommunikationsrenditen konvergieren schnell von anfänglich ÿ800 auf** ÿ300, wie in Abbildung **5a darg Dies zeigt an, dass sich die Kommunikationsqualität verbessert und stabilisiert hat in einem Intervall.**
- **Kollisionszeiten: Die Kollisionszeiten sind die Summe der Kollisionen zwischen UGVs und Hindernissen und zwischen Drohnen und Drohnen in einer durchschnittlichen Runde. Der Kollisionsindikator konvergiert von 540 auf unter 480, wie in Abbildung 5b dargestellt, was darauf hinweist, dass die Die Anzahl der Kollisionen wurde ebenfalls etwas reduziert, und da diese Studie es ermöglicht, UGVs eine bestimmte Anzahl von Kollisionen haben, ist der Kollisionsindikator nicht die Haupt Optimierungsziel.**
- **Außerhalb der Zeiten: Die Außerhalb der Zeiten geben die Anzahl der Male an, die die UGVs die Grenzen verlassen. und aus der Umgebung, die wir eingestellt haben, herauslaufen. Aus Abbildung 5c geht die schnelle Reduzierung der Die Anzahl der Überschreitungen der Grenzen zeigt, dass unsere Forschung deutlich begrenzte ineffektive Grenzverletzungen, was zeigt, dass unsere Studie effektiv ist innerhalb des ausgewiesenen Gebiets tätig.**

Machine Translated by Google

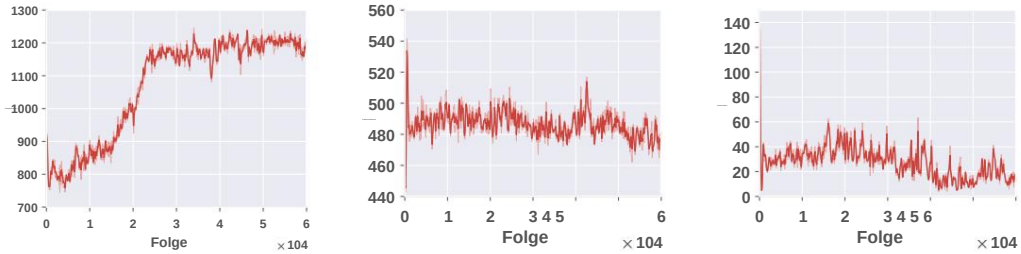


Abbildung 5. Drei Bewertungsindikatoren für die UGV-Navigation. (a) Kommunikationsrückmeldung. (b) Kollision Nummer. (c) Außerhalb der Zeiten.

4.3. Vergleichende GA-MADDPG-Experimente

Um mit dem vorgeschlagenen Algorithmus zu vergleichen und festzustellen, ob der Algorithmus besser funktioniert, bieten wir sieben RL-Ansätze an, die als Basislinien gedacht sind. Die Methoden sind MADDPG [24]: ein klassischer multi-agent deep deterministic policy gradient, R-MADDPG [38]: ein tief rekurrierender Multi-Agenten-Akteur-Kritiker, MAPPO [39]: Multi-Agenten proximale Policy-Optimierung, RMAPPO [39]: eine tief rekurrierende Multi-Agenten-proximale Policy Optimierung, MQMIX [40]: Mellow–Max-Faktorisierung monotoner Werte für tiefe Multi-Agent, MASAC [41]: ein klassischer Multi-Agent Soft Actor–Critic, MAD3PG [42]: ein Multi-Agent Deep Distributional Deterministic Policy Gradient, MATD3 [43]: der Zwilling mit verzögerter tiefen deterministischen Politikgradienten, und RMATD3 [44]: der Zwilling verzögerten tiefen deterministischen Politikgradient mit einer tiefen Rekurrenz. Insbesondere replizieren wir diese Baselines mit den gleichen Simulationsumgebung, um zu gewährleisten, dass das Experiment fair ist.

Die kumulierte Rendite des GA-MADDPG und anderer Algorithmen, die angezeigt wird in Abbildung 6 zeigt die experimentellen Vergleichsergebnisse und unterstreicht die Wirksamkeit von GA-MADDPG-Algorithmen. GA-MADDPG übertrifft die anderen Algorithmen durch eine deutlich höhere Belohnungsrendite von etwa ÿ1200 bei 60.000 Episoden und erreicht damit Konvergenzpunkt. Darüber hinaus, wie in Abbildung 6 gezeigt, sowohl MADDPG als auch R-MADDPG erreichen geringere Belohnungen von etwa ÿ1600 im Vergleich zu GA-MADDPG, was eine starke Beweis für die Wirksamkeit unseres Beitrags: die Verwendung von GA adaptiven Hyperparametern ermöglicht bessere Sprünge aus den lokalen Optima und höhere Belohnungen. Wie in Abbildung 6 dargestellt, in der spezifischen Umgebung, die wir konfiguriert haben, weder der ursprüngliche MADDPG-Algorithmus noch seine Variante mit tiefen rekurrierenden Netzwerken übertrifft GA-MADDPG in den Bereichen Konvergenzgeschwindigkeit und endgültige Konvergenzergebnisse: GA-MADDPG konvergiert in etwa 2000 Episoden, während R-MADDPG in etwa 5000 Episoden konvergiert, und der ursprüngliche Algorithmus MADDPG konvergiert sogar noch schlechter. Von größerer Bedeutung sind unsere experimentellen zeigen, dass MASAC, MAPPO, MAD3PG, MQMIX und RMAPPO auf Herausforderungen stoßen in Erreichen eines wünschenswerten Konvergenzzustands innerhalb der kooperativen Multi-Agenten-Umgebung MASAC benötigte etwa 25.000 Episoden, um zu konvergieren, Stabilisierung bei einem Belohnungswert von ca. ÿ1800. MAPPO und RMAPPO zeigten weniger stabile Konvergenz, mit Belohnungen, die zwischen ÿ2000 und ÿ2500 schwanken. Die Belohnung von MAD3PG konvergierte auf ungefähr ÿ2100. Was MQMIX betrifft, so ist die Belohnung zeigte anfängliche Schwankungen während der ersten 25.000 Episoden, gefolgt von einem stetigen Rückgang danach. Dies unterstreicht die Überlegenheit von GA-MADDPG in Bezug auf Leistung und Wirksamkeit.

Darüber hinaus tendieren bestimmte Algorithmen dazu, zu lokalen Optima zu konvergieren, was die Wirksamkeit unserer Entscheidung, den MADDPG-Algorithmus zu übernehmen und zu verbessern, weiter ver es. Wie in Abbildung 6 dargestellt, kann GA-MADDPG in den ersten 25.000 Episoden lokale Optimalität. Die Einbeziehung des GA-Mechanismus ermöglicht jedoch GA-MADDPG um höhere Belohnungen jenseits dieser Schwelle zu erhalten. Insbesondere MAPPO und MQMIX zeigen eine unterdurchschnittliche Leistung, möglicherweise aufgrund des Mangels an adaptiven Hyperparamete behindert ihre effektive Zusammenarbeit innerhalb der Multi-Agenten-Umgebung und führt

Machine Translated by Google

auf Konvergenzherausforderungen. Daher zeigt diese Beobachtung natürlich die hohe Wirksamkeit der Einbindung von GA in Multi-Agent-RL-Algorithmen. Durch die Einführung von GA können Multi-Agent-Algorithmen das Fallen in lokale Optima effektiver vermeiden, was zu einer verbesserten Konvergenzgeschwindigkeit und besseren Ergebnissen führt. Und die Variation des durch Gleichung (21) berechneten Verlusts **wird durch Abbildung 7 dargestellt , aus der wir die konstante Konvergenz des Verlusts auf nahezu 1800 erkennen können , was die Konvergenz des Algorithmus beweisen kann. Während des Validierungsprozesses zeigt Abbildung 8 mehrere simulierte Pfade von UGVs. Unter optimalen Kommunikationsbedingungen könnte das BS-UGV stationär bleiben, um potenzielle Verluste durch Kollisionen zu vermeiden. In Situationen mit weniger als hervorragender Kommunikation bewegt sich das BS-UGV jedoch proaktiv, um Kommunikationseinschränkungen auszugleichen. Darüber hinaus zeigen Statistiken für die drei Bewertungsindikatoren (Abbildung 5) die Verbesserung des Kommunikationsertrags, die Verringerung der Anzahl von Kollisionen und die Verringerung der Anzahl von Außeneinwirkungen, wenn der Algorithmus konvergiert. Der Return on Communications verbesserte sich von einem Anfangswert von ÿ800 auf ÿ300 gegen Ende des Experiments. Gleichzeitig sank die Kollisionshäufigkeit von 540 auf 470, und die Anzahl externer Ereignisse verringerte sich von 100 auf nahezu null. Dies deutet darauf hin, dass mit der Konv**

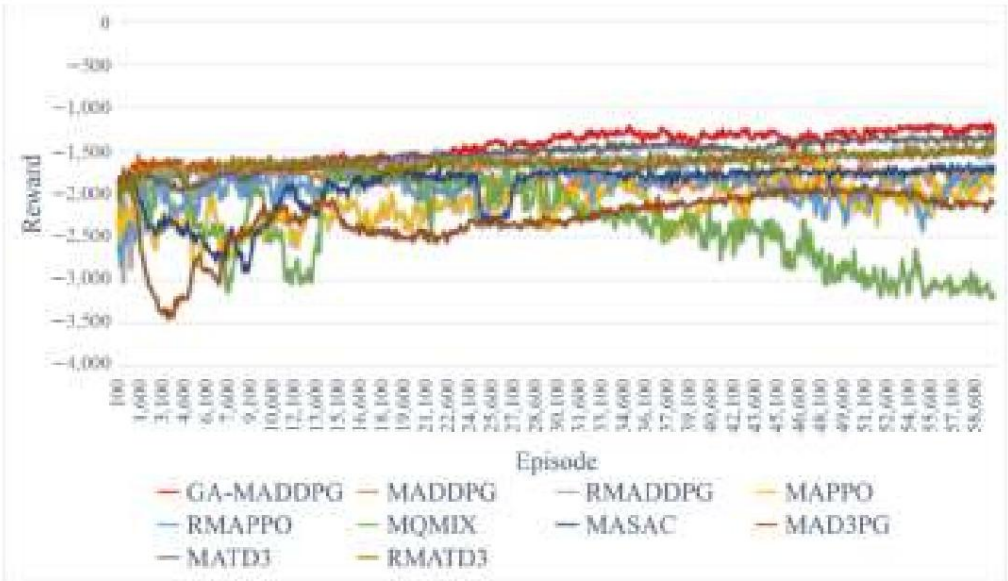


Abbildung 6. Durchschnittliche Kosten des GA-MADDPG und anderer fortgeschrittener Algorithmen.

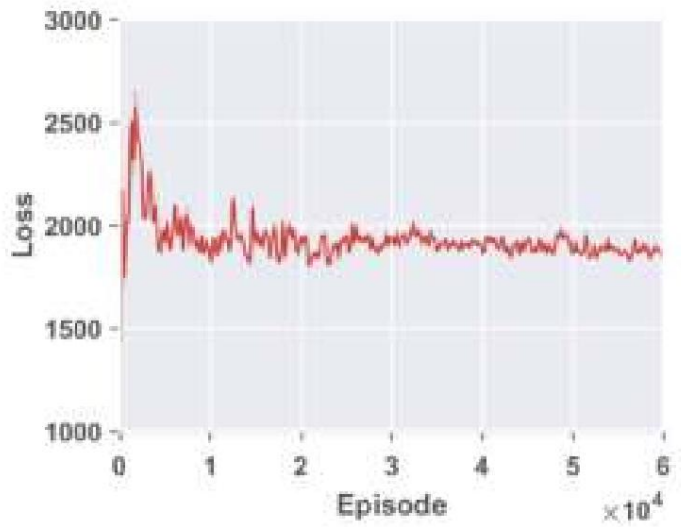
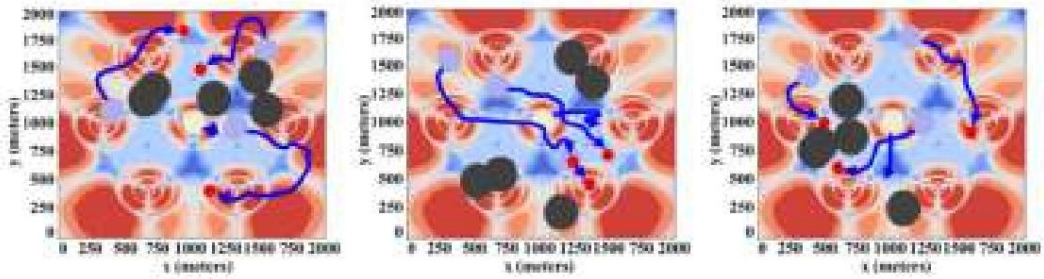


Abbildung 7. Entwicklung der Verlustfunktion.

Machine Translated by Google



(A) **(B)** **(C)**

Abbildung 8. Einige UGV-Pfadkarten basierend auf GA-MADDPG.

4.4. Generalisierungsexperiment von GA-MADDPG **4.4.1.**
Simulation mit unterschiedlichen UGV-Anzahlen

Um die Universalität des vorgeschlagenen GA-MADDPG-Algorithmus in der festgelegten Umgebung weiter zu beweisen, wurden im Rahmen dieser Studie auch zwei weitere Generalisierungsexperimente für die Szene ent Im Experiment wurden unterschiedliche Zahlen von UGVs, Zielpunkten und Hindernissen in der Szene festgelegt, um zu ermitteln, ob der Algorithmus GA-MADDPG weiterhin eine überlegene Leistung erbringen kann. Es ist zu beachten, dass einige Basisalgorithmen in Abschnitt 4.3 schlecht funktionierten oder sogar Schwierigkeiten beim Konvergieren hatten. Daher werden im Generalisierungsexperiment vier Basisalgorithmen verwendet, die in Abschnitt 4.3 relativ stabil sind, darunter MASAC, MAD3PG, MADDPG und seine Variante RMADDPG. Generalisierungsumgebung 1: Die Anzahl der UGVs erhöht sich auf vier, die Anzahl der mobilen Basisstationen beträgt eins, die Anzahl der Zielpunkte erhöht sich auf vier und die Anzahl der Hindernisse erhöht sich auf sieben. Die Bedeutung dieser Umgebungsgestaltung besteht darin, den Schweregrad der Umgebung durch Erhöhung der Anzahl der UGVs und der Anzahl der Hindernisse zu erhöhen.

Aus Abbildung 9 können wir ersehen, dass der GA-MADDPG-Algorithmus trotz der zunehmenden Komplexität der Umgebung in rauen Umgebungen immer einen höheren Konvergenzwert hat und gut auf einen hohen Wert konvergieren kann. Der GA-MADDPG-Algorithmus kann die Konvergenz bis zu einem Belohnungswert von ÿ3000 aufrechterhalten, während die anderen Basisalgorithmen in komplexen Umgebungen keine gute Leistung erbringen oder sogar Schwierigkeiten haben, zu konvergieren, und der höchste Belohnungswert nur bei etwa ÿ3300 liegt . Dies zeigt deutlich, dass der GA-MADDPG-Algorithmus auch **nach zunehmender Umgebungskomplexität noch eine bessere Leistung als andere Algorithmen aufweist .**

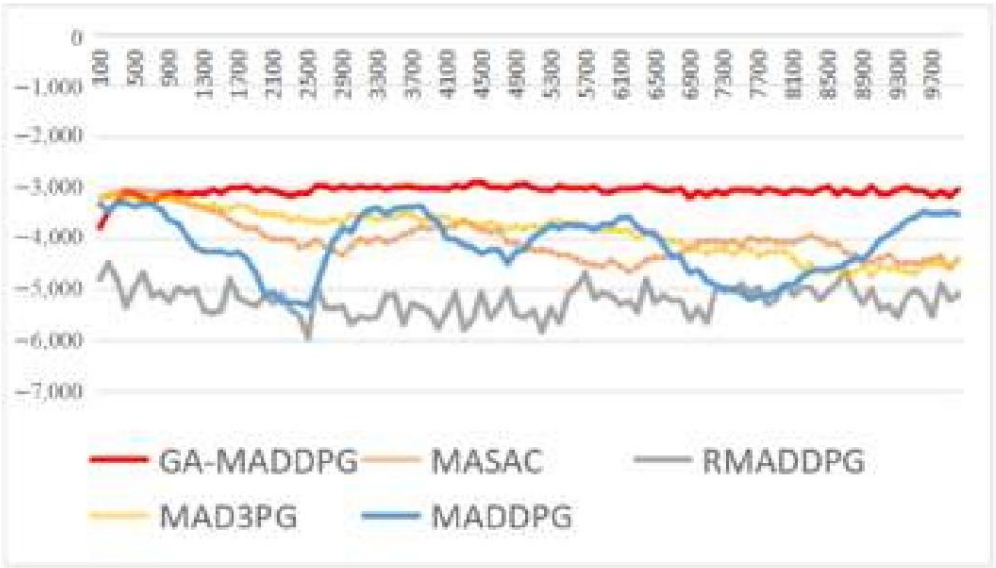


Abbildung 9. Vergleich der Durchschnittskosten der Generalisierungsumgebung 1.

Machine Translated by Google

Allgemeine Umgebung 2: Die Anzahl der UGVs wird auf zwei reduziert, die Anzahl der mobilen Basisstationen beträgt eins, die Anzahl der Zielpunkte wird auf zwei reduziert und die Anzahl der Hindernisse wird auf drei reduziert. Die Bedeutung dieser Umgebungsgestaltung besteht darin, die Einfachheit der Umgebung zu verbessern, indem die Anzahl der UGVs und Hindernisse vereinfacht wird, sodass das UGV das Ziel mit einer größeren Belohnung erreichen

Wie aus Abbildung 10 **ersichtlich , zeigen die Belohnungen der meisten Algorithmen einen guten Aufwärtstrend . Dies liegt daran, dass die verallgemeinerte Umgebung einfachere drei UGVs (einschließlich einer UGV-Basisstation), drei Hindernisse und zwei Zielpunkte verwendet. Der Algorithmus funktioniert in einer einfachen Umgebung besser und die Konvergenz ist einfacher als in der verallgemeinerten Umgebung. Mit abnehmender Anzahl der Fahrzeuge nimmt auch die Anzahl der Kollisionen und Überschreitungen der Grenzen entsprechend ab. Es ist zu beachten, dass der Belohnungswert des Gesamtalgorithmus positiv ist, da die Parameter der Kommunikationsumgebung unverändert bleiben, was normal i Aus Abbildung 10 ist ersichtlich, dass in dieser verallgemeinerten Umgebung die Belohnung des GA-MADDPG-Algorithmus immer die Nase vorn hat, sowohl was die Konvergenzgeschwindigkeit als auch den endgültigen Konvergenzwert betrifft, die viel höher sind als bei den anderen Algorithmen, und dass der endgültige Belohnungswert auf etwa 200 konvergieren kann. Als Basisalgorithmus weist MADDPG außerdem einen höheren Konvergenzwert von etwa 150 auf. Dies zeigt deutlich, dass der GA-MADDPG-Algorithmus auch in einer einfachen Umgebung eine gute Leistung erbringen kann.**

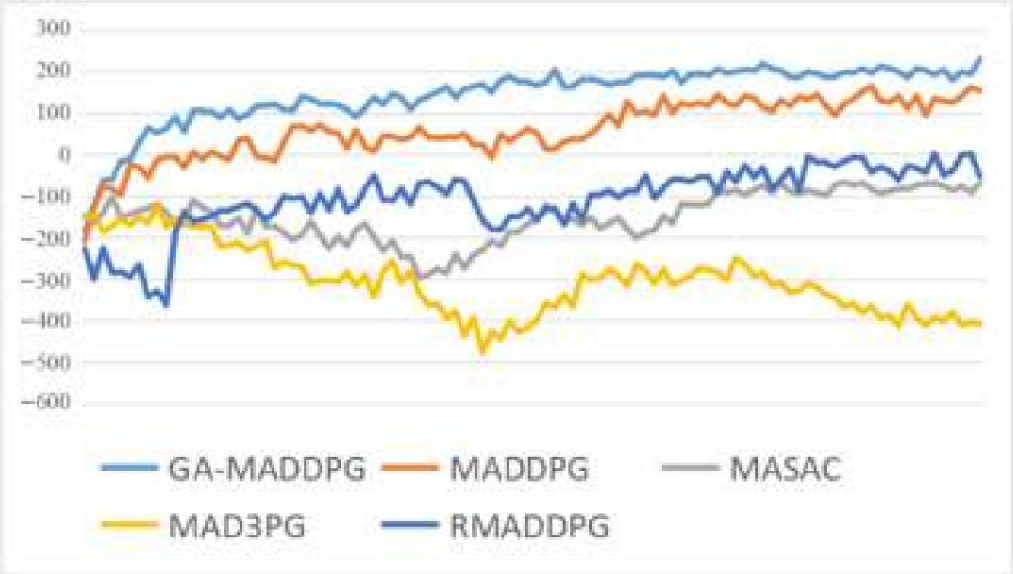


Abbildung 10. Vergleich der Durchschnittskosten der Generalisierungsumgebung 2.

Aus den Abbildungen 9 **und** 10 **ist ersichtlich,** dass **der GA-MADDPG-Algorithmus in den Versuchsumgebungen mit zwei unterschiedlichen Parametereinstellungen trotz Änderungen der Anzahl der UGVs, der Anzahl der Zielpunkte** und der Anzahl der Hindernisse immer noch eine bessere Leistung **als die anderen Algorithmen erbringt, was die Robustheit des GA-MADDPG- Algorithmus und seine Universalität für Umgebungsszenarien deutlich beweist.**

4.4.2. Experimente zur Wirksamkeit der mobilen BS

Die vorherigen Unterabschnitte beweisen die Stabilität und Konvergenz unseres vorgeschlagenen Algorithmus . Außerdem beweist der letzte Abschnitt, dass unser vorgeschlagener Algorithmus im gleichen Szenario überlegen ist. Um die Wirksamkeit der in diesem Artikel vorgeschlagenen mobilen Basisstation besser zu demonstrieren, fügen wir ein zusätzliches Experiment hinzu: Wir ersetzen nur die mobile Basisstation durch eine feste Basisstation, verwenden aber denselben Algorithmus.

Wir verwenden die Kommunikationsrendite als Bewertungsmaßstab. Die Kommunikationsrendite mit einer mobilen Basisstation ist von Beginn des Trainings an besser als die der festen Basisstation, wie in Abbildung 11 dargestellt. **Die Kommunikationsrendite eines einzelnen UGV**

Machine Translated by Google

kann schließlich auf etwa 300 konvergieren, während die der festen Basisstation bei etwa 200 Fuß liegt, was die Wirksamkeit unserer vorgeschlagenen mobilen Basisstation eindeutig beweist.

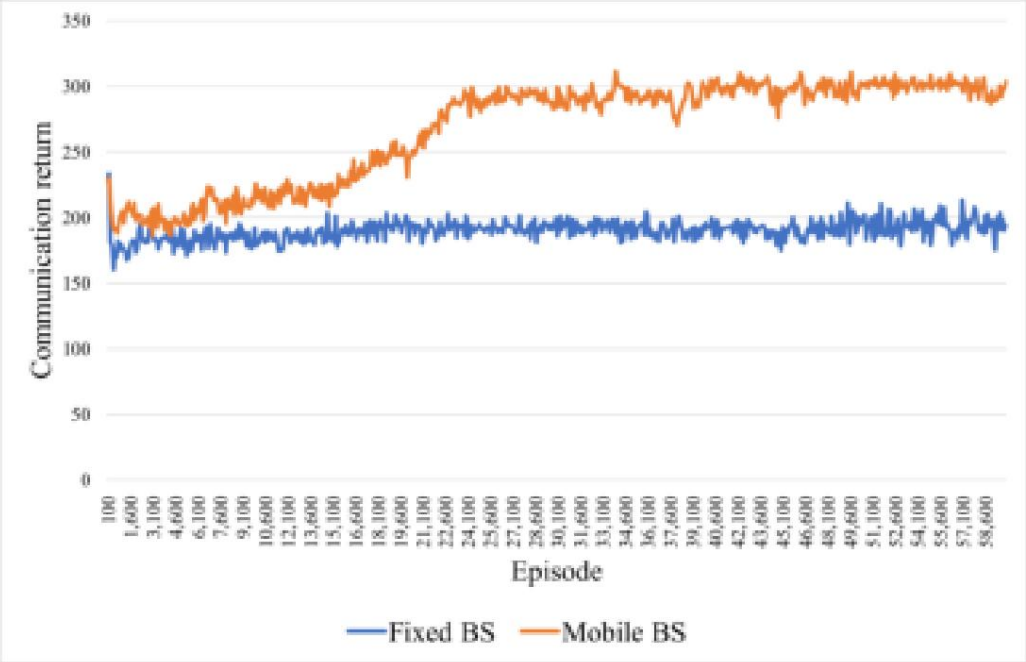


Abbildung 11. Vergleich der Kommunikationsrenditen zwischen mobiler und fester BS.

5. Schlussfolgerungen

In diesem Artikel wird ein kooperatives System für die kooperative Navigation mehrerer UGVs innerhalb eines Kommunikationsabdeckungsbereichs vorgeschlagen. Das System ist als MDP formuliert, um eine optimale Navigationsrichtlinie für die UGVs zu bestimmen, mit dem Ziel, die Gesamtbelohnung zu maximieren. Im Gegensatz zu früheren Studien, die sich auf die Navigation mit fester Abdeckung konzentrierten, führt dieser Artikel einen neuartigen Ansatz ein, indem er eine mobile BS in den Multi-Intelligent-Body-Algorithmus einbezieht . Diese Innovation zielt darauf ab, die Kommunikationsabdeckung zu verbessern und den für intelligente Agenten verfügbaren Lösungsraum zu erweitern. Um das Risiko lokaler Optima zu verringern, führt diese Studie einen adaptiven GA-Hyperparameter-Aktualisierungsmechanismus ein, um das Multi-UGV-Navigationsproblem zu lösen. Wir prägen den Begriff GA-MADDPG, um diesen neuartigen RL-Algorithmus zu bezeichnen. Die Simulationsergebnisse zeigen, dass GA-MADDPG im Vergleich zu anderer

In unserer zukünftigen Forschung möchten wir die folgenden Punkte ansprechen: (1) Um den Modellrealismus zu verbessern, kann man eine traditionelle PID-Steuerung mit Multi-Agent-RL kombinieren und die Navigationsrichtlinie weiter optimieren, indem man die Kontrolle über den Maschinenbetrieb übernimmt. (2) Man kann versuchen, eine neue Architektur zum Erlernen von Richtlinien zu verwenden, beispielsweise durch Verwendung von LSTM (Long Short-Term Memory) und der Transformer-Architektur. LSTM kann das Problem des Gradientenverschwindens und der Gradientenexplosion während des Trainings langer Sequenzen lösen; der Vorteil der Transformer- Architektur besteht darin, dass ihre Aufmerksamkeitsschicht eine Aktionssequenz sehr gut erlernen kann.

Beiträge der Autoren: Forschungsdesign, XL und MH; Datenerfassung, XL; Schreiben – Vorbereitung des Originalentwurfs, XL; Schreiben – Überprüfung und Bearbeitung, XL und MH; Betreuung, MH; Mittelbeschaffung , MH. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Arbeit wurde durch das Sonderprojekt der Academician Innovation Platform der Provinz Hainan (YSPTZX202209) unterstützt.

Datenverfügbarkeitserklärung: Die in der Studie vorgestellten Originalbeiträge sind im Artikel enthalten, weitere Anfragen können an den entsprechenden Autor gerichtet werden.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Machine Translated by Google

Verweise

1. **Afzali, SR; Shoaran, M.; Karimian, G. Ein modifizierter Konvergenz-DDPG-Algorithmus für Robotermanipulation.** Neuronaler Prozess. Lett. **2023,** 55, **11637–11652. [CrossRef]**
2. **Chai, R.; Niu, H.; Carrasco, J.; Arvin, F.; Yin, H.; Lennox, B. Entwurf und experimentelle Validierung einer schnellen Trajektorienplanung und -steuerung auf Basis von Deep Reinforcement Learning für mobile Roboter in unbekannter Umgebung.** IEEE Trans. Neural Netw. Learn. Syst. **2022,** 35, **5778–5792. [CrossRef]**
3. **Dong, X.; Wang, Q.; Yu, J.; Lü, J.; Ren, Z. Neuroadaptives Output-Formation-Tracking für heterogene nichtlineare Multiagentensysteme mit mehreren nicht identischen Leadern.** IEEE Trans. Neural Netw. Learn. Syst. **2022,** 35, **3702–3712. [CrossRef]**
4. **Wang, Y.; Zhao, C.; Liang, J.; Wen, M.; Yue, Y.; Wang, D. Integrierte Lokalisierung und Planung für die Geschwindigkeitsregelung von UGV-Platoons in Infrastrukturfreien Umgebungen.** IEEE Trans. Intell. Transp. Syst. **2023,** 24, **10804–10817. [CrossRef]**
5. **Tran, VP; Perera, A.; Garratt, MA; Kasmarik, K.; Anavatti, SG Abdeckungspfadplanung mit Budgetbeschränkungen für mehrere unbemannte Bodenfahrzeuge.** IEEE Trans. Intell. Transp. Syst. **2023,** 24, **12506–12522. [CrossRef]**
6. **Wu, Y.; Li, Y.; Li, W.; Li, H.; Lu, R. Robustes Lidar-basiertes Lokalisierungsschema für unbemannte Bodenfahrzeuge mittels Multisensor Fusion.** IEEE Trans. Neural Netw. Learn. Syst. **2021,** 32, **5633–5643. [CrossRef]**
7. **Zhang, W.; Zuo, Z.; Wang, Y. Vernetzte Multiagentensysteme: Antagonistische Interaktion, Einschränkung und ihre Anwendung.** IEEE Trans. Neural Netw. Lernen. Syst. **2021,** 33, **3690–3699. [CrossRef] [PubMed]**
8. **Chen, D.; Weng, J.; Huang, F.; Zhou, J.; Mao, Y.; Liu, X. Heuristischer Monte-Carlo-Algorithmus für unbemannte Bodenfahrzeuge in Echtzeit Lokalisierung und Kartierung.** IEEE Trans. Veh. Technol. **2020,** 69, **10642–10655. [CrossRef]**
9. **Unlu, HU; Patel, N.; Krishnamurthy, P.; Khorrami, F. Sliding-window temporal attention based deep learning system for robust Sensormodalitätsfusion für UGV-Navigation.** IEEE Robot. Autom. Lett. **2019, 4, 4216–4223. [CrossRef]**
10. **Lyu, X.; Hu, B.; Wang, Z.; Gao, D.; Li, K.; Chang, L. Ein integrierter fehlertoleranter Navigationsmechanismus für SINS/GNSS/VDM basierend auf einem adaptiven Informationsaustauschfaktor.** IEEE Trans. Instrum. Meas. **2022,** 71, **1–13. [CrossRef]**
11. **Sun, C.; Ye, M.; Hu, G. Verteilte Optimierung für zwei Arten heterogener Multiagentensysteme.** IEEE Trans. Neural Netw. Lernen. Syst. **2020,** 32, **1314–1324. [CrossRef]**
12. **Shan, Y.; Fu, Y.; Chen, X.; Lin, H.; Lin, J.; Huang, K. LiDAR-basierte Methode zur Identifizierung befahrbarer Regionen für das Fahren von UGVs im Gelände .** IEEE Trans. Intell. Veh. **2023, 9, 3544–3557. [CrossRef]**
13. **Garaffa, LC; Basso, M.; Konzen, AA; de Freitas, EP Reinforcement Learning für die Erkundung mobiler Roboter: Eine Untersuchung.** IEEE Trans. Neural Netw. Learn. Syst. **2021,** 34, **3796–3810. [CrossRef]**
14. **Huang, CQ; Jiang, F.; Huang, QH; Wang, XZ; Han, ZM; Huang, WY Dual-Graph-Aufmerksamkeits-Faltungsnetzwerk für die 3D-Punktwolkenklassifizierung.** IEEE Trans. Neural Netw. Learn. Syst. **2022,** 35, **4813–4825. [CrossRef]**
15. **Nguyen, HT; Garratt, M.; Bui, LT; Abbass, H. Überwachtes Deep Actor Network für Imitationslernen in einer Boden-Luft-Koordinationsaufgabe für UAVs und UGVs. In Proceedings der 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, HI, USA, 27. November–1. Dezember 2017; IEEE: Piscataway, NJ, USA, 2017; S. 1–8.**
16. **Han, Z.; Yang, Y.; Wang, W.; Zhou, L.; Gadekallu, TR; Alazab, M.; Gope, P.; Su, C. RSSI-kartenbasiertes Trajektoriendesign für UGV gegen bösartige Funkquellen: Ein Ansatz des bestärkenden Lernens.** IEEE Trans. Intell. Transp. Syst. **2022,** 24, **4641–4650.** **[Querverweis]**
17. **Feng, Z.; Huang, M.; Wu, Y.; Wu, D.; Cao, J.; Korovin, I.; Gorbachev, S.; Gorbacheva, N. Approximation des Nash-Gleichgewichts für ein Anti-UAV-Jamming-Markov-Spiel unter Verwendung eines neuartigen ereignisgesteuerten Multi-Agent-Reinforcement-Learning.** Neural Netw. **2023,** 161, **330–342. [CrossRef]**
18. **Huang, X.; Deng, H.; Zhang, W.; Song, R.; Li, Y. Auf dem Weg zu multimodaler wahrnehmungsbasierter Navigation: Eine tiefe Verstärkung Lernmethode.** IEEE Robot. Autom. Lett. **2021, 6, 4986–4993. [CrossRef]**
19. **Wu, S.; Xu, W.; Wang, F.; Li, G.; Pan, M. Verteilte föderierte Deep Reinforcement Learning-basierte Trajektorienoptimierung für Luft-Boden-Kooperationsnetzwerke für Notfälle.** IEEE Trans. Veh. Technol. **2022,** 71, **9107–9112. [CrossRef]**
20. **Watkins, CJ; Dayan, P. Q-learning.** Mach. Lernen. **1992, 8, 279–292. [CrossRef]**
21. **Tran, TH; Nguyen, MT; Kwok, NM; Ha, QP; Fang, G. Gleitmodus-PID-Ansatz für robuste Low-Level-Steuerung eines UGV. In Proceedings der 2006 IEEE International Conference on Automation Science and Engineering, Shanghai, China, 8.–10. Oktober 2006; IEEE: Piscataway, NJ, USA, 2006; Seiten 672–677.**
22. **Schaul, T.; Quan, J.; Antonoglou, I.; Silver, D. Priorisierte Erfahrungswiedergabe.** arXiv **2015, arXiv:1511.05952.**
23. **Sutton, RS; Barto, AG Reinforcement Learning: Eine Einführung.** IEEE Trans. Neural Netw. **1998, 9, 1054. [CrossRef]**
24. **Lowe, R.; Wu, YI; Tamar, A.; Harb, J.; Pieter Abbeel, O.; Mordatch, I. Multi-Agent-Actor-Critic für gemischte kooperative und kompetitive Umgebungen.** arXiv **2017, arXiv:1706.02275.**
25. **Mirjalili, S.; Mirjalili, S. Genetischer Algorithmus.** Evolutionäre Algorithmen und neuronale Netzwerke: Theorie und Anwendungen; **Springer: Berlin/Heidelberg, Deutschland, 2019; S. 43–55.**
26. **Sehgal, A.; La, H.; Louis, S.; Nguyen, H. Deep Reinforcement Learning mit genetischem Algorithmus zur Parameteroptimierung. In Proceedings der 2019 Third IEEE International Conference on Robotic Computing (IRC), Neapel, Italien, 25.–27. Februar 2019; IEEE: Piscataway, NJ, USA, 2019; S. 596–601.**
27. **Chen, R.; Yang, B.; Li, S.; Wang, S. Ein selbstlernender genetischer Algorithmus basierend auf Reinforcement Learning für flexibles Job-Shop Planungsproblem.** Comput. Ind. Eng. **2020,** 149, **106778. [CrossRef]**

Machine Translated by Google

28. **Alipour, MM; Razavi, SN; Feizi Derakhshi, MR; Balafar, MA Ein Hybridalgorithmus, der einen genetischen Algorithmus und eine Multiagenten-Verstärkungslernheuristik verwendet, um das Problem des Handlungsreisenden zu lösen.** Neural Comput. Appl. **2018,** 30, **2935–2951.** [CrossRef]
29. **Liu, Z.; Chen, B.; Zhou, H.; Koushik, G.; Hebert, M.; Zhao, D. Mapper: Multi-Agenten-Pfadplanung mit evolutionärem Verstärkungslernen in gemischten dynamischen Umgebungen. In Proceedings der 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 25.–29. Oktober 2020; IEEE: Piscataway, NJ, USA, 2020; S. 11748–11754.**
30. **Huang, M.; Lin, X.; Feng, Z.; Wu, D.; Shi, Z. Ein Multi-Agenten-Entscheidungsansatz für die optimale Energiezuweisung in Mikronetzsystemen.** Electr. Power Syst. Res. **2023,** 221, **109399.** [CrossRef]
31. **Qiu, C.; Hu, Y.; Chen, Y.; Zeng, B. Auf einem tiefen deterministischen Policy Gradienten (DDPG) basierende drahtlose Energiegewinnungskommunikation.** IEEE Internet Things J. **2019, 6, 8577–8588.** [CrossRef]
32. **Littman, ML Markov-Spiele als Rahmen für Multi-Agent-Reinforcement-Learning. In Proceedings of the Proc International Konferenz über maschinelles Lernen, New Brunswick, NJ, USA, 10.–13. Juli 1994; S. 157–163.**
33. **Feng, Z.; Huang, M.; Wu, D.; Wu, EQ; Yuen, C. Multi-Agent Reinforcement Learning mit Policy Clipping und Durchschnittsbewertung für UAV-gestütztes Kommunikations-Markov-Spiel.** IEEE Trans. Intell. Transp. Syst. **2023,** 24, **14281–14293.** [CrossRef]
34. **Liu, H.; Zong, Z.; Li, Y.; Jin, D. NeuroCrossover: Ein intelligentes genetisches Locus-Auswahlschema für genetische Algorithmen unter Verwendung bestärkendes Lernen.** Appl. Soft Comput. **2023,** 146, **110680.** [CrossRef]
35. **Köksal Ahmed, E.; Li, Z.; Veeravalli, B.; Ren, S. Auf bestärkendem Lernen basierender genetischer Algorithmus zur Planung von Schulbussen.** J. Intell. Transp. Syst. **2022,** 26, **269–283.** [CrossRef]
36. **Chen, Q.; Huang, M.; Xu, Q.; Wang, H.; Wang, J. Auf Reinforcement Learning basierender genetischer Algorithmus zur Optimierung mehrdimensionaler sionales Datendiskretisierungsschema.** Math. Probl. Eng. **2020,** 2020, **1698323.** [CrossRef]
37. **Yang, J.; Sun, Z.; Hu, W.; Steinmeister, L. Gemeinsame Steuerung von Fertigung und Mikronetzsystem vor Ort über ein neuartiges neuronales Netzwerk integrierte Reinforcement-Learning-Algorithmen.** Appl. Energy **2022,** 315, **118982.** [CrossRef]
38. **Shi, H.; Liu, G.; Zhang, K.; Zhou, Z.; Wang, J. MARL Sim2real Transfer: Verschmelzung physischer Realität mit digitaler Virtualität in Metaverse.** IEEE Trans. Syst. Man Cybern. Syst. **2023,** 53, **2107–2117.** [CrossRef]
39. **Yu, C.; Velu, A.; Vinitsky, E.; Gao, J.; Wang, Y.; Bayen, A.; Wu, Y. Die überraschende Wirksamkeit von ppo in kooperativen Multi-Agenten -Spielen.** Adv. Neural Inf. Process. Syst. **2022,** 35, **24611–24624.**
40. **Rashid, T.; Farquhar, G.; Peng, B.; Whiteson, S. Gewichteter qmix: Erweiterung der monotonen Wertfunktionsfaktorisierung für tiefe Multi-Agent-Verstärkungslernen.** Adv. Neural Inf. Process. Syst. **2020,** 33, **10199–10210.**
41. **Wu, T.; Wang, J.; Lu, X.; Du, Y. Neukonfiguration eines AC/DC-Hybridverteilungsnetzes mit Mikronetzbildung unter Verwendung eines Multi-Agenten- Soft-Actor-Critic-Systems.** Appl. Energy **2022,** 307, **118189.** [CrossRef]
42. **Yan, C.; Xiang, X.; Wang, C.; Li, F.; Wang, X.; Xu, X.; Shen, L. PASCAL: Populationsspezifisches, auf dem Lehrplan basierendes MADRL für kollisionsfreies Flocking mit großen Starrflügel-UAV-Schwärmen.** Aerosp. Sci. Technol. **2023,** 133, **108091.** [CrossRef]
43. **Ackermann, JJ; Gabler, V.; Osa, T.; Sugiyama, M. Reduzierung des Überschätzungsbias in Multi-Agent-Domänen durch Double Zentralisierte Kritiker.** arXiv **2019, arXiv:1910.01465.**
44. **Xing, X.; Zhou, Z.; Li, Y.; Xiao, B.; Xun, Y. Adaptive kooperative Formationsbahnplanung für mehrere UAVs basierend auf einem verbesserten MATD3-Algorithmus des Deep Reinforcement Learning.** IEEE Trans. Veh. Technol. **2024.** [CrossRef]

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Hybrides maschinelles Lernen zur automatisierten Verkehrssicherheitsinspektion der Auckland Harbour Bridge

Munish Rathee 1,*, Boris Baÿcí´c 1,* und Maryam Doborjeh 1,2

1 Fakultät für Ingenieurwissenschaften, Computerwissenschaften und Mathematik, Auckland University of Technology, Auckland 1010, Neuseeland; mgholami@aut.ac.nz oder
2 maryam.gholami.doborjeh@aut.ac.nz Knowledge Engineering and Discovery Research Innovation, Auckland University of Technology, Auckland 1010, Neuseeland

* Korrespondenz: munishonthenet@gmail.com oder munish.rathee@aut.ac.nz (MR); boris.bacic@aut.ac.nz (BB)

Zusammenfassung: Die Auckland Harbour Bridge (AHB) nutzt eine bewegliche Betonbarriere (MCB), um den ungleichmäßigen, bidirektionalen Verkehrsfluss des täglichen Verkehrs zu regulieren. Zusätzlich zum Risiko menschlicher Fehler bei regelmäßigen Sichtprüfungen arbeiten die Mitarbeiter, die die MCB prüfen, unter unterschiedlichen Wetter- und Lichtverhältnissen und verausgaben sich in ergonomisch ungesunden Inspektionshaltungen mit dem zusätzlichen Gewicht der Schutzausrüstung, um Risiken, z. B. herumfliegende Trümmer, zu mindern. Um die Sichtprüfung einer MCB mithilfe von Computer-Vision-Technologie zu ergänzen, stellt diese Studie eine hybride Deep-Learning-Lösung vor, die Kernel-Manipulation mit benutzerdefinierten Transfer-Learning-Strategien kombiniert. Die Videodatenaufzeichnungen wurden unter unterschiedlichen Licht- und Wetterbedingungen (unter der Sicherheitsaufsicht von Branchenexperten) mit einem Hochgeschwindigkeitskamerasystem (120 fps) aufgenommen, das an einem MCB-Transferfahrzeug angebracht war. Bevor ein Sicherheitsrisiko identifiziert wird, z. B. die unsichere Position eines Stifts, der zwei 750 kg schwere Betonsegmente der MCB verbindet, wird eine mehrstufige Vorverarbeitung des räumlich-zeitlichen Interessenbereichs (ROI) durchgeführt, die ein gleitendes Fenster umfasst, bevor die Videobilder mit Diagnoseinformationen identifiziert werden. Diese Studie nutzt die ResNet-50-Architektur, erweitert um 3D-Faltungen, innerhalb des STENet-Frameworks, um räumlich-zeitliche Daten zu erfassen und zu analysieren und so eine Echtzeitüberwachung der Auckland Harbour Bridge (AHB) zu ermöglichen. Angesichts der geringen Anzahl an Sicherheitsanomalien konnten die anfänglichen, von Experten überprüften binären Klassifizierungsergebnisse (82,6 %) für sichere und unsichere (Eingriff erforderlich) Szenarien durch die Einbeziehung synthetischer Daten, Expertenfeedback und Neutraining des Modells auf 93,6 % verbessert werden. Diese Anpassung ermöglichte die optimierte Erkennung falsch positiver und falsch negativer Ergebnisse. In Zukunft wollen

Schlüsselwörter: Anomalieerkennung; Erkennung struktureller Schäden; Verkehrssicherheit; Computer Vision; maschinelles Lernen; Deep Learning; Transferlernen; ARDAD

check for updates

Zitat: Rathee, M.; Baÿcí´c, B.; Doborjeh, M. Hybrid Machine Learning für die automatisierte Verkehrssicherheitsinspektion der Auckland Harbour Bridge. Electronics 2024, 13, 3030. https://doi.org/10.3390/Elektronik13153030

Wissenschaftliche Herausgeber: Jiwei Zhang und Shaozhang Niu

Empfangen: 15. Juni 2024
Überarbeitet: 24. Juli 2024
Akzeptiert: 29. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Einleitung

Die Auckland Harbour Bridge, die sich 1,2 km über den Waitemata Harbour spannt, wurde am 30. Mai 1959 eröffnet. Anfangs bewältigte die Brücke täglich 11.205 Fahrzeuge, heute sind es rund 154.000 Fahrzeuge pro Tag, mit Spitzenwerten von über 200.000 Fahrzeugen aufgrund der Verlagerung des öffentlichen Nahverkehrs [1]. Die Brücke unterstützt die schnelle regionale Entwicklung; die Bevölkerungszahl der North Shore hat sich in den letzten 50 Jahren vervierfacht. Die New Zealand Transport Authority (NZTA), auch bekannt als Waka Kotahi, investiert jährlich bis zu 4 Mio. NZD in ihre Instandhaltung und beschäftigt rund 160 Mitarbeiter für laufende Modernisierungen und Instandhaltung. Wie alle wichtigen Verkehrsinfrastrukturen in Neuseeland ist auch die Auckland Harbour Bridge (AHB) aufgrund von Umwelt- und externen Faktoren mit erheblichen Instandhaltungsproblemen konfrontiert. Diese Probleme tragen zu den insgesamt hohen Instandhaltungskosten der Straßeninfrastruktur des Landes bei, die bis zu 1,1% des neuseeländisc

Dieses Dokument ist wie folgt gegliedert:

Machine Translated by Google

Abschnitt 1 bietet einen Überblick über das Problem und die Forschungsfrage; Abschnitt 2 umfasst einen umfassenden Literaturüberblick über die automatische Erkennung von Straßenmängeln und Anomalien; und Abschnitt 3 schlägt eine neue Methodik auf Grundlage einer hybriden Deep-Learning- Lösung vor, um Sichtprüfungen mithilfe von Computer-Vision-Technologie zu erweitern.

1.1. Hintergrund

Die 1990 installierte bewegliche Betonbarriere (MCB) auf der AHB hat den Verkehrsfluss und die Sicherheit während der Hauptverkehrszeiten verbessert. An jedem Werktag nutzt die AHB zwei Barrier Transfer Machines (BTMs), um die Fahrspurkonfigurationen anzupassen und den Verkehrsfluss während der Hauptverkehrszeiten durch das Bewegen von 750 kg schweren Betonblöcken zu verbessern. Die BTMs, die für die Bewältigung der täglichen Gezeiten unverzichtbar sind, wurden 2009 für jeweils 1,4 Millionen NZD eingeführt. Trotz der hohen Kosten verfügen die BTMs über kein automatisches Überwachungssystem, um die Integrität der beweglichen Betonbarrieren (MCBs) zu gewährleisten. Da es kein Sicherheitsinspektionssystem für die MCB-Barrieren gibt, müssen Mitarbeiter der NZTA über eine Meile unter gefährlichen Bedingungen und inmitten gefährlichen Verkehrs laufen, um die MCBs manuell zu inspizieren und ihre Integrität sicherzustellen [3].

Die Zuverlässigkeit des MCB-Systems hängt von der Integrität der Metallstifte ab, die die Barrieresegmente verbinden (Abbildung 1). Die entscheidende Funktion der Stifte besteht darin, Barrieresegmente zu sichern , die die Fahrspuraufteilung je nach Verkehrsfluss regeln [3]. Fehlfunktionierende oder gelöste Stifte gefährden die Verkehrssicherheit und beeinträchtigen die Effizienz des Systems, was zu potenziellen Verkehrsstörungen und erhöhten Unfallrisiken führt.

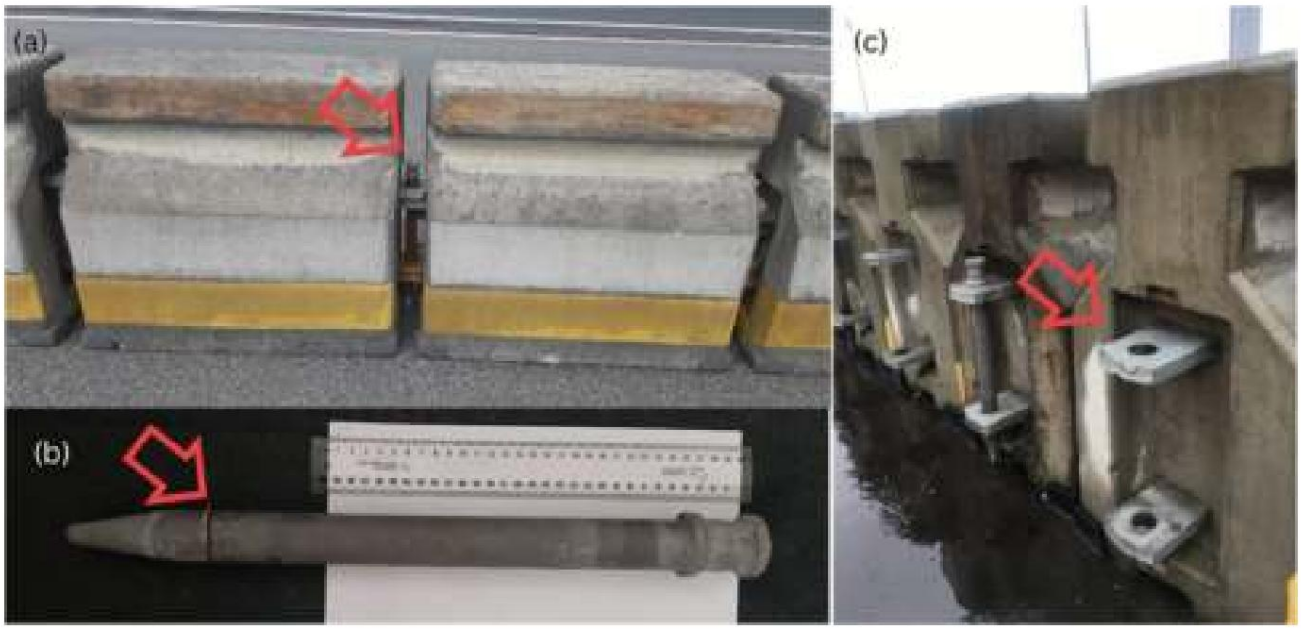


Abbildung 1. Ein bewegliches Betonbarrierensystem und seine Sicherheitsherausforderungen. (a) Stiftposition, die fixiert werden muss, (b) Metallstift mit abnehmbarem Sicherheitsring und (c) bewegliche Betonbarrierenverbindungen ohne Metallstifte.

Frühere Arbeiten zur automatisierten Infrastrukturinspektion sind in dynamischen und komplexen Umgebungen wie der AHB oft unzureichend [4–6]. Diese Forschung zielt darauf ab, das Feld durch die Einbeziehung einer neuartigen hybriden Deep-Learning- und raumzeitlichen Datenanalyse voranzutreiben, die eine genauere und zuverlässigere Erkennung von Sicherheitsanomalien in komplexen Umgebungen wie der AHB ermöglicht. Diese Studie verwendet raumzeitliche Analysen, einsetzbare KI-Algorithmen und halbautomatische synthetische Datengenerierung, um die Überwachung von Verkehrsbarrieren zu verbessern und die Forschung in praktische Lösungen zur Anomalieerkennung in Echtzeit umzuwandeln. Ziel ist es, die mit manuellen Inspektionen verbundenen Risiken zu verringern, die Verkehrssicherheit zu erhöhen (Abbildung 2) und die Effizienz von Barrieresystemen zu steigern. Aufbauend auf dem in **den Jahren 2019–2020 entwickelten Pro**

Machine Translated by Google

automatisiert die Erkennung von Metallstiften in beweglichen Betonbarrieren (MCBs) auf der Auckland Harbour Bridge (AHB).



Abbildung 2. Ausmaß der Herausforderungen bei der manuellen Inspektion und Befestigung des Metallstifts: (a) verschiedene Standorte beweglicher Betonbarrieren auf und um die Auckland Harbour Bridge; (b) Auffinden eines unsicheren Stifts und dessen manuelle Befestigung.

Hybride maschinelle Lernmodelle, die CNNs mit anderen maschinellen Lerntechniken integrieren , können Inspektionen revolutionieren, indem sie die Genauigkeit und Zuverlässigkeit der Erkennung und Klassifizierung anomaler Zustände verbessern [8,9]. Die Deep-Learning-Modelle, die ffi auf dem synthetisch erweiterten Datensatz trainiert hat, verwenden einen merkmalsbasierten Erkennungsansatz, um Videobilder auf Anzeichen von Stiftverschiebungen zu analysieren. Die Möglichkeit zur Echtzeitüberwachung ist für ein effektives Verkehrsmanagement von entscheidender Bedeutung, insbesondere während der Stoßzeiten. Sofortige Warnungen an die Brückenbetreiber ermöglichen rechtzeitige Eingriffe und verhindern, dass potenzielle Sicherheitsprobleme zu schwerwiegenderen Zwischenfällen eskalieren. Die Entwicklung eines solchen automatisierten Überwachungssystems fördert das Bauingenieurwesen und die Verkehrssicherheit, bietet eine skalierbare, effiziente Lösung für eine langjährige Sicherheitsherausforderung und schafft einen Präzedenzfall für ähnliche Anwendungen weltweit, was möglicherweise zu einer breiteren Einführung intelligenter Verkehrsmanagementlösungen in globalen städtischen Umgebungen führt [10].

In dieser Studie wurde ein hybrides maschinelles Lernsystem zur Echtzeit- Erkennung von Anomalien bei Verkehrssicherheitsinspektionen entwickelt, bei dem die Privatsphäre gewahrt bleibt. Die Beiträge dieser Forschung sind wie folgt aufgeführt:

- Erstellt ein Artefakt zur Verkehrssicherheitsanalyse, das auf Szenarien in über 20 Ländern und Hunderten ähnlicher Verkehrsszenen [10] mit beweglichen Betonbarrieren skalierbar ist.
- Führt eine halbautomatische Methode zur Generierung synthetischer Daten unter Verwendung einer neuartigen Hintergrundklontechnik ein . Der neuartige Ansatz behebt Datenspärlichkeit und verbessert das Modelltraining mit Wiederholbarkeitswert für andere Computer Vision-Fallstudien, bei denen Probleme mit dem Datensatzausgleich auftreten.
- Verfeinert Klassifizierungsmethoden, um falsch positive und negative Ergebnisse auszugleichen, wodurch die Erkennungsgenauigkeit von 82,6 % (berichtet in früheren Peer-Review-Studien [11]) auf 93,6 % verbess Das Erreichen dieser 11 % höheren Genauigkeit in komplexen Verkehrsszenen mit chaotischen Hintergründen und Lichtverhältnissen unterstreicht die Eignung des Artefakts für Anwendungen in der realen Welt.
- Erfolgreiches Navigieren durch gefährliche Verkehrsszenen zur Datenerfassung unter Einhaltung der Sicherheitsprotokolle der Branche. Dieser wiederholbare Ansatz bietet eine umfassende Blaupause für die Bewältigung ähnlicher Szenarien, stellt die Zufriedenheit der Beteiligten sicher und erzielt ausreichend Daten.

Machine Translated by Google

1.2. Forschungsfragen und Modellierungskonzepte

In diesem Artikel wird die Entwicklung automatisierter Lösungen vorgestellt, die Computer Vision und Deep Learning nutzen, um die Verkehrssicherheit und Betriebseffizienz durch die Automatisierung der Inspektion von Metallstiften in beweglichen Betonbarrieren zu verbessern. Diese Forschung untersucht das Potenzial visueller Sensoren, Anomalien bei Transportaktivitäten zu erkennen und gleichzeitig die Privatsphäre zu wahren. Unsere Forschung konzentriert sich auf die bewegliche Barriere der Auckland Harbour Bridge und automatisiert die Sicherheitsüberprüfung und verbessert die Anomalieerkennung für kritische, aber unterrepräsentierte Klassen wie unsichere Stiftpositionen . Darüber hinaus wurden Klassifizierungstechniken verfeinert, um falsch positive und falsche negative Ergebnisse auszugleichen und so die Z

1. Können relevante Informationen aus visuellen Sensoren zur Anomalieerkennung extrahiert werden in Transportaktivität unter Wahrung der Privatsphäre?

(a) Was sind die optimalen Methoden, um wichtige Informationen aus visuellen Medien zu extrahieren und Anomalien in der Transportaktivität zu erkennen, ohne die Privatsphäre zu gefährden?

(B) Inwieweit kann die Sicherheitsüberprüfung der beweglichen Barriere der Auckland Harbour Bridge und der umliegenden Bereiche mithilfe von KI-, CV- und DL-Methoden automatisiert werden ?

2. Wie kann die Generierung synthetischer Daten optimiert werden, um Minderheitsklassen, wie z. B. unsichere Pin-Positionen, bei Aufgaben zur Anomalieerkennung im Bereich der Verkehrssicherheit darzustellen?

3. Wie kann die Klassifikationsleistung verbessert werden, um unter Berücksichtigung aktueller und zu erwartender Datenszenarien ein optimales Gleichgewicht zwischen falsch **-positiven und falsch-negativen Ergebnissen aus dem frühen Proof-of-Concept (PoC)** [7] zu erreichen?

2. Literaturübersicht

Die automatische Erkennung von Straßenschäden und Anomalien (ARDAD) nutzt Computer Vision und kombiniert traditionelle und Deep-Learning-Methoden mit unüberwachtem Lernen [12,13]. Herkömmliche unbeaufsichtigte Methoden, die auf Datensätzen mit normalen Bedingungen basieren, führen in komplexen Umgebungen häufig zu hohen Falschalarmraten [14]. In dieser Forschung wird das Spatio-Temporal Enhanced Network (STENet) eingesetzt, um das Problem komplexer Verkehrsszenen zu lösen . Dabei werden zeitliche und räumliche Daten für eine bessere Generalisierung und eine robuste Anomalieerkennung genutzt.

2.1. Hafenbrücke von Auckland

Die New Zealand Transport Agency (NZTA) verwaltet die Auckland Harbour Bridge (AHB), eine achtspurige Autobahn, die täglich über 200.000 Fahrzeuge bewältigt. Die in den 1990er Jahren installierte bewegliche Betonfahrbahnbarriere (MCB) ist unerlässlich, um Unfälle zu verhindern und den Verkehrsfluss während der Stoßzeiten zu optimieren. Das MCB-System, das aus 750 kg schweren Betonblöcken besteht, die durch Metallstifte verbunden sind, neigt dazu, sich aufgrund von Verkehr und Umgebungsvibrationen zu verschieben, was erhebliche Risiken birgt [1,3,15]. Barrier Transfer Machines (BTMs) erleichtern die Bewegung der MCB und passen die Fahrspuren normalerweise viermal täglich an, um den Verkehrsfluss effektiv zu steuern.

Die manuelle Inspektion der Stifte ist arbeitsintensiv und erfolgt unter schlechten ergonomischen Bedingungen, wodurch sie anfällig für menschliches Versagen ist. Obwohl das System die Fahrspuren schnell wechseln kann – etwa 10 Minuten für einen ein Kilometer langen Abschnitt – kann das Fehlen automatisierter Stiftinspektionsmechanismen in den heute bestehenden 1,4 Millionen NZD teuren BTMs als erhebliches Versäumnis und als Gelegenheit angesehen werden, CV in Wartungs- und Sicherheitsprotokollen anzuwenden [1,15,16]. Das System der beweglichen Betonbarrieren (MCB) ist zwar bei der Verwaltung von Gegenverkehrs- und stark befahrenen Fahrspuren wirksam, erfordert jedoch eine kontinuierliche manuelle Überwachung, um seine strukturelle Integrität und Betriebszuverlässigkeit

2.2. Überprüfung der Umfragen zu ARDAD

Die Anwendung künstlicher Intelligenz (KI) und Deep Learning bei der Erkennung von Anomalien und Objekten in der Straßeninfrastruktur hat im letzten Jahrzehnt bedeutende Durchbrüche erzielt **[5,12,13]. KI-Fortschritte haben sich auf Bereiche wie autonomes F**

Machine Translated by Google

Gesichtserkennung und personalisierte Gesundheitsfürsorge [17,18]. Deep-Learning-Techniken, insbesondere Convolutional Neural Networks (CNNs), haben die Bildverarbeitung revolutioniert, indem sie Segmentierungs-, Erkennungs- und Rekonstruktionsaufgaben automatisieren [6]. CNNs verarbeiten und klassifizieren Bilddaten effizient durch Pooling und Faltung, extrahieren Merkmale und reduzieren die Dimensionalität [19–22]. Im Gegensatz zu herkömmlichen Methoden, die auf manuell entwickelten Merkmalen beruhen, lernen Deep-Learning-Modelle Merkmalsdarstellungen direkt aus Daten. Architekturen wie R-CNN und seine Varianten zeichnen sich durch die Objekterkennung aus, indem sie Region-Proposal-Methoden für vorab ausgewählte Interessenbereiche verwenden **[23].**

Darüber hinaus verbessert Transferlernen die Generalisierbarkeit von Modellen, indem es Modelle, die für eine Aufgabe trainiert wurden, auf verwandte, aber unterschiedliche Aufgaben anwendet [24]. Transferlernen nutzt vorab trainierte Modelle, um die Leistung bei neuen Datensätzen zu verbessern, häufig durch Feinabstimmung der letzten Schichten von Netzwerken wie ResNet-50, das bei der Visual Recognition Challenge einen bemerkenswerten Erfolg gezeigt hat **[25]. Der Ansatz unterstreicht die Vielseitigkeit und Anpassungsfähigkeit von Deep-Learning-Modellen im Umgang mit vielfältigen und komplexen visuellen Datenumgebungen.**

2.3. Video- und Bildoptimierungstechniken in ARDAD

Die Auswahl eines geeigneten Farbraums bei der Bildsegmentierung ist entscheidend und oft anwendungsspezifisch, wobei keine Einigkeit über die beste Wahl besteht. Gängige Farbräume sind RGB, LAB, CMY, XYZ, HSV, YCbCr, YIQ, YUV und DHT [26,27]. Beispielsweise ist RGB effektiv bei Änderungen der Bilirubinkonzentration (Gleichung (1)), während CMY und HSV sich bei anderen Aufgaben hervortun **[28].** Zur Optimierung des Prozesses **wurden fortschrittliche Methoden wie eindimensionale Histogramme zur automatischen Farbraumauswahl [29] und zweidimensionale Histogramme zur verbesserten Segmentierung [30] entwickelt. Spezifische Anwendungen wie YCbCr zur Gesichtserkennung [31] und YIQ für Satellitenbilder [32] erfordern maßgeschneiderte Ansätze. Zur Überwachung der Auckland Harbour Bridge (AHB) kann die Transformation der Y-Komponente (Gleichung (1)) von RGB nach YCbCr die Erkennung struktureller Anomalien verbessern, indem der Schwerpunkt auf der Leuchtdichte liegt und Farbabweichungen aufgrund von Licht- oder Wetterbedingungen minimiert werden.**

$$Y = 0{,}299R + 0{,}587G + 0{,}114B \quad (1)$$

Darüber hinaus wurden Techniken wie Histogramm-Schwellenwertverfahren und SVM-basiertes Clustering angewendet, um die Segmentierung durch Strukturierung der Pixeldaten in zusammenhängende Farbgruppen zu verbessern [

2.4. Fortgeschrittene Techniken in der Video- und Bildanalyse für ARDAD

In der Computervision umfasst die Objekterkennung in Videos das Erkennen der Bewegung von Objekten über mehrere Frames hinweg unter Verwendung von Techniken wie Hintergrundsubtraktion, Frame- Differenzierung und optischem Fluss. Das Gaussian Mixture Model (GMM) ist eine gängige Methode zur Hintergrundtrennung, die die Unterscheidung zwischen Vordergrund und Hintergrund verbessert, indem Pixelverteilungen mit Gaussian-Mischungen modelliert werden **[35–37]. Template Matching, ein weiterer praktischer Ansatz, verwendet MATLAB 2022a-Funktionen wie** normxcorr2 **und** regionprops , **um Objekte in aufeinanderfolgenden Frames zu verfolgen , indem der Peak der normalisierten Kreuzkorrelation identifiziert wird [38]. Template Matching, wie durch Gleichung (2) dargestellt, berechnet den intensitätsgewichteten Schwerpunkt eines Objek**

$$xc = \frac{\sum_{i=1}^{N} xi \cdot wi}{\sum_{i=1}^{N} wi} \quad (2)$$

wobei xc die Schwerpunktposition, xi die Pixelposition und wi die Pixelintensität ist.
Durch **die Multiplikation** von **xi mit wi im Zähler wird sichergestellt, dass die Position jedes Pixels anhand seiner Intensität gewichtet wird, wodurch Pixeln mit höheren Intensitäten bei der Schwerpunktberechnung mehr Bedeutung verliehen wird .**

Techniken zur Trennung von Vordergrund und Hintergrund teilen ein Video in statische Hintergründe und dynamische Vordergründe auf. Dabei kommen Modelle zum Einsatz, die sich an Veränderungen in der Szene anpassen, um Bewegungen zu erkennen und Objekte zu erkennen [39,40]. Jüngste Fortschritte im GMM haben Methoden eingeführt, die komplexe Szenen wie Verkehr verarbeiten und die Erkennung von langsamen Be

Machine Translated by Google

oder übergroße Fahrzeuge [41]. Jeder Pixel in der Videosequenz wird als Mischung von K Gauß-Verteilungen modelliert. Die Wahrscheinlichkeit, einen Pixelwert X zum Zeitpunkt t zu beobachten, ist durch Gleichung (3) gegeben.

$$P(Xt) = \ddot{y}_{k=1}^{K} \ddot{y}k,t \cdot N\, Xt\, \mu k,t,\, \ddot{y}k,t \quad (3)$$

Dabei ist **ÿk,t das Gewicht des k th Gauß-Komponente zum Zeitpunkt t, N ist die Gauß- Verteilung, μk,t ist der Mittelwert der k- ten Gauß-Verteilung und ÿk,t ist die Kovarianzmatrix der k-ten Gauß-Verteilung. Bilddifferenzmethoden hingegen erkennen Bewegung durch Vergleichen der Pixelunterschiede von einem Bild zum nächsten, eine einfache aber effektive Methode zur Erkennung bewegter Objekte [42]. Optische Flussmethoden ermitteln Bewegung, indem sie das Muster der scheinbaren Bewegung von Objekten, Oberflächen und Kanten in einer visuellen Szene analysieren, die durch die relative Bewegung zwischen einem Beobachter und der Szene verursacht wird [43]. Beispielsweise synthetisiert Gleichung (4) die Bewegungserkennung unter Verwendung von Bilddifferenzen | Ft ÿ Ft ÿ 1 | zur Erkennung von Änderungen und den optischen Fluss, dargestellt durch ÿlt · vt , zur Analyse von Bewegungsmustern und -geschwindigkeiten, wobei ÿ ihre relativen Beiträge anpasst.**

$$\ddot{y}Mt = | Ft \ddot{y} Ft \ddot{y} 1 | + \ddot{y} \cdot \ddot{y}lt \cdot vt \quad (4)$$

Die Objektverfolgung in der Computervision folgt auf die Objekterkennung und verwendet Farbe, Textur, Form, Größe und Ausrichtung, um Objekte wie Fahrzeuge oder Fußgänger über Videobilder hinweg zu ver Eine robuste Verfolgung ist bei allen Kamerapositionen, Lichtverhältnissen und unübersichtlichen Szenen unerlässlich **[44]. Techniken wie der Kalman-Filter und der Partikelfilter sind Standard für die Verfolgung linearer und nichtlinearer Bewegungen [45].** Die **Blob-Analyse verfolgt Objekte in binärisierten Bildern anhand von Merkmalen wie Fläche und Begrenzungsrahmenabmessungen [46]. Das Bayer-Filtermuster berechnet den Abstand zwischen Objektschwerpunkten und hilft bei der Verfolgung und Messung der Fahrzeugges Der Kalman-Filter verfeinert die Flugbahnvorhersage durch die Kombination von vorhergesagten und tatsächlichen Standorten [47,48]. Für nichtlineare Szenarien verwendet der erweiterte Kalman-Filter die Jacobi-Matrix für Übergangsmatrizen (Gleichung (5)).**

$$\hat{x}k|k = \hat{x}k|k\ddot{y}1 + Kk\, zk\, \ddot{y}\, Hk\hat{x}k|k\ddot{y}1 \quad (5)$$

In Gleichung (5) ist **x^k|k die aktualisierte Zustandsschätzung zum Zeitpunkt k;** x^k|kÿ1 **ist die vorhergesagte Zustandsschätzung zum Zeitpunkt k, basierend auf der Schätzung zum Zeitpunkt k ÿ 1; Kk ist der Kalman-Gewinn, der die Mischung von Vorhersage und Beobachtung bestimmt; zk ist die tatsächliche Messung zum Zeitpunkt k; und Hk ist die Messmatrix, die den Zustandsraum in den**

Der Partikelfilter verbessert die Verfolgung in komplexen Szenarien durch Abtastung und Neuabtastung von Objektmerkmalen über mehrere Frames hinweg und erweist sich in nichtlinearen und unübersichtlichen Umgebungen als effektiver als andere Methoden [49,50]. Er verwendet einen Korrelationspartikelfilter, um Maßstabsabweichungen und Merkmalsabhängigkeiten zu verwalten.

Der Mean-Shift-Algorithmus, der zur Objektverfolgung ein statistisches Farbmodell verwendet, konvergiert iterativ zu Bereichen mit hoher Dichte im Farbraum und nutzt 8-Punkt-Konnektivität, um Objekte zu loka Das Adaptive Local Movement Model (ALMM) berücksichtigt Bewegungen, indem es sich auf regionale Bereiche statt auf das gesamte Objekt konzentriert und sich damit als wirksam gegen Okklusionen und schnelle Bewegungen erweist **[52].**

Die Datenerfassung und manuelle Kennzeichnung in ARDAD-Systemen kann ressourcenintensiv und zeitaufwändig sein. Während Deep-Learning-Ansätze aufgrund ihrer Fähigkeit, traditionelle Methoden zu übertreffen, erhebliche Vorteile bei der Bild- und Videoverarbeitung bieten, erfordern sie umfangreiche Datensätze und Rechenleistung [53]. Alternativ kann die expertengesteuerte Merkmalsextraktion mit traditionellen maschinellen Lernansätzen (ML) bei kleineren Datensätzen und geringerem Rechenaufwand effektiv sein, was sie für anfängliche Zielsystemdesigns geeignet macht [54]. Vortrainierte Deep-Learning-Modelle erleichtern die automatische Merkmalsextraktion und ermöglichen Transferlernen, das Modelle ohne umfangreiche Rechenressourcen [55] und katastrophales Vergessen an neue Funktionen anpasst. Die Fusion von Transferlernen mit traditionellem maschinellem Lernen verbessert die räumlich-zeitliche Klassifizierung, indem vortrainierte Modelle für eine effiziente Merkmalsextraktion genutzt und traditionelle Modelle für die Sequenzanalyse angewendet werden , was eine Möglichkeit zur Verbesserung der Klassifizierungsgenauigkeit bietet (Abbildung 3).

Machine Translated by Google

Die Kombination von Deep Learning und traditionellem maschinellem Lernen wird in verschiedenen Bereichen angewendet, beispielsweise bei der Verkehrsflussvorhersage und der Überwachung des Gesundheitszustands von Nutzpflanzen [56,57].

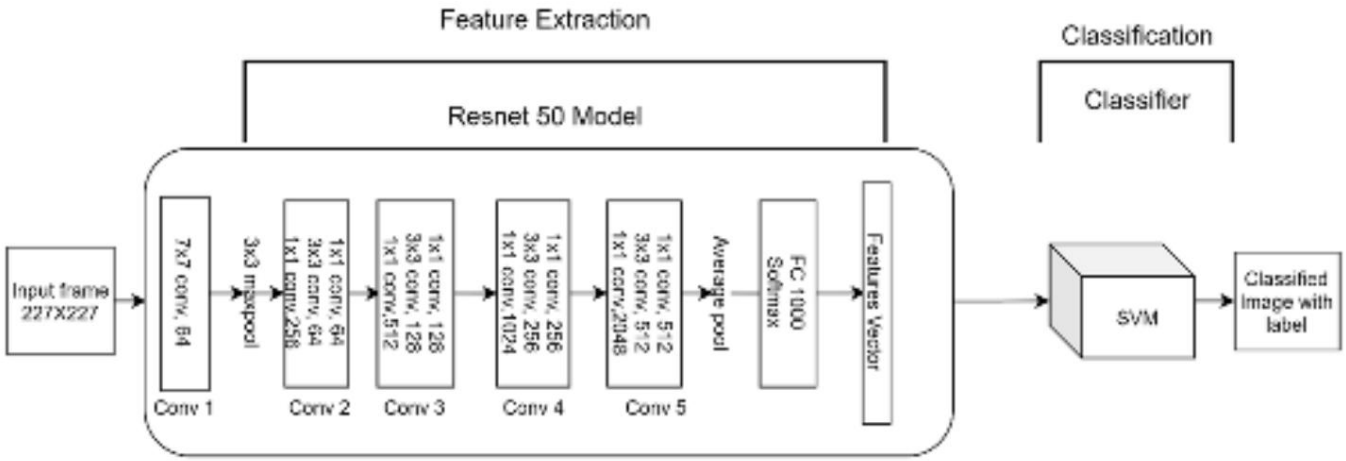


Abbildung 3. Integration von Transferlernen mit traditionellem ML: ResNet 50 extrahiert Merkmale aus Eingabebildern , die ein SVM klassifiziert. Die Integration der Deep-Learning-Funktionen von ResNet 50 zur Generierung von Merkmalsräumen ohne Rückgriff auf Expertenwissen in Kombination mit einem traditionellen Klassifikator (z. B. SVM) bietet die Möglichkeit, die Leistung der Bildklassifizierung zu verbessern.

Die Fusion von Transferlernen mit traditionellem maschinellem Lernen (Abbildung 3) für räumlich -zeitliche Klassifizierungsaufgaben soll die Funktion softmax() (6) ersetzen, die üblicherweise als letzte Schicht eines neuronalen Netzwerks verwendet wird, um hochdimensionale Merkmalsvektoren (ÿR) typischerweise in eine Wahrscheinlichkeit möglicher Ergebnisse umzuwandeln , d . h. normalisiert für die beabsichtigte Ausgabeklassenverteilung.

$$y = \text{softmax}(W_{ff}(X) + b) \quad \hat{y} = \text{softmax}(W_{fRNN}(f_{CNN}(X)) + b) \tag{6}$$

wobei die beim Transferlernen verwendete Verlustfunktion ist

$$L(\ddot{y}) = \frac{1}{N}\sum_{i=1}^{N} L(y_i, f(X_i; \ddot{y})) \quad L(\ddot{y}) = \frac{1}{N}\sum_{i=1}^{N} L_{fCNN}(y_i, f(X_i; \ddot{y})) \tag{7}$$

Hier ist $L(\ddot{y})$ der durchschnittliche Verlust über N Trainingsproben, X_i und y_i sind die Eingabe- und Label-für das i-te Probe, f ist das Modell mit Parametern $\ddot{y}$, initialisiert aus einem vortrainierten Modell, und $\ddot{y}$ ist die Verlustfunktion, wie z. B. der Kreuzentropieverlust für die Klassifizierung. In praktischen Anwendungen verbessert ein solcher hybrider Ansatz die Effizienz und Genauigkeit von räumlich-zeitlichen Modellen in praktischen Anwendungen [56,57].

Zu den wichtigsten Überlegungen aus Sicht der Projektmethodik gehören die Verfügbarkeit und Ausgewogenheit der Datensätze, die Rechenleistungsbeschränkungen der vorgesehenen Zielplattformen und die Integration traditioneller ML-Techniken für verbesserte Dateneinblicke. Die angewandten Methoden umfassen Transferlernen für eine verbesserte Objekterkennung und -klassifizierung, die Nutzung von Deep Learning für die Merkmalsgenerierung und die Verwendung traditioneller ML-Klassifikatoren zur Gewinnung weiterer Erkenntnisse, wodurch ein umfassender Ansatz für die Entwicklung von ARDAD-Lösungen gewährleistet wird.

3. Materialien und Methoden

Dieses Dokument dokumentiert eine kritische Phase im Entwicklungsprozess, den Übergang vom anfänglichen Proof-of-Concept (PoC) [7] zu einem minimal funktionsfähigen Produkt (MVP) für die bewegliche Betonbarriere (MCB) der Auckland Harbour Bridge. Dabei werden zusätzliche Forschungsprobleme im Zusammenhang mit der Erkennung von Anomalien in einem raumzeitlichen Rauschen behandelt .

Um sich den Herausforderungen der realen Welt zu stellen und technologische Lösungen zu entwickeln, musste die Forschungsmethodik erheblich angepasst werden, um den Sicherheitsregeln der NZTA Waka Kotahi in Bezug auf die Aufzeichnung von Videos zu entsprechen, die absichtlich in Temperaturzonen plat

Machine Translated by Google

vorübergehend unsichere Positionen (auf den kurzen Barrieresegmenten im sicheren Arbeitsbereich vor der Harbour Bridge). Die Datenerfassungsgeräte sind mithilfe provisorischer Vorrichtungen sicher an einem BTM befestigt (Abbildung 4 und Tabelle 1). Die Nutzung mehrerer Barrier Transfer Machine (BTM)-Operationen zur Erfassung spärlicher Daten zu Minderheitenklassen war jedoch nicht praktikabel, was zur Entwicklung eines Hintergrundklonierungsansatzes führte, um synthetische Frames zu erstellen, die unsichere Pinpositionen darstellen (Tabelle 2). Ein solcher methodischer Ansatz ist auf mehrdimensionale Problemlösungen für Projekte ausgerichtet, bei denen eine laufende Anpassung an externe Faktoren erforderlich ist [58,59].



Abbildung 4. Datenerfassungs-Setup: iPhone 13 Pro (für LiDAR), Samsung A7 Mobile, Apple iPad 6, externe Powerbank, GoPro-Kameras und Montagezubehör sowie Kamera- und iPhone-Montage auf einem BTM.

Tabelle 1. Übersicht über die Datenerfassung. Ausrüstung, Kameraspezifikationen, Wetterbedingungen, aufgetretene Herausforderungen und technologischer Fortschritt in drei Datenerfassungssitzungen.

Sitzung NEIN.	Ausrüstung	Kamera-Spezifikationen	Wetter Bedingungen	Herausforderungen	Tech. Integration
1	GoPro 8, GoPro 5, Samsung A7 Mobile, Apple iPad 6, Klebeband, Powerbank und Befestigungsstreifen auf dem BTM. Barrier Transfer Machine (BTM) wird zur Aufrechterhaltung optimaler Kamerawinkel während Aufnahmesitzungen verwendet	Auflösung: 720p, Bildrate: 240 fps, Sichtfeld: Schmal, Audio: Nur Wind, Protune: Aktiviert, Weißabgleich: Auto, Farbe: Flach, Verschluss: Auto, ISO-Limit: 6400, Schärfe: Hoch, Audio-Protune: Mittel, Automatische Drehung: Auto	Gelegentlicher Regen und bedeckt	Vibrationen, Kameraheizung, Montage Integritätsprüfungen	Erste Sitzung: Grundlagen für die inkrementelle Datenerfassung
2. Platz	GoPro 8 und 5 an den Vorder- und Hinterarmen von BTM, sonstige Ausrüstung gleich wie beim ersten	Wie bei der ersten Sitzung	Sonnig	Kameraheizung, Abdichtung, Batterie Autonomie bei Hoher Rahmen Preise	Im Mittelpunkt der Sitzung standen sonnige und ideale Lichtverhältnisse

Technische Daten

Machine Translated by Google

Tabelle 1. Fortsetzung

Sitzungsnummer	Ausrüstung	Kamera-Spezifikationen	Wetterbedingungen	Herausforderungen	Tech. Integration
3. Platz	GoPro 9 und iPhone 13 Pro im Hardcase mit 3D Punktwolken-App auf dem Vorderarm des BTM, GPS auf beiden Geräten aktiviert. iPhone 13 Pro genutzt mit einer speziellen Anwendung zur Erstellung von 3D-Punktwolkendaten, die die Tiefe und Qualität der räumlichen Analyse verbessert	GoPro 9: 1080p-Auflösung, 240 fps, Verbesserte Spezifikationen für raue Bedingungen iPhone 13 Pro: Lidar-fähige Kamera zur Erfassung von 3D-Punktwolkendaten und GPS zur Standortverfolgung	Frühmorgens dunkel und regnerisch	Sicherstellen der Kamera Abdichtung, Aufrechterhaltung Batterie Autonomie, Gerätestabilität in rauen Bedingungen	Integration von Lidar und GPS für Fortgeschrittene Erfassung räumlicher Daten; Nutzung von 5G für Echtzeit Datenübertragung. Verwendung der 5G-Technologie wurde als Verbesserung angesehen Datenerfassung und -übertragung in Echtzeit, um sicherzustellen, dass große Datensätze effektiv verwaltet werden können

3.1. Schritt 1: Datenerfassung und synthetische Datengenerierung

Die inkrementelle Datenerfassung begann mit der ersten Sitzung bei gelegentlichem Regen und bewölktem Himmel und schuf damit einen Präzedenzfall für die Belastbarkeit des Prozesses. Während der zweiten Sitzung, die bei sonnigem Wetter durchgeführt wurde, wurden zwei GoPro-Kameras an den Armen der Barrier Transfer Machine (BTM) montiert, wobei technische Herausforderungen wie Vibrationen, Überhitzung der Kamera, die Notwendigkeit der Abdichtung und die Aufrechterhaltung der Akkulaufzeit bei hohen Bildraten deutlich wurden. Die dritte Sitzung wurde bei starkem Regen durchgeführt, was weitere Verbesserungen des Protokolls und des Systemdesigns erforderlich machte. Gleichzeitig wurde die Integration fortschrittlicher Technologien wie Lidar, GPS und 5G-Netzwerke in Betracht gezogen, um die Robustheit und Leistungsfähigkeit der Datenerfassungszyklen

Die vorbeugenden manuellen Inspektionen beweglicher Betonbarrieren (MCBs) vor jedem BTM-Vorgang zur Fahrbahnänderung schlossen die Möglichkeit aus, die erforderlichen unsicheren Stiftpositionen auf natürliche Weise zu erfassen (Abbildung 5). Die Vorsichtsmaßnahme ist verständlich, da während des MCB-Transfervorgangs möglicherweise lose Betonblöcke ein erhebliches Risiko darstellen , das die Integrität der BTM und die Sicherheit der

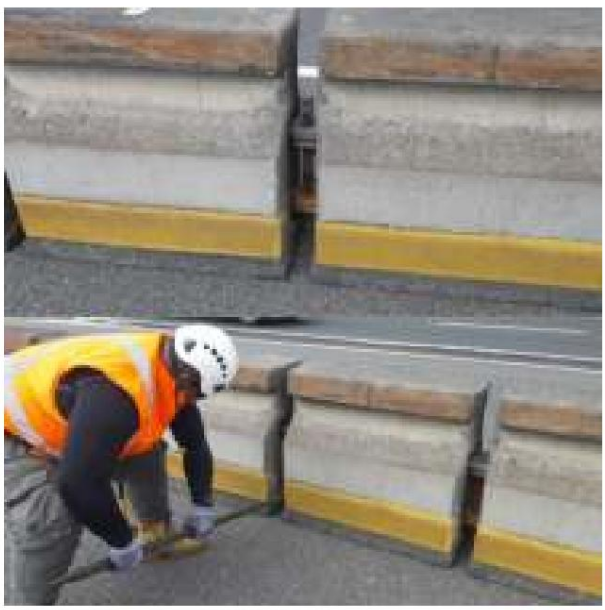


Abbildung 5. Auf Wunsch des Autors wurde der Stift von NZTA-Mitarbeitern manuell aus seiner Position geschoben . Der **Stift**

Abbildung 5. Auf Wunsch des Autors wurde der Stift von NZTA-Mitarbeitern manuell aus seiner Position geschoben. Der Vorgang war anspruchsvoll und arbeitsintensiv, und der Mitarbeiter benötigte mehrere Minuten pro Stift für die Anpassung. Hinweis: Eine solche Option war nicht praktikabel, da sie nicht die natürliche Umgebung des ROI darstellt, in der reale Probleme gelöst werden müssen.

Machine Translated by Google

Tabelle 2. Halbautomatischer Prozess zum Generieren eines vielfältigen Datensatzes synthetischer Pin-Positionen (basierend auf der in Abbildung 6 dargestellten manuellen Klonmethode).

Algorithmus: Halbautomatische synthetische Bildgenerierung für die Pin-Position
Eingabe: Reihe von Pin_OK-Bildern, Pin_OUT-Vorlagenbild
Parameter:
-Normalisierungsmethode für die Ausrichtung
-Segmentierungsmethode
-Verschiebungsberechnungsmethode
-Transformationsmethode
-Hintergrundrekonstruktionsmethode
-Kantenverfeinerungsmethode
Ausgabe: Synthetischer Datensatz mit unterschiedlichen und genau positionierten Pin_Out-Frames
1: Bilder laden 1.1:
Die Serie der Pin_OK-Bilder laden 1.2: Das
Pin_OUT-Vorlagenbild laden 2: Ausrichtung
normalisieren 2.1: Die
Ausrichtung des Pins in der Serie der Pin_OK-Bilder auf eine Standardreferenzposition normalisieren 3: Jedes normalisierte
Pin_OK-Bild verarbeiten 3.1: Den Pin mit einer
regionsbasierten Segmentierungsmethode segmentieren 3.2: Die benötigte
Verschiebung basierend auf der Pin_OUT-Vorlage berechnen 3.3: Die berechnete
Verschiebung anwenden, um die Pin-Position anzupassen und ein synthetisches Pin_Out-Bild zu erstellen 3.4: Den
Hintergrund rekonstruieren, wo der Pin ursprünglich platziert wurde 3.5: Die
Kanten des verschobenen Pins verfeinern, um eine nahtlose Integration zu
gewährleisten 3.6: Die Qualität des synthetischen Bildes visuell
validieren 3.7: Wenn die Qualität akzeptabel ist, das synthetische Bild dem Datensatz
hinzufügen 3.8: Wenn Anpassungen erforderlich sind, die Parameter verfeinern und den
Vorgang wiederholen 4: Für
alle Bilder wiederholen 4.1: Den Vorgang für alle Pin_OK-Bilder in die Serie, um einen umfassenden Datensatz mit unterschiedlichen Pin_Out-Positionen zu erstellen

Angesichts der geringen Anzahl unsicherer Pin-Positionen wurde eine neuartige Methode zur Erstellung synthetischer Daten eingeführt, bei der Hintergrundklone aus Original-Videobildern erstellt werden (Ab Durch das Klonen können verschiedene Darstellungen der Minderheitsklasse erzeugt und so die verzerrte Datenverteilung behoben werden. Erste Versuche führten zu unvollständigen Frames mit ungleichmäßigen Rändern, aber der Prozess wurde durch iterative Verfeinerung verfeinert, was zu hochpräzisen synthetischen Frames führte, die den Datensatz erheblich verbesserten. Das Hinzufügen der synthetischen Frames erleichterte die Feinabstimmung der Modellpräzision und des Rückrufs, was für die Wirksamkeit automatisierter Inspektionssysteme entscheidend ist und eine höhere Empfindlichkeit gegenüber falschen Negativen gewährleistet – eine Priorität für Verkehrssicherheitssysteme. Der Klonansatz erwies sich als entscheidend, um die durch manuelle Datenerfassungsmethoden auferlegten Einschränkungen zu umgehen und ebnete den Weg für eine sicherere, effizientere Methode zum Trainieren robust

Wie in Abbildung 6 dargestellt, dauert das Erstellen synthetischer Frames mit einem Bildeditor mindestens 20 Minuten pro Frame, was den Bedarf an einer effizienteren und automatisierteren Lösung unterstreicht (Tabelle 2). Der Prozess beginnt mit dem Laden einer Reihe von Bildern, auf denen sich der Pin in einer sicheren Position (Pin_OK) befindet, und eines Vorlagenbilds, auf dem sich der Pin in einer unsicheren Position (Pin_OUT) befindet. Die Methode umfasst mehrere wichtige Schritte: Normalisieren der Ausrichtung der Pins in jedem Pin_OK- Bild auf eine standardmäßige Referenzposition, Segmentieren der Pins mithilfe einer regionsbasierten Segmentierungsmethode und Berechnen der Verschiebung

Machine Translated by Google

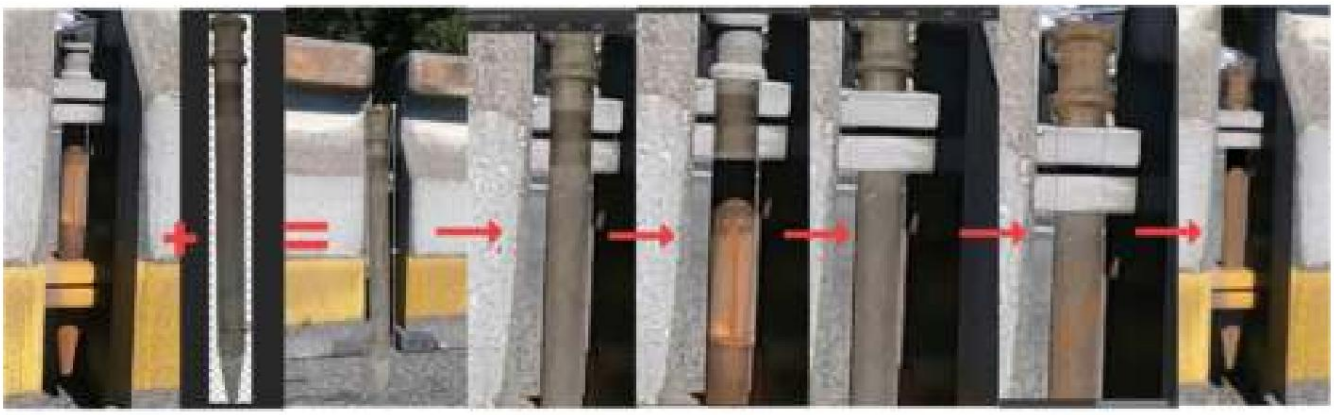


Abbildung 6. Generierung synthetischer Daten: Das Bild veranschaulicht die verschiedenen Schritte der Manipulation eines Videobilds zur Erstellung synthetischer Bilder, die einen Metallstift in unsicheren Positionen zeigen. Ausgehend von einem Einzelbild des Metallstifts umfasst der Prozess die Anpassung seiner Ausrichtung, Position und Umgebungskontext, um realistische, unsichere Szenarien zu generieren.

3.2. Schritt 2: Datenerweiterung

Die Datenerweiterung war notwendig, um mit spärlichen und unausgewogenen Datensätzen arbeiten zu können. Wir haben auch häufig verwendete Ansätze in Betracht gezogen, um die Trainingsrobustheit neuronaler Netzwerke zu verbessern, indem zusätzliche Trainingsdaten aus vorhandenen Datensätzen generiert und so Überanpassungen vermieden werden [60]. Die angewandten Erweiterungstechniken umfassten geometrische und affine Transformationen, bei denen Rotation, Größenänderung, Spiegelung, Verschiebung und Scherung verwendet wurden, um die Bildstruktur zu ändern, ohne den Inhalt zu verändern [37] (Abbildung 7). Die Transformationen wurden mithilfe der Image Processing Toolbox von MATLAB implementiert, die die effiziente Anwendung solcher Techniken erleichterte. Darüber hinaus wurden in diese Forschung Methoden einbezogen, um realistische Variationen in den Datensatz einzubringen, indem Rausch- und Unschärfeeffekte hinzugefügt wurden , insbesondere mithilfe von Gaußschen und Salt-and-Pepper-Rauschmustern. Solche Effekte werden mithilfe der Funktion imnoise von MATLAB zur Rauschzugabe und imgaussfilt für Gaußsche Unschärfe angewendet. Solche Änderungen simulierten potenzielle reale Unvollkommenheiten in den Daten und halfen dem Netzwerk dabei, zu lernen, mit solchen Unregelmäßigkeiten effektiv umzugehen.



Abbildung 7. Illustrative Ergebnisse von Datenerweiterungstechniken zeigen die Bandbreite der Bildtransformationen, die zur Verbesserung des Modelltrainings angewendet werden. Dazu gehören geometrische Transformationen wie FFI-Rotation und -Skalierung, affine Transformationen mit verschiedenen Translationen und Scherungen sowie visuelle Effekte wie Gaußsche Unschärfe, Rauschaddition und anspruchsvolle Farbanpassungen. Die Modifikationen sind entscheidend, um das neuronale Netzwerk auf die vielfältigen und realistischen Szenarien vorzubereiten, die in praktischen Anwendungen auftreten, Xu, Yoon [61].

Zu den weiteren Innovationen gehört die Einführung moderner Farbtransformationstechniken . Mithilfe der Funktion jitterColorHSV von MATLAB wurden die Bilder zufällig angepasst.

Machine Translated by Google

änderungen in Helligkeit, Kontrast, Farbton und Sättigung, wodurch die Palette der visuellen Daten erweitert wird dem das Modell während des Trainings ausgesetzt war. Zusätzlich wurde eine neuartige Farbtransformation Ein Ansatz, der Tabellensuchmethoden und 3D-Farbrauminterpolation kombiniert, ermöglichte für hochwertige Farbverarbeitung in Echtzeit.

Der umfassende Ansatz zur Datenerweiterung diversifiziert den Trainingssatz und verbessert die Generalisierungsfähigkeiten des Modells, die für ähnliche Anwendungen wichtig sind. einschließlich Smart-City-Kontexten und der Anwendung von Technologien zur Verbesserung der Benutzerfreundlichkeit von Räume, in denen menschliche Aktivitäten stattfinden können.

3.3. Schritt 3: Datenverteilungsanalyse nach der Förderung von Minderheiten

Synthetische Frames glichen den Datensatz aus und verbesserten die Klassifizierungsgenauigkeit auf einer relativ kleiner, manuell beschrifteter Datensatz (Abbildung 8). Dieser Ansatz war vorteilhaft unter restriktive Bedingungen, die die Erhebung von Daten zu Minderheitenklassen einschränken. Darüber hinaus erleichterte Feinabstimmung von Modellleistungsmetriken wie Präzision und Rückruf, mit Fokus auf auf die Minimierung falsch-negativer Ergebnisse aufgrund ihrer kritischen Bedeutung für die Sicherheit.

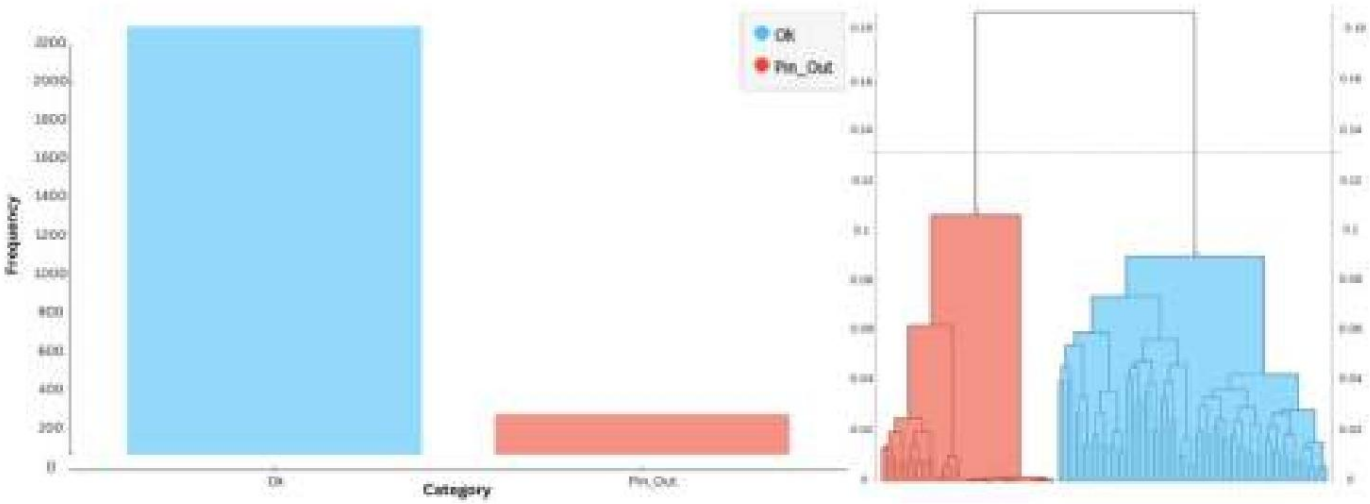


Abbildung 8. Das Dendrogramm, das aus den im Balkendiagramm dargestellten Daten abgeleitet wurde, zeigt zwei Cluster (Pin_OK (rot) und Pin_Out (blau)) und eine visuelle Trennung der generierten mehrdimensionalen Merkmalsraum.

Aus etwa 30.000 Bildern wurden 2300 Bilder ausgewählt, die sichere Pin-Positionen zeigten. kuratiert. Um die Herausforderung des begrenzten Vorkommens unsicherer Pin-Zustände anzugehen – ein häufiges Problem beim Training robuster Erkennungssysteme – 210 synthetische Bilder simulieren Es wurden unsichere Pin-Positionen generiert und dem Datensatz hinzugefügt (Tabelle 3).

Tabelle 3. Datenverteilung: Übersicht über die Anzahl der für das Training verwendeten Pin_OK- und Pin_Out-Bilder und Validierung über verschiedene Videoaufzeichnungssitzungen hinweg.

Schulungsvideo Sitzungsnummer (Tabelle 1)	Trainingsdaten		Validierungsdaten	
	Pin_OK Bilder	Pinbelegung Bilder	Pin_OK Bilder	Pinbelegung Bilder
1	155	32	40	5
2	725	80	180	20
2	800	40	200	10
3	400	16	100	5

ÿ ÿ

Machine Translated by Google

Durch die Schaffung unterschiedlicher Pin-Out-Positionen konnte die Balance zwischen falsch-positive und falsch-negative Ergebnisse, was für die Verbesserung der Genauigkeit und Robustheit des Modells von entscheidender Bedeutung ist (Tabelle 4).

Tabelle 4. Erste Klassifizierungsergebnisse aus Datenclustern aus Abbildung 8. Übernommen von Baÿci´c, [7] .

Verwirrungstabelle				
Modell			**Tatsächlich**	
			PIN_OK	**PIN_OUT**
Logistische Regression	**PRtD**	**PIN_OK**	**[98,5 %**	**1,5 %**
		PIN_OUT	**0**	**100%]**
Neurales Netzwerk	**DaCTtD**	**PIN_OK**	**[98,9 %**	**1,1 %**
		PIN_OUT	**0**	**100%]**
SVM		**PIN_OK**	**[98,1 %**	**1,9 %**
		PIN_OUT	**0**	**100%]**

Synthetische Frames mit unterschiedlichen Pin_Out-Positionen verbesserten die Feinabstimmung von False-Positive- (FP) und Falsch-Negativ-Verhältnisse (FN). Ziel ist es, FP > FN mithilfe eines Datensatzes mit Grenzbereiche unsicherer Pinpositionen, was die Klassifizierungsgenauigkeit verbessert. Die Genauigkeit wurde gemessen mit der folgenden Formel:

$$\text{Genauigkeit} = \frac{TP + TN}{TP + FP + TN + FN} \quad (8)$$

TP repräsentiert die wahren Positiven oder korrekt identifizierten Pin_Out-Frames und TN repräsentiert die wahren Negativen oder korrekt identifizierten Pin_OK-Frames. Präzision, das Verhältnis der korrekt vorhergesagten positiven Beobachtungen zur Gesamtzahl der vorhergesagten positiven Beobachtungen ist definiert als wie folgt:

$$\text{Genauigkeit} = \frac{TP}{TP + FP} \quad (9)$$

Der umfassende Ansatz umfasste eine sorgfältige dreifach geschichtete Kreuzvalidierung Prozess, wie in Tabelle 5 beschrieben.

Tabelle 5. Die Kreuzvalidierungstests und Bewertungsergebnisse wurden seit der ersten Untersuchung aktualisiert.

3-fach geschichtete Kreuzvalidierung		
Modell	**Präzision**	**Abrufen**
Parameter der logistischen Regression: Regularisierung: Ridge (L2), C = 1	**0,995**	**0,995**
Mehrschichtiges Perzeptron (MLP)-Parameter: Versteckte Schichten: 2 Neuronen: Aktivierungsfunktion: ReLu Löser: Adam Alpha: 0,02 Max. Iterationen: 200 Backpropagation-Algorithmus	**0,995**	**0,995**
Parameter der Support Vector Machine (SVM): C = 1,0, ÿ = 0,1 Kernel: Linear Numerische Toleranz: 0,001 Max. Iteration: 100	**0,985**	**0,984**

Machine Translated by Google

3.4. Schritt 4: Region of Interest (ROI) ermitteln

Unser Forschungsteam untersuchte **bei der Entwicklung eines automatisierten Pin-Erkennungssystems für Verkehrsmanagementanwendungen umfassend verschiedene Methoden zur Erkennung von Interessenbereichen (ROI) . Jede Methode wurde auf ihre Fähigkeit hin untersucht, Pin- Standorte in Videobildern unter verschiedenen Umgebungsbedingungen genau zu identifizieren. Das Durchgehen verschiedener Methoden verdeutlichte die Herausforderungen und das Potenzial jedes Ansatzes und führte schließlich zu einer besonders effektiven Lösung. Zunächst nutzten wir bei unserer Untersuchung Methoden zur Region-Proposal-Ermittlung, die Begrenzungsrahmen um potenzielle ROIs Obwohl der Ansatz zunächst vielversprechend war, waren häufige manuelle Anpassungen erforderlich, um eine präzise Erkennung und Zählung zu gewährleisten. Dies machte ihn für dynamische Anwendungen in der realen Welt, bei denen Automatisierung von größter Bedeutung ist, weniger praktikabel . Für die adaptive Hintergrundmodellierung wurde ein Gaussian Mixture Model (GMM) eingesetzt** (Abbildung **9B). Das GMM war gut darin, sich bewegende Pins von statischen Hintergründen zu segmentieren, indem es die Wahrscheinlichkeitsdichteverteilung der Pixelintensitäten modellierte. Trotz seiner Wirksamkeit bei der Vordergrunddifferenzierung erforderte das GMM eine umfangreiche Nachbearbeitung, um jeden Pin genau zu lokalisieren, was seinen eigenständigen Nutzen bei der präzisen ROI-Erkennung einschränkte. Wir untersuchten auch die farbbasierte Segmentierung mithilfe von K-Means-Clustering innerhalb der Farbräume RGB, LAB und HSV – letzteres wird besonders bevorzugt, da es widerstandsfähiger gegen Lichtschwankungen ist (Abbildung 9A). Die Methode nutzte die einzigartigen Farbsignaturen der Pins, war jedoch hinsichtlich der Spezifität eingeschränkt und erfass**

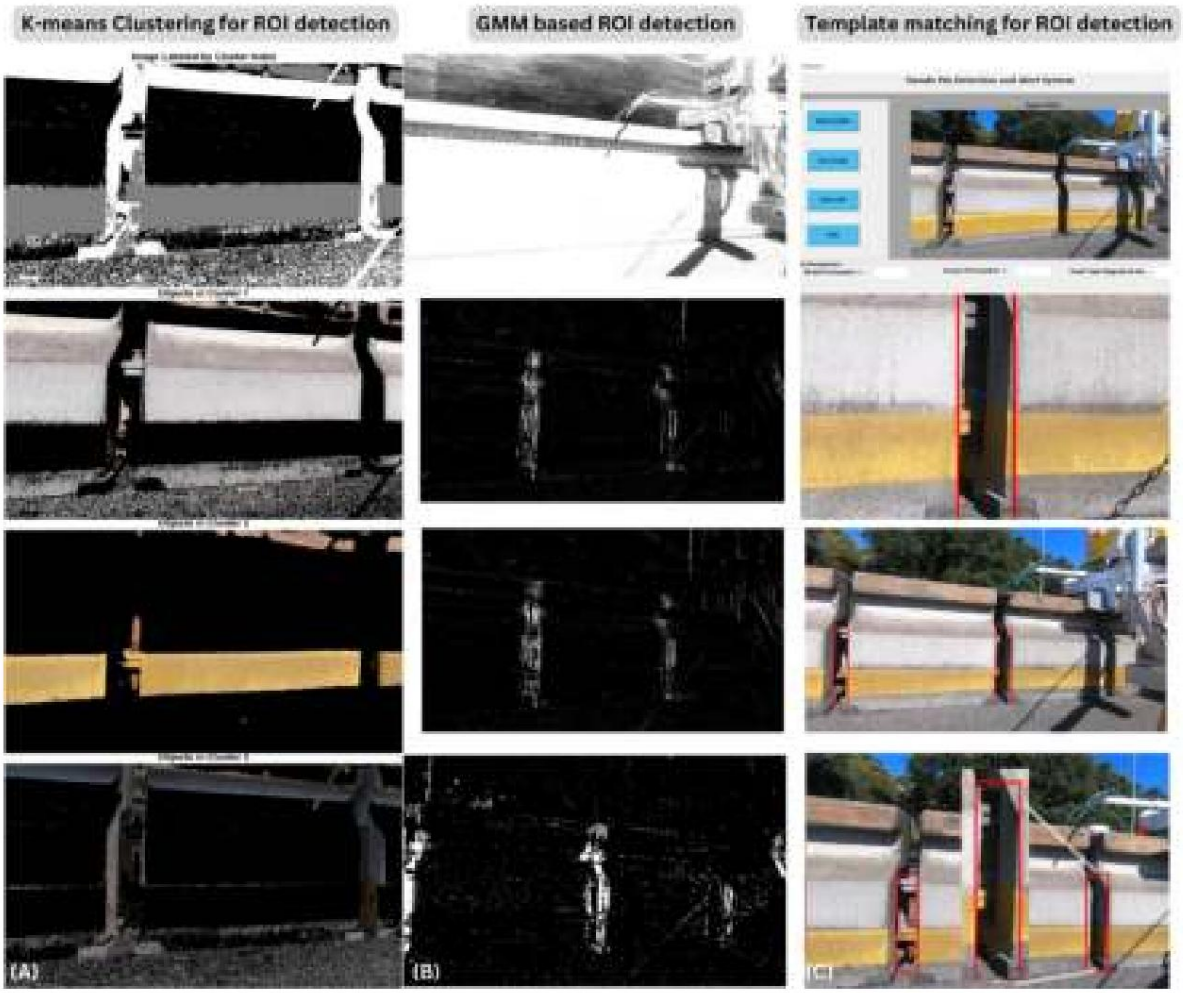


Abbildung 9. Vergleichende Analyse vielversprechender ROI-Erkennungstechniken bei der automatischen Pin-Erkennung: (1) Farbbasierte Segmentierung mittels K-Means-Clustering im HSV-Farbraum, wobei die einzigarti

Machine Translated by Google

Farbsignaturen von Pins bei gleichzeitiger Bewältigung von Herausforderungen hinsichtlich der Spezifität. (2) Die auf dem Gaussian Mixture Model (GMM) basierende Erkennung veranschaulicht die Vordergrund-Hintergrund-Segmentierung und die Verfolgung der Pin-Bewegung zwischen Betonblöcken. (3) Regionprops-Anwendung, die Template-Matching und automatische Bounding-Box-Erkennung für eine präzise Lokalisierung und weniger manuelle Eingriffe bei der Erkennung und Beschriftung des Pin-ROI vorführt.

Die in Abbildung **9** **getesteten Methoden führten dazu, dass wir die** Regionprops **-Funktion von MATLAB übernahmen, die sich als überlegen erwies, wenn es darum ging, die Präzisions- und Automatisierungsprobleme früherer Methoden zu lösen. Die** Regionprops **analysiert jede verbundene Komponente in binären Videobildern und quantifiziert effektiv die Flächen- und Begrenzungsrahmenkoordinaten für jeden erkannten ROI. Die Funktion verbesserte die Genauigkeit der Pin-Erkennung und ermöglichte eine erhebliche Reduzierung manueller Eingriffe durch Automatisierung d**

Die zugrunde liegenden mathematischen Konzepte der Regionprops- **Funktion sind wesentlich für ihre Leistung (Gleichungen (8)–(13)).**

Fläche: Die Fläche einer Region wird als Anzahl der darin enthaltenen Pixel berechnet.

$$\text{Fläche} = \ddot{y}(i,j)\ddot{y}R\ 1 \tag{10}$$

In dieser Gleichung stellt die Fläche die Gesamtzahl der Pixel in der Region R dar. Jedes Pixel innerhalb der Region trägt einen Wert von 1 zur Gesamtfläche bei, wodurch die Pixel effektiv gezählt werden.

Schwerpunkt: Der Schwerpunkt wird durch Mittelung der Positionen aller Pixel in der Region wie folgt bestimmt:

$$\text{Schwerpunktx} = \frac{1}{N}\ \ddot{y}(i,j)\ddot{y}R\ i,\ \text{Schwerpunkt} = \frac{1}{N}\ \ddot{y}(i,j)\ddot{y}R\ j \tag{11}$$

Begrenzungsrahmen: Der Begrenzungsrahmen, der die Region umschließt, wird durch die minimalen und maximalen x- und y-Koordinaten der Pixel definiert.

$$\text{Begrenzungsbox} = [\text{xmin, ymin, xmax} \ddot{y} \text{xmin} + 1, \text{ymax} \ddot{y} \text{ymin} + 1] \tag{12}$$

Länge der Haupt- und Nebenachse: Für Regionen, die durch Ellipsen approximiert werden, werden die Längen der Haupt- und Nebenachse aus den Eigenwerten der Kovarianzmatrix des Pixels abgeleite Koordinaten.

$$\text{Länge der Hauptachse} = 2\ \frac{\overline{\ddot{y}1}}{N},\ \text{Nebenachsenlänge} = 2\ \frac{\overline{\ddot{y}2}}{N} \tag{13}$$

Orientierung: Die Orientierung einer solchen Ellipse, die den Winkel zwischen der x-Achse angibt **und der Hauptachse wird unter Verwendung der zweiten Momente der Region berechnet.**

$$\text{Orientierung} = \frac{1}{2}\arctan\left(\frac{2\mu 11}{\mu 20 \ddot{y} \mu 02}\right) \tag{14}$$

Umfang: Der Umfang wird durch die Summierung der Abstände zwischen aufeinanderfolgenden Grenzpixeln gemessen.

$$\text{Umfang} = \ddot{y}\text{Grenzpixel Abstand zwischen aufeinanderfolgenden Grenzpixeln} \tag{15}$$

Die Integration von Regionprops **in unsere Methode verbessert die Automatisierung der Pin-Erkennung erheblich. Die Verbesserung bietet hohe Präzision und Effizienz, die für Verkehrsmanagementsysteme, die Zuverlässigkeit und schnelle Verarbeitung erfordern, von entscheidender B**

Unsere erfolgreiche Umsetzung unterstreicht die Bedeutung eines methodischen Ansatzes in der technischen Forschung und zeigt, wie durch iteratives Testen und die Integration verschiedener Techniken komplexe Probleme der realen Welt gelöst werden können. Diese Studie unterstreicht die Wirksamkeit der Kombination traditioneller Bildverarbeitungsmethoden mit fortschrittlichen Analysetools und setzt einen neuen Standard für die zukünftige Forschung im Bereich automatisierter Verkeh

Machine Translated by Google

Die Ergebnisse bieten Forschern und Praktikern einen soliden Rahmen zur Verbesserung Automatisierte Erkennungssysteme in allen technischen Bereichen.

3.5. Schritt 5: Modelltraining und Validierung

Das Computersystem (Tabelle 6) bot ausreichend Hardwarekapazitäten für das Training der Modell im Kern des entwickelten „Unsafe Pin Detection and Alert System“-Systems.

Tabelle 6. Ein Entwicklungssystem mit paralleler Verarbeitungsarchitektur der NVIDIA GPU.

Systemkonfiguration	
Prozessor	**Intel Core i7 Prozessor**
Erinnerung	**32 GB RAM**
Festplatte	**512 GB Solid-State-Laufwerk**
Grafik	**NVIDIA GeForce RTX2070 Super 8 GB GFX**
Betriebssystem	**Windows 10**

ResNet-50, erweitert mit 3D-Faltungen im STENet-Framework, wurde nach einem umfangreichen Vergleich mit anderen Modellen wie Squeezenet, GoogLeNet ausgewählt. InceptionV3 und Mobilenetv2. ResNet-50 wird aufgrund seiner optimalen Balance zwischen Rechenleistung und Leistung, bestätigt durch zufriedenstellende GPU-RAM-Bewertungen während erster Experimente (Tabelle 7).

Tabelle 7. Vergleich vortrainierter tiefer neuronaler Netzwerke basierend auf der Auflösung der Eingabebilder, Anzahl von Parametern, Tiefe und Modellgröße.

Vortrainierte tiefe neuronale Netzwerke				
Modell	**Eingabebild Auflösung**	**Parameter (1.000.000)**	**Tiefe**	**Größe**
AlexNet	**227 × 227**	**61**		**227**
SqueezeNet	**227 × 227**	**1,24**	**8**	**4.6**
GoogleNet	**224 × 224**	**7**	**18 22**	**27**
Einführung v3	**299 × 299**	**23,9**	**23,9**	**48**
MobileNet v2	**224 × 224**	**3,5**	**3,6**	**53**
Resnet 50	**224 × 224**	**25,6**	**50**	**96**

Verschiedene Umweltelemente, wie rostfarbenes Laub, passten gut zum Aussehen von verrosteten Metallstiften, was die visuelle Unterscheidung erschwert. Ebenso unverrostete Metallstifte hatten die gleiche Farbe wie die Straßenoberfläche und die Reifen vorbeifahrender Fahrzeuge, was den Identifizierungsprozess weiter erschwert. Die Metallbarrieren im Hintergrund hatten oft Formen, die dem horizontalen Profil der Metallstifte ähnelten, während die dynamische Farben vorbeifahrender Fahrzeuge sorgten für intensive Hintergrundvariationen. Darüber hinaus Neben der Beschriftung der Pin-Status wurden für jedes Bild Bereiche von Interesse (ROIs) definiert mit die Regionprops- **Methode zur Unterscheidung zwischen relevanten Merkmalen und potenziellem Hintergrund Rauschen. Diese Methode, die jede verbundene Komponente in binären Videobildern analysiert zur Quantifizierung der Fläche und der Begrenzungsrahmenkoordinaten für jeden erkannten ROI, deutlich verbesserte Genauigkeit der Pin-Erkennung und reduzierte manuelle Eingriffe. Die kombinierte Die Berücksichtigung von Form- und Farbhomogenität war ein wesentlicher Faktor bei der Ausbildung des AHB Pin-Erkennungsmodell (Abbildung 10).**

Darüber hinaus sind bestimmte Teile des Systems der beweglichen Betonbarrieren (MCB) eng ähnelten der Form der Metallstifte, was zu weiterer Verwirrung führte (Abbildung 11(3)). Diese Faktoren machten die Erkennungsaufgabe anspruchsvoll, da das System unterscheiden musste zwischen mehrere visuell ähnliche Elemente unter verschiedenen Bedingungen. Der Datensatz wurde dann eine Reihe von Bildvergrößerungsprozessen zur Simulation verschiedener Betriebsbedingungen, nicht durch die erste Videoaufnahme abgedeckt. Die Prozesse umfassten geometrische Transformationen

Machine Translated by Google

wie Drehungen, Skalierungen, Verschiebungen und das Hinzufügen von Bildstörungen wie Gaußscher Unschärfe und Rauschen, die für die gute Leistung der Trainingsmodelle unter praktischen Einsatzbedingungen entscheidend sind.



Abbildung 10. Darstellung der Hintergrundherausforderungen bei der ROI-Erkennung.

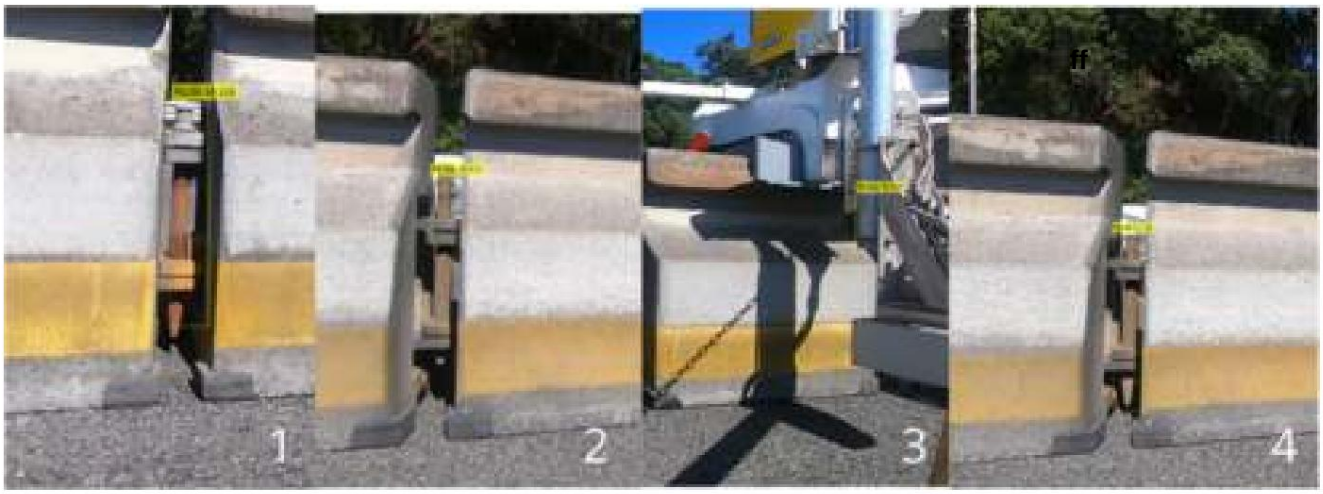


Abbildung 11. Beispiele für die Pin-Erkennung gemischt mit der Erkennung falsch-positiver Ergebnisse (3) veranschaulichen ein hohes Maß an Hintergrundrauschen, das den Merkmalen der ROI ähnelt.

Eine kritische Komponente des Trainingsprozesses war die Anpassung der Lernrate, ein Parameter, der die Konvergenz und die endgültige Leistung des Modells erheblich

Machine Translated by Google

Es wurden Lernraten von 0,01 bis 0,00001 mit ResNet-50 getestet (Abbildung 12). Die optimale Rate wurde durch eine Reihe von Versuchen ermittelt, bei denen die Modellgenauigkeit und die Verlustmetriken ausgewertet wurden. Die Experimente nutzten ROC-Kurven, um die Kompromisse zwischen True-Positive- und False-Positive-Raten bei verschiedenen Schwellenwerteinstellungen visuell darzustellen , was eine fundierte Auswahl des Modells mit der besten Leistung unter den gegebenen Trainingsbedingungen ermöglichte. Die Ergebnisse zeigten durchweg, dass eine Lernrate von 0,0001 das beste Gleichgewicht zwischen Trainingsgeschwindigkeit und Modellgenauigkeit bot.

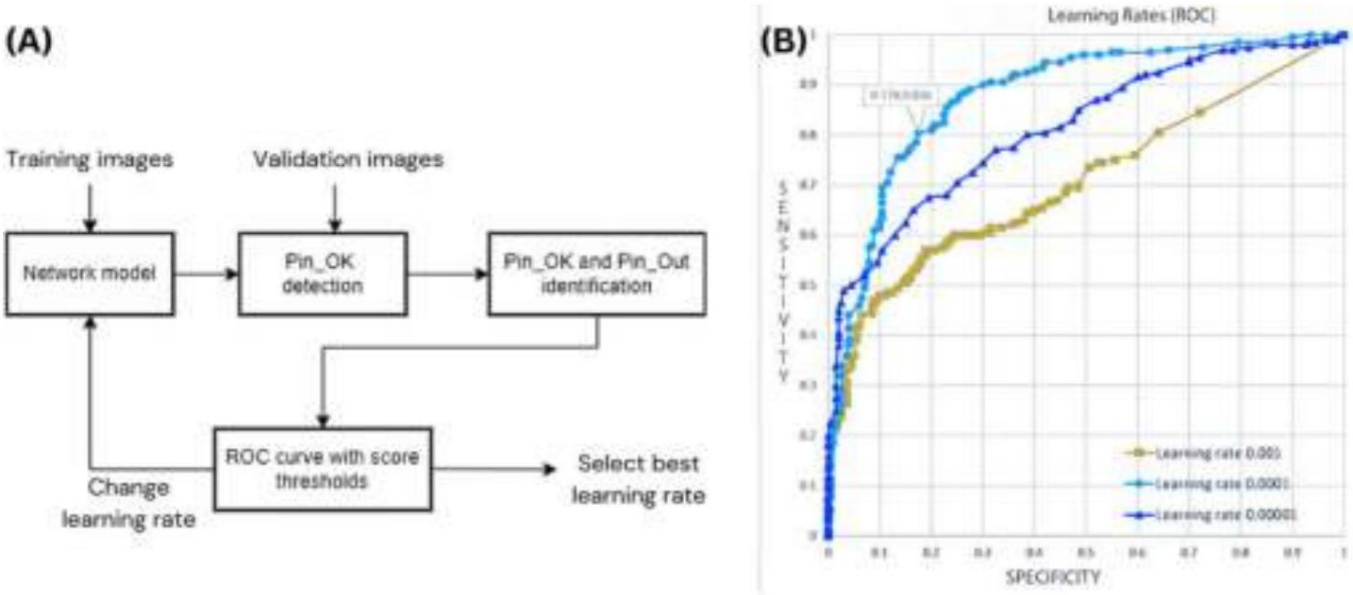


Abbildung 12. Der Trainingsprozess zum Finden des besten Netzwerkmodells basierend auf der Lernrate: (A) Ablauf des ff- Trainingsprozesses und (B) ROC-Kurven für verschiedene Lernraten, die die Leistung der Lernraten zeigen . Die Lernrate von 0,0001 wurde gewählt, da sie ein Gleichgewicht zwischen Trainingsgeschwindigkeit und Genauigkeit herstellt, eine bessere Leistung bietet und gleichzeitig den hohen Rechenaufwand und -zeitaufwand vermeidet.

Zusammenfassend lässt sich sagen, dass diese Forschung die Auswahl, Vorbereitung und das Training eines Deep- Learning-Modells detailliert beschreibt, das in der Lage ist, den Status von Metallstiften in beweglichen Betonbarrieren zu erkennen und zu klassifizieren . Der methodische Ansatz, der die Erstellung eines umfassenden und vielfältigen Trainingsdatensatzes durch reale und synthetische Bilder und strenge Modelltrainingsprotokolle betont , **stellt eine skalierbare und effiziente Lösung zur Verbesserung der Sicherheit der öffentlichen Infrastruktur durch fortschrittliche KI-Techniken dar.**

3.6. Schritt 6: Modellanalyse und Test

Die STENet-Architektur, die vom robusten Framework von ResNet-50 inspiriert ist und auf diesem aufbaut, bietet erhebliches Potenzial für Aufgaben zur räumlich-zeitlichen Anomalieerkennung, wie sie beispielsweise für die Überwachung der Auckland Harbour Bridge (AHB) erforderlich sind. Obwohl ResNet-50 ursprünglich für die Bildklassifizierung entwickelt wurde, ist es aufgrund seiner starken räumlichen Merkmalsextraktionsfunktionen eine solide Grundlage für weitere Verbesserungen bei der Analyse vo STENet integriert Komponenten von ResNet-50 mit 3D-Faltungen, um räumliche und zeitliche Informationen zu erfassen und so die direkte Analyse von Bewegung und dynamischen Änderungen innerhalb von Videosequenzen zu ermöglichen. Weitere Verbesserungen umfassen die Nutzung der räumlichen Merkmalsextraktionsfunktionen und rekurrenter neuronaler Netzwerke (RNNs) von ResNet-50, um zeitliche Sequenzen zu verfolgen , oder den Einsatz zeitlicher Pooling-Methoden, um Videosegm

Durch die Anwendung von Strategien zur Modellvereinfachung, Edge Computing und inkrementellem Lernen verbessert STENet die Systemleistung und die Reaktionsfähigkeit auf neue Anomalien. Dadurch eignet sich STENet besonders für Anwendungen zur kontinuierlichen Überwachung und Verkehrsüberwachung . Tabelle 8 zeigt die Leistungskennzahlen von STENet im Vergl Durch die Integration von Anpassungen der Kernfunktionen von ResNet-50 erkennt STENet effektiv

Machine Translated by Google

Abweichungen im Erscheinungsbild und Verhaltensmuster im Laufe der Zeit, wodurch ein robuster Rahmen geschaffen wird zur Echtzeit-Anomalieerkennung in unterschiedlichen Betriebsumgebungen (Abbildung 13).

Tabelle 8. Das Spatio-Temporal Enhanced Network (STENet) erreicht eine hervorragende Genauigkeit von 95,2 % Score und ein F1-Score von 94,8 %, was seine außergewöhnliche Fähigkeit zeigt, die **Herausforderungen von Hintergrundrauschen und kleinen ROIs zu meistern ,** mit einem bemerkenswerten ROC-AUC von 98,5 %, was seine **Robustheit bei der Klassendifferenzierung bei räumlich-zeitlichen Aufgaben.**

Modell	Genauigkeit	Präzision	Abrufen	F1-Ergebnis	ROC-AUC	Ausbildung Zeit	Inferenz Zeit	Anzahl der Parameter
STENet	95,2 %	94,5 %	95,2 %	94,8 %	98,5 %	4 h	80 ms	25 M
VGG-19	88,0 %	88,5 %	89,0 %	88,7 %	94,0 %	6,5 h	90 ms	142 M
ResNet-50	90,0 %	90,3 %	91,0 %	90,6 %	95,8 %	4 h	70 ms	25,6 M
InceptionV3	91,0 %	91,2 %	91,5 %	91,3 %	96,0 %	3,5 h	65 ms	23,8 M
SqueezeNet	85,0 %	85,5 %	86,0 %	85,7 %	92,0 %	2,5 h	35 ms	1,25 M

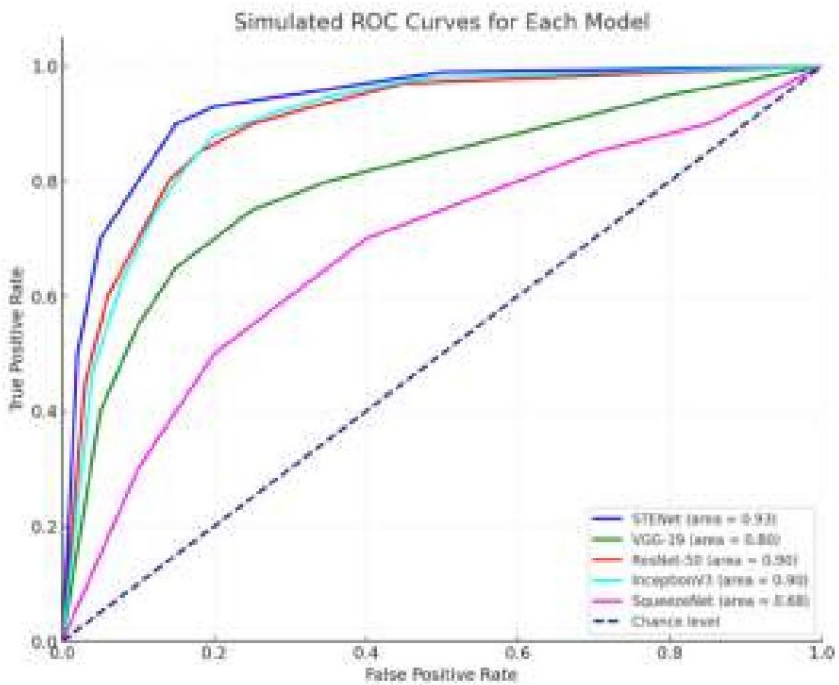


Abbildung 13. ROC-Kurvenvergleich von Tabelle 8.

4. Ergebnisse und Diskussion

Die Wirksamkeit des Pin-Erkennungs- und Warnsystems, das sowohl Pin-Regionen umfasst of Interest (ROI)-Erkennung und -Verfolgung neben Pin-Statuserkennung und Alarmfunktionen , wurde kritisch bewertet. Die Bewertungen basierten auf zuvor beschriebenen Me Besonderes Augenmerk gilt den Betriebsbedingungen, die die Systemleistung beeinflussen.

4.1. Vorbereitungen, Datenaufzeichnungsprotokoll, Ergebnisse und Erkenntnisse

Die Leistung des Systems wurde durch die Analyse aufgezeichneter Videos Bild für Bild beurteilt. Rahmen, der sich auf die Verfolgung des ROI der Pins konzentriert. Mehrere Umwelt- und Betriebsfaktoren wurden identifiziert, die die Ergebnisse negativ beeinflussen könnten.

- **Lichtverhältnisse: Bewölkte Wetterbedingungen behinderten die Sicht erheblich, was die Erkennung von Stiften zwischen Betonblöcken erschwerte und die Erkennung und Zählvorgänge unzuverlässig.**

Machine Translated by Google

• Hintergrundbewegungen: Bewegte Fahrzeuge führten zu erheblichem Rauschen im Hintergrund , was die Erkennungsgenauigkeit erheblich beeinträchtigte. Während der gräuliche Farbton der Straße und der Betonblöcke dazu beitrug, den Hintergrund zu vereinfachen, störten die vorbeifahrenden großen Fahrzeuge die visuelle Klarheit. Darüber hinaus verschmolzen die Rostfarben und das Gelb der beweglichen Betonbarrieren (MCBs) und Fragmente zerbrochener Betonbarrieren und Metallbarrieren oft mit den Pin-Regionen von Interesse (ROIs), was die Tracking-Genauigkeit weiter verringerte. • Schatten und Beleuchtung: Schatten, die auf Pin-ROIs geworfen wurden, wurden manchmal als bewegte Objekte fehlinterpretiert. Dunklere Schatten könnten die Erkennungs- und Trackingsysteme beeinträchtigen , insbesondere bei hellem Sonnenlicht. • Vibrationsverzerrungen: Betriebsgeschwindigkeiten über 6 km/h verursachen übermäßige Vibrationen, wodurch die Bilder unscharf werden und Hintergrundgeräusche entstehen, was die Erkennungsprozesse weiter erschwert.

Um wetterbedingte Schwankungen zu mildern, wurden mehrere bei sonnigem Wetter aufgenommene Videos analysiert. Das erste Video, das ein breiteres Feld vom Vorderarm der Barrier Transfer Machine (BTM) aus erfasste, zeigte ungefähr zwei ROIs pro Frame, aufgenommen an einem klaren Nachmittag. Das zweite Video, das vom Hinterarm der BTM aus aufgenommen wurde, zeigte ein engeres Sichtfeld mit manchmal nur teilweiser Sichtbarkeit eines ROI mit einem einzelnen

4.2. Synthetische Rahmen erstellen

Das Konzept der Erstellung synthetischer Frames stieß auf erhebliche Herausforderungen. Die anfänglichen Klontechniken waren unzureichend und führten zu einem nicht lebensfähigen Minderheitsdatensatz, der eine Umstellung von Klassifizierungstechniken auf traditionelle mathematische Ansa Anschließende Konsultationen mit Grafikexperten führten zur Einführung fortgeschrittener Methoden mit Photoshop und Gimp zur Erzeugung synthetischer Rahmen (Abbildung 14).

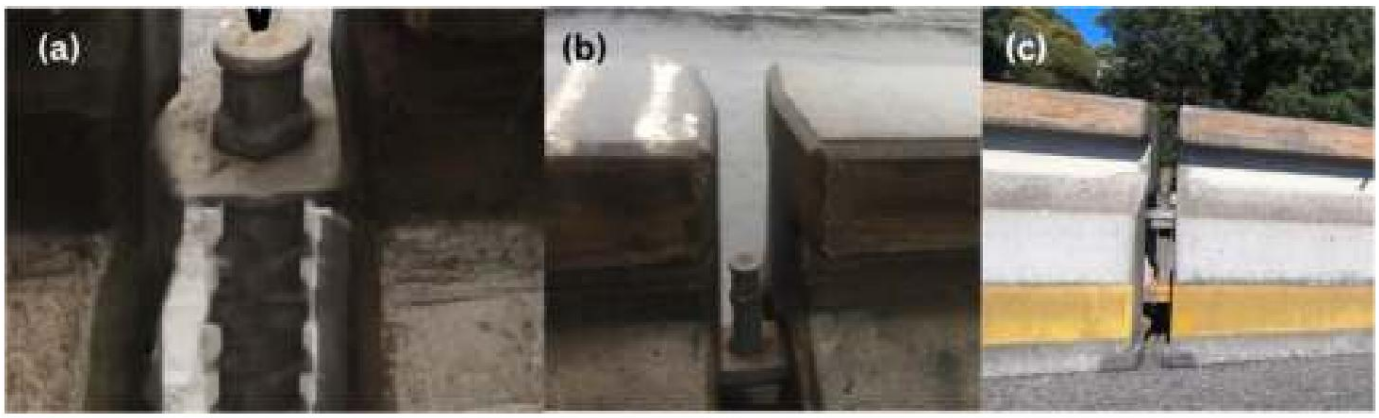


Abbildung 14. Klonvorgang zum Erstellen synthetischer Rahmen, die die Position „PIN OUT“ darstellen. (a) Fehlgeschlagene Versuche bis zum Erfolg (b,c); zwei unterschiedliche Winkel: (b) Kamera am hinteren Arm der Barrier Transfer Machine (BTM) montiert und (c) Kamera am vorderen Arm der BTM montiert, um die Modelliergenauigkeit zu verbessern.

Während der Proof-of-Concept-Phase wurden etwa 40 synthetische Frames aus Videomaterial erstellt, das mit einem Mobiltelefon aufgenommen wurde. Ein zweiter Stapel von über 200 synthetischen Frames wurde später aus Videoaufnahmen höherer Qualität erstellt, was eine deutliche Verbesserung der Frame-Klarheit zwischen der ersten und der späteren Produktion zeigt. Angesichts des arbeitsintensiven Charakters der manuellen Frame-Erstellung wurden Anstrengungen unternommen, den Prozess mithilfe des Aktionsfelds von Photoshop zu automatisieren. Die Komplexität der Hintergrundelemente verhinderte jedoch eine vollständige Automatisierung. Obwohl MATLAB Pixel-Klontechniken bietet, sind weitere Forschungsarbeiten erforderlich, um eine vollständig automatisierte und robuste Methode zur Generierung synthetischer Frames zu entwickeln. Die Analyse unterstreicht die kritischen Abhängigkeiten der Umgebungsbedingungen von der Leistung von Pin-Er

Machine Translated by Google

technische Verbesserungen bei der Generierung synthetischer Daten zur Unterstützung eines robusten Systemtrainings und einer Validierung.

4.3. Systemprototyp: Pin-Tracking-, Zähl- und Warnfunktion **Die Entwicklung des**

Metallstift-Erkennungs- und Warnsystems basierte auf der Anforderung, dass dem Endbenutzer, der als technisch nicht versiert gilt, eine vereinfachte Schnittstelle für eine effiziente Systembedienung zur Verfügung steht. Der Systemprototyp mit der Bezeichnung „Unsafe Pin Detection and Alert System“ (Abbildung 15) wurde mit dem MATLAB App Designer [62] entwickelt und implementiert , wobei Backend-Operationen mit einer grafischen Benutzeroberfläche (GUI) integriert wurden.



Abbildung 15. Grafische Benutzeroberfläche für „Unsafe Pin Detection and Alert System“.

Systemfunktionen Die App

ermöglicht es dem Endbenutzer, ein Video oder eine Bildsequenz zur Analyse hochzuladen. Beim Hochladen nutzt die App ein auf Deep Learning basierendes Pin-Status-Erkennungsnetzwerk, um den Video-Feed zu analysieren. Das System zählt die Metallstifte, und wenn es einen Stift in einer unsicheren Position erkennt, löst es einen Alarm aus und zeigt den spezifischen Metallstift-ROI auf dem Bildschirm an.
Die zweite Funktion unterstützt die Live-Feed-Analyse, bei der der Benutzer die App mit einem Live-Kamera-Feed verbinden kann. Das Pin-Erkennungsnetzwerk wird aktiviert, um eingehende Frames in Echtzeit zu analysieren, während der Live-Feed an die Überwachungs-App des Unsafe Pin Detection and Alert System übertragen wird (Abbildung 15). Wie bei der voraufgezeichneten Analyse verfolgt und zählt das System die Metallstift -ROIs, warnt den Benutzer und zeigt die ROI-Nummer an, wenn ein unsicherer Pin erkannt wird.
Die Benutzeroberfläche des Metallstift-Erkennungs- und Warnsystems (siehe Abbildung 15) ist intuitiv und leicht zu navigieren, sodass Benutzer das System auch ohne technische Vorkenntnisse bedienen können. Das Design und die Betriebslogik der App wurden speziell auf die Anforderungen der NZTA zugeschnitten und ermöglichen individuelle Änderungen, um bestimmten Betriebsanforderungen oder Aktualisierungen gerecht zu werden.

Abschließend lässt sich sagen, dass die App des Warnsystems theoretisch einsatzbereit ist. Sofern Interesse besteht und **die NZTA weitere Unterstützung bietet,** besteht **die Möglichkeit, sie zu einem vollständigen Produkteinsatz weiterzuentwickeln .** Zukünftige Arbeiten werden sich auf die Verfeinerung der Funktionalität der App , **die Verbesserung ihrer Leistung und die Sicherstellung ihrer Installation vor Ort und Integration in vorhandene Infrastruktursysteme konzentrieren. Solche Verbesserungen hängen jedoch von der Verfügbarkeit weiterer Ressourcen und dem anhaltenden Interesse der Beteiligten ab.**

Machine Translated by Google

4.4. Ergebnisse

Die Bewertung der Leistung des STENet-Modells zielte darauf ab,
Finden und Lokalisieren des Pin_Out-Status mithilfe der beschriebenen Workstation-Spezifikationen
Wir haben die Modelle auf die Region of Interest (ROI) des Pins angewendet, indem wir sie Bild für
Bild verfolgten und in Echtzeit zählten. Dabei haben wir jedem Pin-ROI einen Index zugewiesen, um
Kontinuität über alle Frames hinweg. Die Detektormodellkonfigurationen sind in Tabelle 9 zusammengefasst.
die Parameter wie Batchgröße, Anzahl der Epochen, Lernrate und Optimierer detailliert beschreibt
Auswahlmöglichkeiten. Zwei Optimierungstechniken wurden eingesetzt, Stochastic Gradient Descent mit
Momentum (SGDM) und Adam, beide ausgewählt aufgrund ihrer Fähigkeit, die Modellleistung zu verbessern
durch effiziente Orientierung in der Optimierungslandschaft [63,64].

Tabelle 9. Konfiguration des Detektormodells. Die Tabelle listet die Batchgröße, die Anzahl der Epochen, das Lernen
Rate und Optimierer, der zum Trainieren des Detektormodells verwendet wird. Zwei verschiedene Optimierungstechniken,
Stochastischer Gradientenabstieg mit Momentum (SGDM) und adaptiver Momentschätzung (Adam),
wurden eingesetzt, um die Modellleistung zu verbessern.

Modell	Batchgröße	Epoche	Lernrate	Optimierer
Detektor	64	100	0,0001	Sgdm
	64	100	0,0001	Adam

Die Genauigkeit des Pin-ROI-Trackings wurde durch Vergleich der Konsistenz von Indexwerten bewertet.
Zuordnung von Frame zu Frame. Abweichungen in der Indexkontinuität wurden als Fehler in
Tracking. Der Datensatz wurde in 70 % Trainingsrahmen, 10 % Validierung und 20 % Tests unterteilt
Rahmen, um das Modell unter verschiedenen Bedingungen robust zu bewerten. Der Trainingsprozess mit
SGDM erreichte schnellere Trainingszeiten als Adam, wie in Tabelle 10 gezeigt. Die Validierung
der Modelle wiesen geringere Root Mean Square Error (RMSE) und Validierungsverluste auf
mit SGDM, was auf einen effizienteren Optimierungspfad hinweist.

Tabelle 10. SGDM-Trainings- und Validierungsprozess.

Epoche	Wiederholung	Zeit Verstrichen	Mini-Charge (RMSE)	Validierung (RMSE)	Minibatch Verlust	Validierung
10	150	15.00.05 Uhr	0,91	0,87	0,8285	0,7910
20	300	30.00.20	0,74	0,83	0,5384	0,6761
30	450	00.45.12	0,66	0,78	0,4141	0,6329
40	600	01.00.15	0,62	0,76	0,3968	0,5809
50	750	14.01.50 Uhr	0,56	0,77	0,3182	0,4150
60	900	28.01.58	0,53	0,76	0,2708	0,4060
70	1050	01.44.55	0,52	0,74	0,2408	0,3716
80	1200	01.58.21	0,50	0,76	0,2018	0,3208
90	1350	13.02.43 Uhr	0,48	0,75	0,1808	0,2710
100	1500	27.02.23	0,47	0,74	0,1280	0,2219

Die Ergebnisse der Pin-ROI-Klassifizierung und die Genauigkeit der Boundary-Box-Erkennung durch
Die trainierten STENet-Systeme sind in Abbildung 16 dargestellt. Das System war sehr effektiv bei der Erkennung
sowohl Pin_OK als auch Pin_Out Status, auch wenn Pin ROIs teilweise verdeckt oder eng
positioniert, was traditionell die Erkennungsgenauigkeit in Frage stellt. Die Genauigkeit
verringerte sich in Frames, in denen die Pins zu eng beieinander oder teilweise außerhalb des Feldes standen
der Ansicht.

Machine Translated by Google



Abbildung 16. Die ausgewählten Pin-ROI-Videobilder zeigen generierte Overlays mit Begrenzungsrahmen.

Während das System unter idealen Betrachtungsbedingungen eine stabile Leistung zeigte, Die Erkennungsgenauigkeit variierte je nach Sichtfeld und Lichtverhältnissen, was die Notwendigkeit eines vielfältigeren Trainingsdatensatzes unterstreicht. Die Erstellung synthetischer I Die Erweiterung des Datensatzes wurde untersucht, aber die Automatisierung des Prozesses bleibt eine Herausforderung für zukünftige Arbeiten, da sich die manuelle Rahmenerstellung als zeitintensiv erwies. Tabelle 11 zeigt die Klassifizierungsmetriken – Genauigkeit, Präzision und Rückruf – nach dem Training, wobei die Die insgesamt starke Leistung des Modells.

Tabelle 11. Leistung des STENet, wenn ResNet 50 als Klassifikator verwendet wird.

Klassen	Genauigkeit		Präzision		Abrufen	
	Ausbildung	Validierung	Trainingsvalidierung		Zug	Validierung
Pin_Ok	0,952	0,945	0,942	0,940	0,940	0,945
Pinbelegung	0,925	0,922	0,910	0,900	0,900	0,890

Trotz der Variationen aufgrund der begrenzten Vielfalt in den Daten der Minderheitenklassen ist der Klassifikator behält in beiden Klassen eine hohe Genauigkeit und Präzision bei. Die konstante Leistung Metriken für die Klassen „Pin_Ok“ und „Pin_Out“ demonstrieren die Robustheit des Modells und Zuverlässigkeit bei der Identifizierung und Klassifizierung von Mehrheits- und Minderheitsklasseninstanzen.

Die Ergebnisse unterstreichen die praktische Anwendung von Deep Learning-Modellen auf Metallstifte Erkennungsaufgaben, was die Notwendigkeit weiterer Verbesserungen beim Modelltraining unterstreicht und

Machine Translated by Google

Generierung synthetischer Daten, um verschiedene Betriebsszenarien effektiv zu handhaben. Die wichtigsten Beiträge der Forschung zu ARDAD sind wie folgt aufgeführt:

- **Räumlich-zeitliche Analyse für die automatische Überwachung von Verkehrsbarrieren auf der Auckland Harbour Bridge und andere Verkehrsstandorte, die dieselbe Barriere zur Steuerung des Verkehrsflusses verwenden;**
- **Umwandlung eines PoC [7] in ein MVP mit einsetzbaren KI-Algorithmen für E ARDAD als Beispiel für die Umsetzung von Forschung in die Praxis;**
- **Halbautomatische Methoden zur synthetischen Datengenerierung, um Modelle des maschinellen Lernens für komplexe ARDAD-Aufgaben zu verbessern und sich mit der Spärlichkeit und Seltenheit kritischer Verkehrsereignisse zu befassen;**
- **Integration von maschinellem Lernen mit Kernelmanipulation zur dynamischen Anomalieerkennung , um die Präzision aktueller ARDAD-Systeme zu verbessern und die durchschnittliche Erkennungsgenauigkeit von 0,826 auf 0,939 zu erhöhen;**
- **Teilnahme an interdisziplinärer Zusammenarbeit, um die Fortschritte von ARDAD an die Anforderungen der Interessengruppen anzupassen und computergestützte Forschung mit Verkehrsmanagementlösungen zu verbinden .**

4.5. Diskussion

Die ersten Versuche, Hochgeschwindigkeitsvideos aus öffentlichen Verkehrsmitteln und Privatwagen aufzunehmen, stellten erhebliche Herausforderungen dar. Verkehrsflussvorhersagen zeigten, dass die Geschwindigkeit auf der Auckland Harbour Bridge gelegentlich unter 20 km/h fallen könnte, was im Idealfall die Aufnahme von Bildern mit 240 fps und perfekt ausgerichteten Pins ermöglichen würde. Diese Bedingungen wurden jedoch selten erfüllt, und die Verwendung von Verkehrsstaus während der Hauptverkehrszeiten brachte aufgrund unregelmäßiger Stillstandszeiten und Einschr Angesichts der kritischen Sicherheitsanforderungen an der Auckland Harbour Bridge wurden alle Datenerfassungsbemühungen von Experten der NZTA überwacht, die auch Zugang zur Barrier Transfer Machine (BTM) und zum Einsatzort gewährten und die erforderlichen Sicherheitseinweisungen durchführten. Um die schmalen Lücken zwischen den beweglichen Betonsegmenten zu sehen, wurden hochauflösende Kameras (GoPro 5, GoPro 8 und GoPro 9) auf der BTM Die Aufnahmen wurden unter verschiedenen Wetterbedingungen und zu verschiedenen Tageszeiten gemacht, um verschiedene Einsatzszenarien zu erfassen. Die Schwierigkeit, falsch positionierte Pins zu finden und aufzuzeichnen, behinderte den Forschungsprozess erheblich. Nach zahlreichen erfolglosen Versuchen wurden synthetische Frames als praktikable Lösung gewählt. Der NZTA wurde ein Zwischenbericht vorgelegt , der hierarchisches Clustering und eine visuelle Trennung von Merkmalsvektoren in Bezug auf die Minderheitsausgabeklasse unter Verwendung von auf Pearson- und Cosinus Solche rechnerisch anspruchsvolleren Maße wurden gegenüber einfacheren wie dem euklidischen Verfahren ausgewählt, da die aus CNNs extrahierten Merkmale im Verhältnis zur Anzahl der Stichproben von Minderheitsklassen eine hohe Dimensionalität aufweisen und Informationen extrahiert werden müssen, die invariant hinsichtlich der Lichtverhältnisse, des Niederschlags oder der Hintergrundfarben vorbeifahre Klassifizierungen über die binäre Klassifizierung hinaus (Stift in oder außerhalb der Position) bleiben unerforscht, wie etwa Szenarien, in denen der Stift-ROI vollständig verdeckt ist oder Metallstifte teilweise herausragen. Die Ausweitung der binären Klassifizierung auf ein Mehrklassensystem könnte es zukünftigen Systemen ermöglichen, verschiedene Arten von Schäden zu erkennen, die unterschiedliche Wartungsmaßnahmen erfordern. Um solche Verbesserungen zu unterstützen, wären zusätzliche Datensätze erforderlich, die ein breiteres Spektrum an Anomalien erfassen, und mehr synthetische Daten, und zwar gemäß den in unserem Algorithmus zur Erstellung synthetischer Daten in Tabelle 2 beschriebenen Methoden. Die erreichte durchschnittliche Erkennungsgenauigkeit betrug 0,93, was angesichts der zahlreichen Herausforderungen, die während des Modelltrainings aufgetreten sind, lobenswert ist. Im Vergleich zu anderen regionenbasierten Detektoren bot unser Hybridmodell eine höhere Genauigkeit und eine überlegene Verarbeitungsgeschwindigkeit und verarbeitete 40 bis 45 Bilder pro Sekunde mit einer Genauigkeit von bis zu 93,6 %. Die Integrations- und Tiefenverkettungsebenen verbesserten die Erkennung kleinerer Objekte, indem sie Bilddetails auf niedriger Ebene in den Erkennungsprozess einbezogen, was durch eine Abfolge von Faltungs-, ReLU- und Batch-Normalisierungsebenen erleichtert wurde. Die MATLAB-App bietet eine robuste Plattform für die Ausweitung der Forschung auf zukünftige Anwendungen. Im Hinblick auf den Umgang mit dem Unbekannten und der Forschungsunsicherheit lohnt es sich für die Forschungsger

Machine Translated by Google

Zu solchen Überlegungen können Änderungen beim Personal der Industriepartner, mögliche Lockdowns aufgrund der Pandemie, staatliche Förderung und Richtlinienaktualisierungen gehören, die Flexibilität bei der Projektplanung und Datenerfassung erfordern.

Eine Zusammenfassung der praktischen Aspekte dieser Studie ist wie folgt: • Die Technologie bietet eine kostengünstige automatisierte Lösung für Fahrspur- und allgemeinen Verkehr Sicherheit, die menschliche Inspektionen ergänzt, aber nicht ersetzt.

- **Das System erhöht die Inspektionsfrequenz, verbessert die Privatsphäre und ermöglicht die Erstellung von digitalen Aufzeichnungen für analytische Erkenntnisse zur Verkehrssicherheit.**
- **Zukünftige wissenschaftliche Bemühungen werden einen umfangreicheren Datensatz verwenden, um sich auf die Entwicklung adaptiver Modelle und die Leistungssteigerung zu konzentrieren. Zusätzliche Datenvisualisierung und hybride methodische Ansätze werden untersucht.**
- **Für unseren Industriepartner, die NZTA, ebnet das Projekt den Weg für eine unabhängige Softwareentwicklung und eine mögliche Systemintegration in umfassendere Smart-City-Infrastrukturen .**
- **Der Übergang von einem minimal funktionsfähigen Produkt (MVP) zu Produktionssystemen wird umfangreiche Tests, Codeoptimierungen und möglicherweise einen Übergang von MATLAB zu Python beinhalten, um die Integrationsfähigkeiten der Rechenleistung zu verbessern und die Möglichkeit einer Bindung des Anbieters an proprietäre Infrastrukturen und Datenverarbeitung außerhalb der nationalen Gerichtsbarkeit zu minimieren.**
- **Diese Studie legt den Grundstein für zukünftige Innovationen in der Verkehrsmanagementtechnologie und versetzt die NZTA in die Lage, die Fortschritte bei der laufenden Modernisierung ihrer Verkehrsinfrastruktur und ihren Smart-City-Initiativen zu nutzen.**

Bei zukünftigen Systemweiterentwicklungen wird die Möglichkeit bestehen, die Pin-Statusverfolgung aus verschiedenen Blickwinkeln zu ermöglichen, möglicherweise unter Einbeziehung zusätzlicher Datenerfassungsprotokolle und -technologien, um die Vielseitigkeit und den Anwendungsbereich des Systems zu erweitern.

5. Schlussfolgerungen und zukünftige Arbeiten

Die Auckland Harbour Bridge spielt eine entscheidende Rolle in der Infrastruktur von Auckland. Der Verkehr ist ungleichmäßig, aber vorhersehbar und weist während der morgendlichen und abendlichen Hauptverkehrszeiten eine Umkehr des Volumens auf. Bewegliche Betonbarrieren (MCBs) haben sich bei der Bewältigung von Engpässen im Nahverkehr als wirksam erwiesen. Die Anfälligkeit der Brücke für verschiedene Arten von Vibrationen, insbesondere im erhöhten Mittelteil, gibt jedoch Anlass zu Sicherheitsbedenken und erfordert häufige Inspektionen der Bolzen. Weitere Gesundheits- und Sicherheitsbedenken sind die ausschließliche Abhängigkeit von manuellen Inspektionen und das Potenzial menschlicher Fehler im Zusammenhang mit der Sicherheit des Inspektionspersonals am Arbeitsplatz, einer ungesunden ergonomischen Haltung

Um die Sicherheit und die Häufigkeit der Sicherheitsüberwachung des MCB zu erhöhen, haben wir ein datenschutzfreundliches automatisiertes Überwachungssystem entwickelt, das von den auf der Auckland Harbour Bridge gesammelten Daten auf ähnliche Kontexte wie Verkehrsflussregelung und Sicherheitsüberwachungsanwendungen übertragbar ist **[10]. Eine neuartige Technik zur Generierung synthetischer Frames wurde eingeführt, um verschiedene unsichere Pin-Positionen zu simulieren und so die inkrementelle Modellentwicklung und Leistungsoptimierung zu unterstützen. Diese Forschung hat erfolgreich gezeigt, dass der Prototyp unsichere Pin-Positionen direkt aus Live-Feeds und zuvor aufgezeichneten Video-Frames unter unterschiedlichen Lichtbedingungen (wie strahlendem Sonnenschein, starkem Regen und „weichen“ Lichtbedingungen am frühen Morgen) erkennen kann. Der Mangel an Video-Frames, die einen Pin_Out-Status zeigen, wurde durch die Einführung einer Methode zur Erstellung synthetischer Bilder zur Verbesserung des Modellierungsprozesses behoben. Die erwartete Gesamtleistung des Systems für die Erkennung von Pin-Regionen, die Frame-Auswahl und die Pin-Klassifizierung wurde auf über 80 % geschätzt, wobei einzelne Modelle bei einem begrenzten Datensatz eine Genauigkeit von bis zu 99 % erreichen, wie in Tabelle 5 dargestellt. Diese Ergebnisse erfordern eine weitere Validierung anhand eines größeren und ausgewogeneren zukünftigen Datensatzes. Das System zur Erkennung des Pin-Status und zur Alarmierung wies eine wünschenswerte Präzision und Genauigkeit auf, wobei einige Leistungseinbußen auf die Unausgewogenheit des Datensatzes, unterschiedliche Lichtverhältnisse sowie Kamerawinkel und -entfernungen zurückzuführen waren. Die Ergebnisse eines kleineren Datensatzes mit Beschriftung lassen darauf schlie**

Machine Translated by Google

manuelle Kennzeichnung. Die Integration eines Hybridmodells erleichterte die Analyse und sorgte für Flexibilität für zukünftige Modellanpassungen mit minimalen Anforderungen an die Datenkennzeichnung.

Zukünftige Arbeiten umfassen die Sammlung weiterer Videodaten, darunter zusätzliche Videos von den Wartungsteams der NZTA und AHB aufgezeichnet. Die erweiterte Datenerfassung ist Es wird erwartet, dass das grundlegende System gestärkt wird und dazu beiträgt, universell anwendbare ARDAD-Systeme für ähnliche Verkehrssicherheitskontexte weltweit weiterzuentwickeln. Während Die Ergebnisse bei Erkennung und Klassifizierung sind vielversprechend, es besteht jedoch erhebliches Potenzial für weitere Fortschritte bei der Integration des Pin-ROI-Trackings in das Warnsystem. Zukünftige Iterationen des Systems kann auch fortschrittliche Technologien wie LiDAR und GPS nutzen, die werden in modernen Mobilgeräten immer häufiger eingesetzt. Die Entwicklung zusätzlicher Die Fähigkeiten erfordern umfangreiches Systemtraining und Anpassung an angereicherte Datensätze die verschiedene Pin-Bedingungen und Szenarien erfassen, was möglicherweise zu robusteren und reaktionsschnelle Verkehrsmanagementlösungen.

Autorenbeiträge: Konzeptualisierung, MR; Software, MR; Validierung, MR; Datenkuratierung, MR; Schreiben – Originalentwurf, MR; Schreiben – Überprüfung und Bearbeitung, BB und MD; Überwachung, BB und MD. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Forschung erhielt keine externe Finanzierung.

Datenverfügbarkeitserklärung: Datensatz auf Anfrage bei den Autoren erhältlich.

Danksagungen: Wir danken Gary Bonser, Martin Olive, David Ranby, und Angela Potae von der NZTA und Auckland System Management (ASM) für ihre unschätzbare Hilfe beim Transport, Sicherheitseinweisungen, Videoaufzeichnungen, Projekteinblicke und laufende Betreuung. Wir danken außerdem der Auckland University of Technology (AUT) für die Finanzierung und das PhD-Stipendium Support und Bereitstellung des Zugriffs auf Computerhardware, Bibliothek und Aufnahmegeräte. Wir möchten auch die Dokumentations- und Softwareentwicklungsbemühungen der Industrie würdigen, Hochschulen und Open-Source-Communitys, die MATLAB, Orange Data-Mining, SqueezeNet, TensorFlow, ImageNet, Google Cloud und OpenCV.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte vorliegen.

Notationen

Die folgende Tabelle fasst die wichtigen Symbole und mathematischen Notationen zusammen, die in diesem Papier:

Symbol	Beschreibung
Y	Luminanzkomponente im YCbCr-Farbraum
xc	Schwerpunktposition eines Objekts beim Template-Matching
xi	Pixelposition beim Template-Matching
wi	Pixelintensität beim Template-Matching
P(Xt)	Wahrscheinlichkeit der Beobachtung eines Pixelwertes X zum Zeitpunkt t in einem Gaußschen Mischmodell
wk,t	Gewicht der k-ten Gauß-Komponente zum Zeitpunkt t
μk,t	Mittelwert der k-ten Gauß-Verteilung zum Zeitpunkt t
ÿk,t	Kovarianzmatrix der k-ten Gauß-Verteilung zum Zeitpunkt t
ÿMt	Bewegungserkennungsmetrik kombiniert Frameunterschiede und optischen Fluss
Ft	Rahmen zum Zeitpunkt t
ÿlt	Gradient des Bildes zum Zeitpunkt t
vt	Optischer Flussvektor zum Zeitpunkt t
ÿ	Gewichtungsfaktor in der Bewegungserkennungsgleichung
Kk	Kalman-Verstärkung zum Zeitpunkt k
zk	Tatsächliche Messung zum Zeitpunkt k
Hk	Messmatrix, die den Zustandsraum in den Messraum abbildet

Machine Translated by Google

Verweise

1. **New Zealand Transport Agency Waka Kotahi. Auckland Harbour Bridge Factsheet. 2024. Online verfügbar: https://www.nzta. govt.nz/assets/site-resources/content/about/docs/auckland-harbour-bridge-factsheet.pdf (abgerufen am 10. April 2024).**
2. **Te Waihanga New Zealand Infrastructure Commission. Neuseelands Infrastrukturvermögenswert, Investitionen und Abschreibungen, 1990–2022. 2023. Online verfügbar: https://tewaihanga.govt.nz/our-work/research-insights/build-or-maintain (abgerufen am 10. April 2024).**
3. **New Zealand Transport Agency Waka Kotahi. So verschieben Sie eine Betonbarriere auf einer Autobahn. 2024. Online verfügbar: https: //www.nzta.govt.nz/media-releases/how-to-move-a-concrete-autoway-barrier/ (abgerufen am 10. April 2024).**
4. **Yang, X.; Zhang, J.; Liu, W.; Jing, J.; Zheng, H.; Xu, W. Automatisierung bei der Erkennung, Diagnose und Behandlung von Verkehrspannen.** J. Road Eng. **2024, 4, 1–26. [CrossRef]**
5. **Bai, D.; Li, G.; Jiang, D.; Yun, J.; Tao, B.; Jiang, G.; Sun, Y.; Ju, Z. Methoden zur Erkennung von Oberflächendefekten bei Industrieprodukten mit unausgewogenen Proben: Ein Überblick über die Fortschritte in den 2020er Jahren.** Eng. Appl. Artif. Intell. **2024,** 130, **107697. [CrossRef]**
6. **Trilles, S.; Hammad, SS; Iskandaryan, D. Anomalieerkennung basierend auf künstlicher Intelligenz der Dinge: Eine systematische Literaturzuordnung .** Internet Things **2024,** 25, **101063. [CrossRef]**
7. **Baÿcí´c, B.; Rathee, M.; Pears, R. Automatisierte Inspektion beweglicher Fahrbahnbegrenzungen für die Verkehrssicherheit der Auckland Harbour Bridge. In Proceedings of the Neural Information Processing: 27. Internationale Konferenz, ICONIP 2020, Bangkok, Thailand, 23.–27. November 2020; Proceedings, Teil I 27. Springer: Berlin/Heidelberg, Deutschland, 2020; S. 150–161.**
8. **Klarák, J.; Andok, R.;** Malík, **P.; Kuric, I.; Ritomský, M.; Klaÿcková, I.; Tsai, H.-Y. Von der Anomalieerkennung bis zur Fehlerklassifizierung.** Sensoren **2024,** 24, **429. [CrossRef] [PubMed]**
9. **Lozano-Ramírez, NE; Sánchez, O.; Carrasco-Beltrán, D.; Vidal-Méndez, S.; Castañeda, K. Digitalisierung und Nachhaltigkeit in lineare Projekttrends: Eine bibliometrische Analyse.** Sustainability **2023,** 15, **15962. [CrossRef]**
10. **Wikipedia. Barriere-Transfermaschine. 2024. Online verfügbar: https://en.wikipedia.org/wiki/Barrier_transfer_machine (Zugriff am 20. April 2024).**
11. **Rathee, M.** Sicherheitsabschirmung der beweglichen Betonbarriere der Harbour Bridge von Auckland; **Auckland University of Technology: Auckland, Neuseeland, 2021.**
12. **Baccari, S.; Hadded, M.; Ghazzai, H.; Touati, H.; Elhadef, M. Anomalieerkennung in vernetzten und autonomen Fahrzeugen: Eine Untersuchung, Analyse und Forschungsherausforderungen.** IEEE Access **2024,** 12, **19250–19276. [CrossRef]**
13. **Rathee, M.; Baÿcí´c, B.; Doborjeh, M. Automatische Erkennung von Straßenschäden und Anomalien für die Verkehrssicherheit: Eine systematische Übersicht.** Sensors **2023,** 23, **5656. [CrossRef] [PubMed]**
14. **Cui, Y.; Liu, Z.; Lian, S. Eine Untersuchung zu unüberwachten Anomalieerkennungsalgorithmen für industrielle Bilder.** IEEE Access **2023,** 11, **55297–55315. [CrossRef]**
15. **Cottrell, BH Evaluation of a Movable Concrete Barrier System. [Technischer Bericht] 1994. Online verfügbar: https://rosap.ntl.bts.gov/Ansicht/Punkt/19352 (abgerufen am 9. April 2024).**
16. **Poe, CM Bewegliche Betonbarrieren als Ansatz für die Gestaltung und den Betrieb einer Gegenverkehrsspur für private Nutzfahrzeuge.** Transp. Res. Rec. **1991, 40–54.**
17. **Chirayil Nandakumar, S.; Mitchell, D.; Erden, MS; Flynn, D.; Lim, T. Anomalieerkennungsmethoden in autonomen Robotern Missionen.** Sensors **2024,** 24, **1330. [CrossRef]**
18. **Galvão, YM; Castro, L.; Ferreira, J.; Neto, FBdL; Fagundes, RAdA; Fernandes, BJ Anomalieerkennung in Smart Houses für das Gesundheitswesen: Jüngste Fortschritte und Zukunftsperspektiven.** SN Comput. Sci. **2024, 5, 136. [CrossRef]**
19. **Es-Swidi, A.; Ardchir, S.; Elghoumari, Y.; Daif, A.; Azouazi, M. Verkehrsstaus und Erkennung von Straßenanomalien mithilfe der Bildverarbeitung von CCTVs, Herausforderungen und Chancen. In** der International Conference on Advanced Intelligent Systems for Sustainable Development; **Springer: Berlin/Heidelberg, Deutschland, 2022; S. 92–105.**
20. **Gao, J.; Zuo, F.; Ozbay, K.; Hammami, O.; Barlas, ML Ein neuer Ansatz zur Überwachung von Fahrspuren am Straßenrand und zur Abschätzung der Auswirkungen von illegalem Parken basierend auf der Warteschlangentheorie und Computervision für Kameras mit niedriger Auflösung und niedriger Bildrate.** Transp. Res. Teil A Policy Pract. **2022,** 162, **137–154. [CrossRef]**
21. **Kim, S.; Anagnostopoulos, G.; Barmpounakis, E.; Geroliminis, N. Visuelle Erweiterungen und Anomalieerkennung im pNEUMA Experiment mit einem Schwarm Drohnen.** Transp. Res. Part C Emerg. Technol. **2023,** 147, **103966. [CrossRef]**
22. **Yi, K.; Luo, K.; Chen, T.; Hu, R. Ein verbessertes YOLOX-Modell und eine Domänentransferstrategie für nächtliche Fußgänger- und Fahrzeugbeobachtungen. Erkennung.** Appl. Sci. **2022,** 12, **12476. [CrossRef]**
23. **Wang, J.; Wang, X.; Hao, R.; Yin, H.; Huang, B.; Xu, X.; Liu, J. Inkrementelles Template-Nachbarschafts-Matching zur 3D-Anomalieerkennung .** Neurocomputing **2024,** 581, **127483. [CrossRef]**
24. **Pan, Q.; Bao, Y.; Li, H. Transferlernbasierte Datenanomalieerkennung zur strukturellen Gesundheitsüberwachung.** Struct. Health Monit. **2023,** 22, **3077–3091. [CrossRef]**
25. **Yan, P.; Abdulkadir, A.; Luley, P.-P.; Rosenthal, M.; Schatte, GA; Grewe, BF; Stadelmann, T. Eine umfassende Untersuchung zum Deep Transfer Learning zur Anomalieerkennung in industriellen Zeitreihen: Methoden, Anwendungen und Richtungen.** IEEE Access **2024,** 12, **3768–3789. [CrossRef]**
26. **Bharambe, U.; Bhangale, U.; Narvekar, C. Rolle der Mehrzieloptimierung bei der Bildsegmentierung und -klassifizierung. In** Computational Intelligence in Image and Video Processing; **Chapman and Hall/CRC: Boca Raton, FL, USA, 2023; S. 317–340.**
27. **Gonzalez, RC** Digitale Bildverarbeitung; **Pearson Education India: Hoboken, NJ, USA, 2009.**

Machine Translated by Google

28. Yuan, Y.; Huang, J.; Yu, J.; Tan, JKS; Chng, KZ; Lee, J.; Kim, S. Anwendung von Algorithmen des maschinellen Lernens zur genauen Bestimmung des Bilirubinspiegels auf in vitro künstlich erzeugten Gewebephantombildern. Sci. Rep. **2024,** 14, **5952.** [CrossRef] [PubMed]

29. Kilicaslan, M.; Tanyeri, U.; Demirci, R. Bildabruf mithilfe eines eindimensionalen Farbhistogramms, das mit Entropie erstellt wurde. Adv. Electr. Comput. Eng. **2020,** 20, **79–88.** [CrossRef]

30. Mittal, H.; Pandey, AC; Saraswat, M.; Kumar, S.; Pal, R.; Modwel, G. Eine umfassende Untersuchung der Bildsegmentierung: Clustering- Methoden, Leistungsparameter und Benchmark-Datensätze. Multimed. Tools Appl. **2022,** 81, **35001–35026.** [CrossRef]

31. Vansh, V.; Chandrasekhar, K.; Anil, C.; Sahu, SS Verbesserte Gesichtserkennung mit YCbCr und Adaboost. In Proceedings der 5. Internationalen Konferenz über Computational Intelligence im Data Mining (ICCIDM 2018), Burla, Indien, 15.–16. Dezember 2018; Springer: Berlin/Heidelberg, Deutschland, 2020; S. 689–699.

32. Han, H.; Han, C.; Lan, T.; Huang, L.; Hu, C.; Xue, X. Automatische Schattenerkennung für multispektrale Satellitenfernerkundung Bilder in invarianten Farbräumen. Appl. Sci. **2020,** 10, **6467.** [CrossRef]

33. Sahu, Y.; Tripathi, A.; Gupta, RK; Gautam, P.; Pateriya, RK; Gupta, A. Eine CNN-SVM-basierte computergestützte Diagnose von Brustkrebs unter Verwendung der Histogramm-K-Means-Segmentierungstechnik. Multimed. Tools Appl. **2023,** 82, **14055–14075.** [CrossRef]

34. Kollem, S.; Reddy, KR; Rao, DS Ein optimierter SVM-basierter possibilistischer Fuzzy-C-Means-Clusteralgorithmus für die Tumorsegmentierung . Multimed. Tools Appl. **2021,** 80, **409–437.** [CrossRef]

35. Agrawal, S.; Natu, P. Ein verbessertes, auf der Gaussian Mixture Method basierendes Hintergrundsubtraktionsmodell zur Erkennung bewegter Objekte in Außenszenen. In Proceedings der 2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, Indien, 15.–17. September 2021; IEEE: Piscataway, NJ, USA, 2021; S. 1–8.

36. Rakesh, S.; Hegde, NP; Gopalachari, MV; Jayaram, D.; Madhu, B.; Hameed, MA; Vankdothu, R.; Kumar, LS Bewegtes Objekt Nachweis mittels modifizierter GMM-basierter Hintergrundsubtraktion. Meas. Sens. **2023,** 30, **100898.** [CrossRef]

37. Zhou, Q.; Situ, Z.; Teng, S.; Chen, G. Vergleichende Wirksamkeit der Datenerweiterung mit traditionellen Ansätzen gegenüber StyleGANs bei der automatischen Erkennung von Kanalfehlern. J. Water Resour. Plan. Manag. **2023,** 149, **04023045.** [CrossRef]

38. Zuehlke, D.; Henderson, TA; McMullen, S. Maschinelles Lernen durch Template-Matching, angewandt auf die Objektverfolgung in Videodaten. Artif. Intell. Mach. Learn. Multi-Domain Oper. Appl. **2019,** 11006, **110061S.**

39. Ge, Y.; Zhang, J.; Ren, X.; Zhao, C.; Yang, J.; Basu, A. Tiefes Variationstransformationsnetzwerk zur Vordergrunderkennung. IEEE Trans. Schaltungssysteme. Videotechnik. **2020,** 31, **3544–3558.** [CrossRef]

40. Cao, Q.; Wang, Z.; Long, K. Verkehrsvordergrunderkennung an komplexen städtischen Kreuzungen mithilfe eines neuartigen Hintergrundwörterbuch- Lernmodells. J. Adv. Transp. **2021,** 2021, **3515512.** [CrossRef]

41. Zhang, Y.; Zheng, W.; Leng, K.; Li, H. Hintergrundsubtraktion mithilfe einer adaptiven lokalen Mediantexturfunktion bei Beleuchtung verändert urbane Verkehrsszenen. IEEE Access **2020, 8, 130367–130378.** [CrossRef]

42. Feng, J.; Zeng, D.; Jia, X.; Zhang, X.; Li, J.; Liang, Y.; Jiao, L. Cross-Frame-Keypoint-basierte und räumlich bewegungsinformationsgesteuerte Netzwerke zur Erkennung und Verfolgung von fahrenden Fahrzeugen in Satellitenvideos. ISPRS J. Photogramm. Remote Sens. **2021,** 177, **116–130** [Querverweis]

43. Yang, G.; Ramanan, D. Verbesserung des optischen Flusses zum 3D-Szenenfluss durch optische Erweiterung. In Proceedings der IEEE/CVF- Konferenz zu Computer Vision und Mustererkennung 2020, Seattle, WA, USA, 13.–19. Juni 2020; S. 1334–1343.

44. Chen, X.; Jia, Y.; Tong, X.; Li, Z. Forschung zur Fußgängererkennung und Deepsort-Verfolgung vor intelligenten Fahrzeugen basierend auf Deep Learning. Sustainability **2022,** 14, **9281.** [CrossRef]

45. Sun, M.; Davies, ME; Proudler, IK; Hopgood, JR Adaptiver Kernel-Kalman-Filter. IEEE Trans. Signal Process. **2023,** 71, **713–726.** [Querverweis]

46. Chaurasiya, RK; Gondane, PM; Acharya, B.; Khan, MI Automatischer Straßenverkehrsanalysator mit Hintergrundsubtraktion, Blob- Analyse und Tracking-Algorithmen. In Proceedings der 7. Internationalen Konferenz über Computeranwendungen in der Elektrotechnik – jüngste Fortschritte (CERA) 2023, Roorkee, Indien, 27.–29. Oktober 2023; IEEE: Piscataway, NJ, USA, 2023; S. 1–6.

47. Sajeed, MA; Kelouwani, S.; Amamou, A.; Alam, MZ; Agbossou, K. Schätzung des Spurverlassens von Fahrzeugen auf städtischen Straßen unter Verwendung von GIS- Informationen. In Proceedings der 2021 IEEE Vehicle Power and Propulsion Conference (VPPC), Gijon, Spanien, 25.–28. Oktober 2021; IEEE: Piscataway, NJ, USA, 2021; S. 1–7.

48. Momin, KA; Barua, S.; Jamil, MS; Hamim, OF Kurzzeit-Verkehrsflussvorhersage mittels Kalman-Filterung. AIP Conf. Proc. **2023,** 2713, **040011.**

49. Moghaddasi, SS; Faraji, N. Ein Hybridalgorithmus basierend auf Partikelfilter und genetischem Algorithmus zur Zielverfolgung. Expert Syst. Appl. **2020,** 147, **113188.** [CrossRef]

50. Chen, S.; Huang, L.; Chen, H.; Bai, J. Mehrspurige Erkennung und Verfolgung mithilfe eines zeitlich-räumlichen Modells und einer Partikelfilterung. IEEE Trans. Intell. Transp. Syst. **2021,** 23, **2227–2245.** [CrossRef]

51. Nissimagoudar, P.; Algur, N.; Bonageri, N.; Chavan, A.; Koppa, A.; Iyer, NC Mehrfachfahrzeugverfolgung mithilfe des Meanshift-Algorithmus und 8-Punkt-Konnektivität. In Proceedings der International Conference on Soft Computing and Pattern Recognition, Online, 14.–16. Dezember 2022; Springer: Berlin/ Heidelberg, Deutschland, 2022; S. 3–13.

52. Zhang, B.; Li, Z.; Perina, A.; Del Bue, A.; Murino, V.; Liu, J. Adaptive lokale Bewegungsmodellierung für robustes Objekt-Tracking. IEEE Trans. Circuits Syst. Video Technol. **2016,** 27, **1515–1526.** [CrossRef]

53. O'Mahony, N.; Campbell, S.; Carvalho, A.; Harapanahalli, S.; Hernandez, GV; Krpalkova, L.; Riordan, D.; Walsh, J. Deep Learning vs. traditionelle Computervision. In Proceedings of the Advances in Computer Vision: 2019 Computer Vision Conference (CVC), Las Vegas, NV, USA, 2.–3. Mai 2019; Springer: Berlin/Heidelberg, Deutschland, 2020; Band 1, S. 128–144.

Machine Translated by Google

54. Zhang, W.; Li, H.; Li, Y.; Liu, H.; Chen, Y.; Ding, X. Anwendung von Deep-Learning-Algorithmen in der Geotechnik: Eine kurze kritische Rezension. Artif. Intell. Rev. **2021,** 54, **5633–5673.** [CrossRef]

55. Long, G.; Zhang, Z. Deep encrypted traffic detection: Ein Framework zur Anomalieerkennung für verschlüsselten Datenverkehr basierend auf parallelen automatische Merkmalsextraktion. Comput. Intell. Neurosci. **2023, 3316642.** [CrossRef] [PubMed]

56. Butt, UM; Ullah, HA; Letchmunan, S.; Tariq, I.; Hassan, FH; Koh, TW Nutzung von Transferlernen für räumlich-zeitliche Aktivitätserkennung aus Videosequenzen. Comput. Mater. Contin. **2023,** 74, **5017–5033.**

57. Ouhami, M.; Hafiane, A.; Es-Saady, Y.; El Hajji, M.; Canals, R. Computer Vision, IoT und Datenfusion zur Erkennung von Pflanzenkrankheiten mithilfe von maschinellem Lernen: Eine Umfrage und laufende Forschung. Remote Sens. **2021,** 13, **2486.** [CrossRef]

58. Carlson, MP; Bloom, I. Die zyklische Natur der Problemlösung: Ein sich entwickelnder mehrdimensionaler Problemlösungsrahmen. Educ Stud. Math. **2005, 58 , 45–75.**

59. Vallenga, D.; Grypdonck, MHF; Hoogwerf, LJR; Tan, FIY Aktionsforschung: Was, warum und wie? Acta Neurol. Belg. **2009,** 109, **81–90 .**

60. Farady, I.; Lin, C.-Y.; Chang, M.-C. PreAugNet: Verbessern Sie die Datenerweiterung für die industrielle Defektklassifizierung mit kleinskaligen Trainingsdaten. J. Intell. Manuf. **2024,** 35, **1233–1246.** [CrossRef]

61. Xu, M.; Yoon, S.; Fuentes, A.; Park, DS Eine umfassende Übersicht über Bilderweiterungstechniken für Deep Learning. Mustererkennung . **2023,** 137, **109347.** [CrossRef]

62. Mathworks. Entwickeln Sie Apps mit App Designer. 2021. Online verfügbar: https://www.mathworks.com/help/matlab/app-designer.html (abgerufen am 10. April 2024).

63. Kingma, DP; Ba, J. Adam: Eine Methode zur stochastischen Optimierung. arXiv **2014, arXiv:1412.6980.**

64. Sutskever, I.; Martens, J.; Dahl, G.; Hinton, G. Über die Bedeutung von Initialisierung und Momentum beim Deep Learning. In Proceedings der International Conference on Machine Learning, Atlanta, GA, USA, 16.–21. Juni 2013; PMLR; S. 1139–1147.

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Kommentar

Kommentar zu Sakli, H. Zylindrischer Wellenleiter auf Ferritsubstrat, gesteuert durch ein extern angelegtes Magnetfeld. Elektronik 2021, 10, 474

Afshin Moradi

Fakultät für Technische Physik, Technische Universität Kermanshah, Kermanshah 6715685420, Iran; a.moradi@kut.ac.ir

Zusammenfassung: Sakli untersuchte kürzlich die Ausbreitung elektromagnetischer Wellen in metallischen zylindrischen Wellenleitern, die mit longitudinal magnetisiertem Ferrit gefüllt sind, und konzentrierte sich dabei auf die TEz- (transversal elektrisch) und TMz- Modi (transversal magnetisch) relativ zur z-Achse. Dieser Kommentar hebt hervor, dass das vorgeschlagene System die Ausbreitung der TEz- und TMz- Wellen im Allgemeinen nicht unterstützen kann, wodurch die von Sakli erzielten Hauptergebnisse ungültig werden.

Schlüsselwörter: anisotrope Materialien; Antenne; zylindrische Wellenleiter; Ferrite; Ausbreitung

check for updates

Zitat: Moradi, A. Kommentar zu Sakli, H. Zylindrischer Wellenleiter auf Ferritsubstrat, gesteuert durch ein extern angelegtes Magnetfeld. Elektronik **2021,** 10, **474.** Elektronik **2024,** 13, **3031. https://doi.org/10.3390/Elektronik13153031**

Wissenschaftlicher Redakteur: Flavio Canavero

Empfangen: 9. Mai 2024
Überarbeitet: 11. Juli 2024
Akzeptiert: 29. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 beim Autor. Lizenznehmer MDPI, Basel, Schweiz. Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https://creativecommons.org/licenses/by/4.0/).

1. Einleitung

Kürzlich untersuchte Sakli [1] die Ausbreitung von TEz- und TMz -Moden in metallischen zylindrischen Wellenleitern, die mit verlustfreiem, longitudinal magnetisiertem Ferrit gefüllt sind. Er demonstrierte, wie man Dispersionsdiagramme erhält und diskutierte den Einfluss anisotroper Parameter auf Dispersionseigenschaften und Grenzfrequenzen. Zusätzlich präsentierte er numerische Ergebnisse für die TEz- und TMz- Moden.

Basierend auf der Theorie der mit Ferrit gefüllten zylindrischen Wellenleiter, die zu Beginn der 50er Jahre des letzten Jahrhunderts entwickelt wurde **[2,3],** sollte jedoch bei **der vorgeschlagenen Ausbreitung in metallischen zylindrischen Wellenleitern eine Hybridwelle erwartet werden ,** und daher **können sich** TEz- **und TMz** -Wellen nicht ausbreiten. Darüber hinaus unterstützt **ein metallischer zylindrischer Wellenleiter, der mit homogenen anisotropen Materialien gefüllt ist, im Allgemeinen nur Hybridmodi, wie in unserer jüngsten Studie [4] gezeigt wurde. Dies bedeutet, dass die oben genannten Ergebnisse für die Ausbreitung von TEz- und TMz- Modi, die von Sakli [1]** abgeleitet wurden , **ungültig sind. Beachten wir, dass es auch gut belegt ist, dass in allgemeinen Fällen die Trennung elektromagnetischer Wellen in TE- und TM- Modi in metallischen rechteckigen Wellenleitern nicht**

Wir sind der Ansicht, dass fehlerhafte Veröffentlichungen korrigiert werden sollten, um sicherzustellen, dass neue Forscher auf den korrekten Ergebnissen früherer Studien aufbauen können. Aus dieser Motivation heraus haben wir unseren Kommentar verfasst, in dem wir die Fehler in [1] identifizieren.

2. Die strenge elektromagnetische Analyse

Betrachten wir einen metallischen zylindrischen Wellenleiter, der mit longitudinal magnetisiertem Ferrit gefüllt ist, wie in Abbildung 1 von [1] dargestellt. Wir sind an Lösungen für geführte Moden im Wellenleiter interessiert, die sich in z-Richtung ausbreiten. Betrachten wir daher die Form der elektromagnetischen Welle, die sich im Wellenleiter ausbreitet, als

$$E(r, \ddot{y}, z, t) = E(r, \ddot{y})e\, j(\ddot{y}t\ddot{y}kz\ z) \quad , \tag{1}$$

$$H(r, \ddot{y}, z, t) = H(r, \ddot{y})e\, j(\ddot{y}t\ddot{y}kz\ z) \quad , \tag{2}$$

Machine Translated by Google

wobei kz **die Ausbreitungskonstante entlang der z-Richtung ist. Durch Erweitern der Maxwell-Rougel-Gleichungen (d. h. Gleichungen (1) und (2) in [1]) und Verwenden von Gleichung (3) in [1] erhalten wir**

$$\begin{matrix} \ddot{y} & \frac{1}{R}\frac{\ddot{y}Ez}{\ddot{y}\ddot{y}} & \frac{\ddot{y}E\ddot{y}}{\ddot{y}z} \\ & \frac{\ddot{y}Er}{\ddot{y}z} & \frac{\ddot{y}Ez}{\ddot{y}r} \\ \ddot{y}\ddot{y}\frac{1}{R} & \frac{\ddot{y}(rE\ddot{y})}{\ddot{y}r} & \frac{\ddot{y}Er}{\ddot{y}\ddot{y}}\frac{1}{R} \end{matrix} \ddot{y}\ddot{y}\ddot{y} = \ddot{y}\ddot{y}\ddot{y}\mu 0\ddot{y} \begin{matrix} \mu Hr\ddot{y} & j\ddot{y}H\ddot{y} & \ddot{y} \\ j\ddot{y}Hr + \mu H\ddot{y} & & \ddot{y} \\ \mu rzHz & & \ddot{y}, \end{matrix} \quad (3)$$

$$\begin{matrix} \ddot{y} & \frac{1}{R}\frac{\ddot{y}Hz}{\ddot{y}\ddot{y}} & \frac{\ddot{y}H\ddot{y}}{\ddot{y}z} \\ & \frac{\ddot{y}Hr}{\ddot{y}z} & \frac{\ddot{y}Hz}{\ddot{y}r} \\ \ddot{y}\ddot{y}\frac{1}{R} & \frac{\ddot{y}(rH\ddot{y})}{\ddot{y}r} & \frac{\ddot{y}Hr}{\ddot{y}\ddot{y}}\frac{1}{R} \end{matrix} \ddot{y}\ddot{y}\ddot{y} = j\ddot{y}\ddot{y}0\ddot{y}r f \ddot{y} \begin{matrix} Äh & \ddot{y} \\ E\ddot{y} & \ddot{y} \\ Ez & \ddot{y}. \end{matrix} \quad (4)$$

Setzen wir die Gleichungen (1) und (2) in die Gleichungen (3) und (4) ein und führen einige Manipulationen durch, so erhalten wir

$$Er = \frac{kz}{K2_C} \ddot{y}jK2 \frac{\ddot{y}Ez}{\ddot{y}\ddot{y}_r}\frac{1}{\ddot{y}r}\frac{\ddot{y}Ez}{\ddot{y}r} + F c\mu \frac{1}{r} \frac{\ddot{y}Ez}{\ddot{y}\ddot{y}} \quad \frac{1}{K2_C} \quad A1 + jA2 \frac{\ddot{y}Hz}{\ddot{y}r}\frac{1}{\ddot{y}\ddot{y}}\frac{\ddot{y}Hz}{\ddot{y}\ddot{y}} \quad \frac{1}{R}, \quad (5)$$

$$E\ddot{y} = \ddot{y} \frac{kz}{K2_C} F \frac{\ddot{y}Ez}{\ddot{y}r} + jK2 \quad c\mu \frac{1}{r}\frac{\ddot{y}Ez}{\ddot{y}\ddot{y}} + \frac{1}{K2_C} \frac{\ddot{y}Hz}{\ddot{y}r}\frac{1}{\ddot{y}\ddot{y}}\frac{\ddot{y}Hz}{\ddot{y}\ddot{y}} jA2 \quad A1 \frac{1}{R}, \quad (6)$$

$$Std. = \frac{\ddot{y}\ddot{y}0\ddot{y}r f}{K2_C} F \frac{\ddot{y}Ez}{\ddot{y}r} + jK2 \quad c\mu \frac{1}{r}\frac{\ddot{y}Ez}{\ddot{y}\ddot{y}} \quad \frac{1}{K2_C} jkzK \quad c\mu \frac{2}{r}\frac{\ddot{y}Hz}{\ddot{y}r}\frac{\ddot{y}Hz}{\ddot{y}\ddot{y}}\frac{1}{\ddot{y}}Fkz \quad \frac{1}{R}, \quad (7)$$

$$H\ddot{y} = \frac{\ddot{y}\ddot{y}0\ddot{y}r f}{K2_C} \ddot{y}jK2 \quad c\mu \frac{\ddot{y}Ez}{\ddot{y}r}\frac{\ddot{y}Ez}{\ddot{y}\ddot{y}_r}\frac{1}{\ddot{y}r} + F \quad \frac{1}{K2_C} Fkz \frac{\ddot{y}Hz}{\ddot{y}r} + jkzK \quad c\mu \frac{2}{r}\frac{\ddot{y}Hz}{\ddot{y}\ddot{y}}, \quad (8)$$

Dabei werden der Einfachheit halber einige Symbole eingeführt, und zwar die folgenden:

$$\begin{aligned} K_{c\mu}^2 &= \ddot{y}r^2 k_0^2 \mu\ddot{y}k_z^2, \\ F &= \ddot{y}r f_0^2 k^2 \ddot{y}, \\ K_C^4 &= K_{c\mu}^2 \ddot{y}F, \\ A1 &= \frac{Fk_z^2}{\ddot{y}\ddot{y}0\ddot{y}r f}, \\ A2 &= \frac{K_c^2 + k_z^2 K_{c\mu}^2}{\ddot{y}\ddot{y}0\ddot{y}r f}. \end{aligned} \quad (9)$$

Wir möchten den Leser daran erinnern, dass die übrigen Parameter in den in dieser Arbeit verwendeten Gleichungen in Referenz [1] definiert wurden . Beachten Sie, dass die Längskomponenten die folgenden gekoppelten Gleichungen erfüllen:

$$\frac{1}{R}\frac{\ddot{y}}{\ddot{y}r} R\frac{\ddot{y}}{\ddot{y}r} + \frac{1}{R^2}\frac{\ddot{y}}{\ddot{y}\ddot{y}\mu 0\mu rzK + \ddot{y}\ddot{y}2}\frac{2}{C} Hz = \ddot{y}j \frac{kzF}{A2} \frac{1}{R}\frac{\ddot{y}}{\ddot{y}r} R\frac{\ddot{y}}{\ddot{y}r} + \frac{1}{R^2}\frac{\ddot{y}}{2\ddot{y}\ddot{y}2} Ez, \quad (10)$$

$$\frac{1}{\ddot{y}rr}\frac{\ddot{y}}{\ddot{y}r} r\frac{\ddot{y}}{\ddot{y}r} + \frac{1}{R^2}\frac{\ddot{y}}{\ddot{y}\ddot{y}2} + \frac{K_C^2}{K2_{c\mu}} Ez = \frac{jFkz}{K2\, c\mu\ddot{y}\ddot{y}0\ddot{y}r f}\frac{1}{\ddot{y}rr}\frac{\ddot{y}}{\ddot{y}r} r\frac{\ddot{y}}{\ddot{y}r} + \frac{1}{R^2}\frac{\ddot{y}}{2\ddot{y}\ddot{y}2} Hz, \quad (11)$$

so dass im Allgemeinen ein Hybridmodus erforderlich ist. Beachten Sie, dass die rechten Seiten der Gleichungen (10) und (11) in der Analyse von Sakli [1] vernachlässigt wurden; daher ist seine Aussage, dass TEz- und TMz -Modi separat in einem metallischen zylindrischen Wellenleiter unterstützt werden können, der mit longitudinal magnetisiertem Ferrit gefüllt ist, falsch.

Die Entkopplung von Ez **von** Hz **tritt jedoch in zwei besonderen Fällen auf. Im ersten** Fall können die Gleichungen (10) und (11) in zwei unabhängige Gleichungen von Ez **und** Hz **aufgeteilt werden , wenn die Magnetisierung gleich Null ist, d. h. wenn wir**

$$\ddot{y} = 0. \quad (12)$$

Machine Translated by Google

In diesem Fall werden die Gleichungen (10) und (11)

$$\frac{1}{\ddot{y}_{rr}}\frac{\ddot{y}}{\ddot{y}_{r}}\,r\frac{\ddot{y}}{\ddot{y}_{r}} + \frac{1}{R^{2}}\frac{\ddot{y}}{\ddot{y}\ddot{y}2}\,\mathbf{2}\mu rz + \frac{}{\mu}\,f^{\,2}\,k\,0\mu\,\ddot{y}\,k\,\ddot{y}r\,\frac{\mathbf{2}}{z}\,Hz = \mathbf{0}, \quad \textbf{(13)}$$

$$\frac{1}{R}\frac{\ddot{y}}{\ddot{y}r}\,r\frac{\ddot{y}}{\ddot{y}r} + \frac{1}{R^{2}}\frac{\ddot{y}^{2}}{}\,\mathbf{2}\,k\,0\mu\,\ddot{y}\,k + \ddot{y}r\,f_{\ddot{y}\ddot{y}2}\,\frac{\mathbf{2}}{z}\,Ez = \mathbf{0}, \quad \textbf{(14)}$$

und daher können TEz- **und TMz** -Modi getrennt in einem metallischen zylindrischen **Wellenleiter unterstützt werden, der mit longitudinal „unmagnetisiertem“ Ferrit gefüllt ist. Im zweiten Fall tritt Entkopplung auf, wenn wir die azimutalen Modi betrachten, d. h. wenn wir**

$$kz = \mathbf{0}. \quad \textbf{(15)}$$

Setzen wir in den Gleichungen (10) und (11) die Ausbreitungskonstante gleich Null, erhalten wir

$$\frac{1}{R}\frac{\ddot{y}}{\ddot{y}r}\,R\frac{\ddot{y}}{\ddot{y}r} + \frac{1}{R^{2}}\frac{\ddot{y}}{\mathbf{2}\,\ddot{y}\ddot{y}2} + \ddot{y}r\,f\,\mu rzk\,\frac{\mathbf{2}}{\mathbf{0}}\,Hz = \mathbf{0}, \quad \textbf{(16)}$$

$$\frac{1}{\ddot{y}_{rr}}\frac{\ddot{y}}{\ddot{y}_{r}}\,r\frac{\ddot{y}}{\ddot{y}_{r}} + \frac{1}{R^{2}}\frac{\ddot{y}}{\ddot{y}\ddot{y}2}\,\mathbf{2}\,\ddot{y}r\,f + \frac{}{\mu}\,\mu^{\,2\,2\,\ddot{y}\,\ddot{y}}\,k_{\mathbf{0}}^{\mathbf{2}}\,Ez = \mathbf{0}. \quad \textbf{(17)}$$

In diesem Fall treten die TEz- und TMz- Modi mit den Komponenten (Hz, Er , *Eÿ)* **bzw.** (Ez, Hr , *Hÿ)* **auf .**

3. Schlussfolgerungen

In dieser Arbeit haben wir zunächst den allgemeinen mathematischen Rahmen für die Ausbreitung elektromagnetischer Wellen durch einen metallischen zylindrischen Wellenleiter verwendet, der mit einem entlang seiner Länge magnetisierten Ferritmaterial gefüllt ist (wie in den Referenzen [2] und [3] beschrieben). Unsere Analyse hat gezeigt, dass dieser Systemtyp in den meisten Situationen keine klare Trennung der TEz- und TMz-Modi zulässt. Es gibt nur zwei spezielle Szenarien, in denen diese Trennung auftreten könnte. Daher sind die wichtigsten Ergebnisse, die Sakli in Referenz [1] präsentiert , ungültig.

Finanzierung: Diese Forschung erhielt keine externe Finanzierung.

Erklärung zur Datenverfügbarkeit: Die Daten, die die Ergebnisse dieser Studie stützen, sind im Artikel verfügbar.

Interessenkonflikte: Der Autor erklärt, dass keine Interessenkonflikte vorliegen.

Verweise

1. **Sakli, H. Zylindrischer Wellenleiter auf Ferritsubstrat, gesteuert durch ein extern angelegtes Magnetfeld.** Electronics **2021,** 10, **474.** [Querverweis]
2. **Kales, ML Moden in Wellenleitern mit Ferriten.** J. Appl. Phys. **1953,** 24, **604.** [CrossRef]
3. **Gamo, H. Die Faraday-Rotation von Wellen in einem kreisförmigen Wellenleiter.** J. Phys. Soc. Jpn. **1953, 8, 176.** [CrossRef]
4. **Moradi, A.; Bait-Suwailam, MM Kommentar zu: Verbesserte Kopplung von Licht aus Quellen unterhalb der Wellenlänge in eine hyperbolische Metamaterialfaser.** J. Light. Technol. **2024,** 42, **5435–5436.** [CrossRef]
5. **Moradi, A. Kommentar zu steuerbaren, mit Metamaterial beladenen Wellenleitern, die Rückwärts- und Vorwärtswellen unterstützen.** IEEE Trans. Antennas Propag. **2023,** 72, **3858–3859.** [CrossRef]
6. **Moradi, A.; Bait-Suwailam, MM Magnetostatische Wellen in metallischen rechteckigen Wellenleitern, gefüllt mit uniaxialen Medien mit negativer Permeabilität.** J. Appl. Phys. **2024,** 135, **153102.** [CrossRef]

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultieren.

Machine Translated by Google

Elektronik

MDPI

Artikel

Identifizierung des Strommasttyps und charakteristische Parameter Berechnung aus luftgestützten LiDAR-Daten

Shengxuan Zu 1,2, Linong Wang 1,2,*, Shaocheng Wu 1,2, Guanjian Wang 1,2 und Bin Song[1,2]

[1] Technisches Forschungszentrum des Bildungsministeriums für Blitzschutz und Erdungstechnologie, Fakultät für Elektrotechnik und Automatisierung, Universität Wuhan, Wuhan 430072, China; 2022202070038@whu.edu.cn (SZ); wushaocheng@whu.edu.cn (SW); wangguanjian@whu.edu.cn (GW); 00007609@whu.edu.cn (BS)

[2] Fakultät für Elektrotechnik und Automatisierung, Wuhan-Universität, Wuhan 430072, China * Korrespondenz: wangln@whu.edu.cn; Tel.: +86-138-7120-8599

Zusammenfassung: Die Rekonstruktion dreidimensionaler (3D) Modelle von Stromanlagen spielt bei der Weiterentwicklung digitaler Zwillingsstromnetze eine immer wichtigere Rolle. Um ein hochpräzises Modell zu rekonstruieren, ist es entscheidend, vor der Modellierung den Masttyp und seine erforderlichen Parameterinformationen Diese Studie schlägt eine verbesserte Methode zur Identifizierung von Masttypen auf der Grundlage von Ähnlichkeitsmessungen und eines linear transformierten Datensatzes vor. Zunächst wird die Identifizierung von Punktwolken mithilfe der Mastformkurve vereinfacht. Anschließend wird die Ähnlichkeit zwischen der Kurve und den Kurven innerhalb des Datensatzes mithilfe einer Ähnlichkeitsmessung ausgewertet, um den Masttyp zu bestimmen. Es wird eine neuartige Methode zur Berechnung der charakteristischen Parameter der Mastpunktwolken vorgeschlagen. Die horizontalen und vertikalen Verteilungseigenschaften der Mastpunktwolken werden analysiert, um wichtige Segmentierungspositionen basierend auf ihren Typen zu identifizieren. Merkmalspunkte werden aus wichtigen Segmentierungspositionen abgeleitet, um die charakteristischen Parameter zu berechnen. Abschließend werden die 3D-Modelle der Masten auf der Grundlage der berechneten Werte rekonstruiert. Die experimentellen Ergebnisse zeigten, dass die Hausdorff-Distanz im Vergleich zu anderen Ähnlichkeitsmessungen die beste Wirkung als Ähnlichkeitsmessung unter Verwendung des linear transformierten Datensatzes hatte, mit einem Gesamtbewertungs-F-Score von 86,4 %. Der maxi

Schlüsselwörter: luftgestütztes LiDAR; Strommast; Ähnlichkeitsmessung; 3D-Rekonstruktion

check for updates

Zitat: Zu, S.; Wang, L.; Wu, S.; Wang, G.; Song, B. Strommasttyp Identifikation und Charakteristik Parameterberechnung aus der Luft LiDAR-Daten. Electronics **2024,** 13, **3032. https://doi.org/10.3390/ Elektronik13153032**

Wissenschaftlicher Herausgeber: Ahmed Abu-Siada

Empfangen: 17. Juni 2024
Überarbeitet: 22. Juli 2024
Akzeptiert: 30. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

Mit dem rasanten Wachstum der Volkswirtschaft steigt auch die Nachfrage nach grundlegenden Energiequellen wie Elektrizität kontinuierlich an. Hochspannungsleitungen, die als wesentliche Infrastruktur für die Stromübertragung über große Entfernungen dienen, spielen eine entscheidende Rolle für die nationale Wirtschaftsentwicklung und die tägliche Produktion [1–3]. Eine regelmäßige Überwachung und Wartung der Stromübertragungskorridore ist unerlässlich, um den sicheren und stabilen Betrieb des Stromnetzes zu gewährleisten. Die herkömmliche Inspektion von Übertragungsleitungen wird in der Regel von Leitungsinspektoren durchgeführt, die die Stromanlagen visuell oder mit Handgeräten inspizieren und auf der Grundlage ihrer Erfahrung potenzielle Gefahren identifizieren. Aufgrund verschiedener Einschränkungen und des Fehlens dreidimensionaler (3D) Daten entlang der Leitungen sind manuelle Inspektionen jedoch mit einem erheblichen Arbeitsaufwand, geringer Präzision und potenziellen

In den letzten Jahren wurde die luftgestützte Light Detection and Ranging (LiDAR)-Technologie häufig bei Strominspektionen eingesetzt, da sie die schnelle und präzise Erfassung dichter 3D-Punktwolken innerhalb von Stromübertragungskorridoren ohne Einschränkungen durch Licht und Gelände ermöglicht **[8–10]. Als wesentlicher Bestandteil von Übertragungsleitungen spielt der Strommast eine entscheidende Rolle bei der Gewährleistung der Sicherheit von Hochspannungsleitungen. Das rekonstruierte 3D- Modell des Mastes kann grundlegende Daten und Modellunterstützung für die Durchführung multiphysikalischer Feldsimulationsanalysen, die Simulatio**

Machine Translated by Google

Betriebsmethoden. Daher ist es notwendig, eine genaue und effiziente Methode zu entwickeln, um Pyloninformationen aus Punktwolken für die Rekonstruktion von 3D-Modellen zu extrahieren.

Derzeit konzentriert sich die Forschung zu Mastpunktwolken vor allem auf die Segmentierung von Mastpunkten aus den Punktwolken, die durch luftgestütztes LiDAR erfasst wurden [11–16]. Und es gibt viele Methoden zur Mastrekonstruktion auf Grundlage von Punktwolkendaten, die als datengesteuert [17], modellgesteuert [10] und hybridgesteuert [18] klassifiziert werden können. Der datengesteuerte Ansatz ist im Allgemeinen eine Bottom-up-Strategie. Er verarbeitet die Daten direkt, ohne dass die Eigenschaften des rekonstruierten Objekts vorausgesetzt werden müssen. Han et al. [17] schlugen eine datengesteuerte Methode zur Modellierung von Strommasten vor. Bei dieser Methode wurden Strommastpunktwolken lokalisiert und extrahiert, indem die Verbindungspunkte von Linienpaaren genutzt wurden. Das 3D-Modell des Strommastes wurde entsprechend der 3D-Linienfunktion erstellt, die aus der Konturverfolgung des Binärbildes gewonnen wurde. Diese Methode erfordert eine hohe Datenqualität, und es ist schwierig, die Maststruktur zu rekonstruieren, wenn die erhaltenen **Punktwolken viele Rauschpunkte enthalten . Verglichen mit dem datengesteuerten Ansatz verfolgt der modellgesteuerte Ansatz eine Top-down-Strategie und erfordert eine vorab vervollständigte Modellbibliothek. Li et al. [10] haben den Pylon in drei relativ einfache Teile unterteilt: Fuß, Körper und Kopf. Der Kopf wurde rekonstruiert, indem das entsprechende Modell aus der vorgefertigten Modellbibliothek gesucht wurde , und der Körper wurde rekonstruiert, indem die Schnittlinien der angepassten Seitenebenen berechnet wurden . Das Experiment deutete darauf hin, dass mit diesem Ansatz eine automatische 3D-Modellierung von Kopf und Körper des Pylons effektiv erreicht werden kann. Die Rekonstruktion des Pylonfußes erforderte jedoch eine interaktive Operation. Da es schwierig ist, die Rekonstruktionsanforderungen komplexer Objekte nur mit daten- oder modellgesteuerten Methoden zu erfüllen, wurden hybride Methoden vorgeschlagen, die die beiden oben genannten Methoden kombinieren. Die Methode wendet geeignete Strategien entsprechend den unterschiedlichen Anforderungen an die Strukturmodellierung an, wodurch die Modellierungsgenauigkeit verbessert werden kann. Zhou et al. [18] haben den Pylon in Kopf und Körper unterteilt. Sie rekonstruierten den Pylonkörper mithilfe einer datengesteuerten Strategie und den Kopf mithilfe einer modellgesteuerten Strategie mithilfe**

Um die Probleme der oben genannten Modellierungsmethoden zu lösen, werden in **diesem Dokument** die Pylonentypen und erforderlichen **Parameterinformationen aus 3D-Punktwolken verwendet, um die Pylone zu rekonstruieren .**

Die bestehenden Methoden zur Identifizierung der Masttypen werden hauptsächlich in regelbasierte Methoden und Methoden des maschinellen Lernens eingeteilt **[19]. Bei der ersten Methode werden die Merkmale aus den Mastpunktwolken extrahiert und dann die Masttypen anhand der Unterschiede in den Merkmalen identifiziert. Qiao et al. [20] schichteten die Mastkopfpunktwolken und berechneten dann die vertikale Füllrate für jede einzelne Schicht. Sie klassifizierten Masten in zwei Typen basierend auf der Position der Schicht mit der höchsten Füllrate. Chen et al. [21] segmentierten die Mastkopfpunktwolken basierend auf Punktverteilungsmerkmalen und projizierten sie dann auf die Y0Z0- Ebene, um ein Bild zu erstellen. Das Bild der Mastkopfkontur wurde durch Integration der Bildverarbeitungsmethode gewonnen. Schließlich wurde der Masttyp basierend auf der Anzahl der Pixel innerhalb der Kontur bestimmt. Obwohl die beiden oben genannten Methoden Masten unterscheiden können, ist die Anzahl der identifizierbaren Masttypen sehr begrenzt. Silva F. et al. [22] schlugen eine Klassifizierungsmethode basierend auf Ähnlichkeit vor. Sie nutzten Messgrößen für die Punktwolkendistanz, um die Ähnlichkeit zwischen den Punktwolken der Masten und grundlegenden Re Diese Methode erfordert die Verarbeitung großer Datenmengen und reagiert empfindlich auf Schwankungen in der Dichte der Stichprobenpunkte. Bei der zweiten Methode wird die Ermittlung der Masttypen auf Grundlage eines maschinellen Lernalgorithmus realisiert. Zhou et al. [18] definierten zunächst eine 3D-parametrisierte Modellbibliothek von Mastköpfen. Dann wurden die Mastkopftypen mit dem Formkontextalgorithmus ermittelt und die relevanten Parameter der Mastköpfe mithilfe eines Simulated-Annealing-Algorithmus geschätzt. Chen et al. [23] extrahierten Merkmale auf Grundlage von Punkthöhenhistogrammen und Frontalprojektion und verwendeten schließlich die Klassifizierungsmethode Support Vector Machine (S**

Machine Translated by Google

Algorithmus wurde verwendet, um die relevanten Parameter der Pylonenköpfe zu schätzen. Chen et al. [23] extrahierten Merkmale basierend auf Punkthöhenhistogrammen und Frontalprojektion und verwendeten schließlich die Klassifizierungsmethode Support Vector Machine (SVM), um Kopfmerkmalsvektorproben zu trainieren und zu klassifizieren. Wang et al. [24] projizierten innere und äußere Konturpunkte in gerasterte Bilder, um Histogramm orientierter Gradienten (HOG) Merkmale zu extrahieren. Dann projizierten Wang et al. [24] innere und äußere Ko

verwendeten diese als Eingaben für den SVM-Klassifikator zur Typidentifikation. Wenn jedoch das SVM

Da der Algorithmus große Datensätze verarbeitet, kann das Training aufgrund seines höheren SVM

verarbeitet, ist die Rechenkomplexität und der Speicherbedarf höher.

Zusammenfassend lässt sich sagen, dass die derzeitigen Methoden zur Identifizierung von Pylontypen unter

Erkennungseffektivität, Anfälligkeit für Schwankungen in der Stichprobendichte, hohe Zusammenfassend

leiden die aktuellen Methoden zur Identifizierung von Masttypen unter begrenzten

Rechenkomplexität und lange Trainingszeiten. Darüber hinaus gibt es keine geeignete Methode zur Berechnung der Parameter der Pylon-Punktwolken. Zur Lösung der

gibt es keine geeignete Methode zur Berechnung der Parameter der Pylon-Punktwolken. Zur Lösung der oben genannten Probleme gibt es keine geeignete Methode

Probleme, diese Studie schlägt eine verbesserte Methode zur Identifizierung von Pylontypen vor, die auf der

Ähnlichkeitsmessung und eine neuartige Methode zur Berechnung charakteristischer Parameter von

Punktwolken. Fehler! Referenzquelle nicht gefunden. Abbildung 1 zeigt die Ähnlichkeitsmessung des

Diagramm dieses Dokuments. Zunächst werden die Punktwolken mit Nullmittelwert, Punktwolken

Normalisierung, Verschiebung und Umleitung, bevor die Pylon-Informationen erhalten werden

die Masttypen identifiziert werden. Zweitens werden Masttypen identifiziert, indem die

Punktwolken generiert und Formkurven berechnet werden und

werden. Dann Kurven innerhalb des Datensatzes. Dann werden die charakteristischen Pylonpositionen bestimmt,

Anzahl, Dichte, Füllrate und Formparameter der berechneten Punktwolken bestimmt. Im

Merkmalspunkte aus den Punktwolken abgeleitet, um

berechnen. Abschließend werden die Pylon-3D-Modelle rekonstruiert auf

Parameter. Abschließend werden die 3D-Modelle der Masten auf Basis der berechneten Werte rekonstruiert.

Werte.

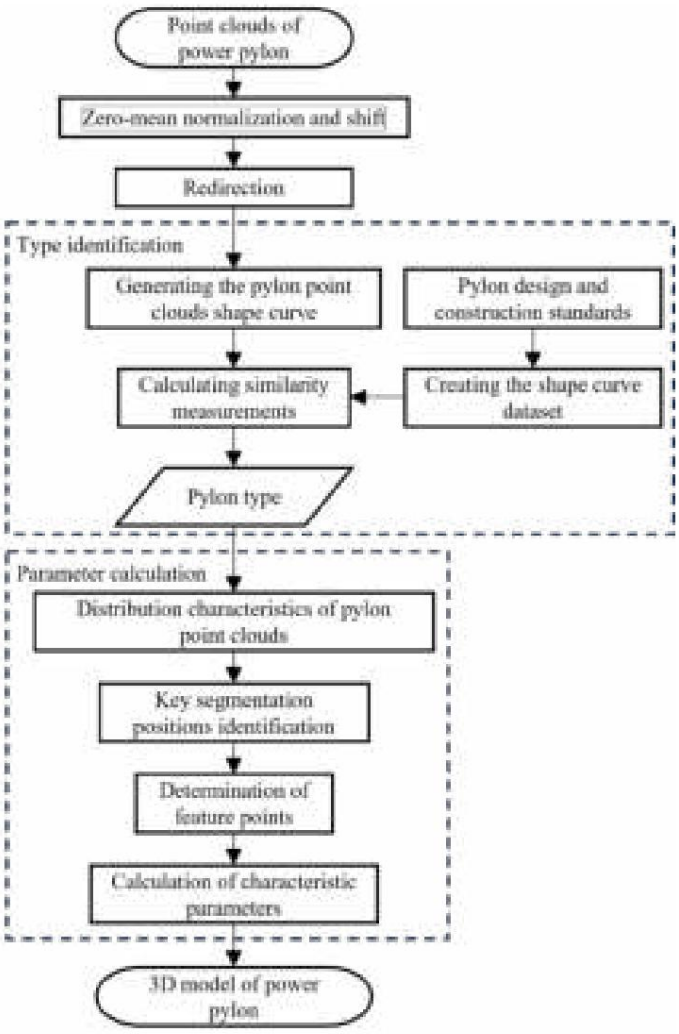


Abbildung 1. Gesamtflussdiagramm der Typidentifizierung und Parameterberechnung.

Machine Translated by Google

Die wichtigsten Neuerungen und Beiträge dieser Studie sind wie folgt: 1.

Diese Studie schlägt eine verbesserte Methode zur Identifizierung von Masttypen vor, die auf der Ähnlichkeitsmessung und einem linear transformierten Datensatz basiert. Durch den Vergleich der Auswirkungen von vier Ähnlichkeitsmessungen auf die Identifizierung von Masttypen wird eine Ähnlichkeitsmessung erhalten, die für die Identifizierung verschiedener Masttypen besser geeignet i

2. Zum ersten Mal wird eine neuartige Methode zur automatischen Berechnung charakteristischer Parameter unter Verwendung von LiDAR-Daten aus der Luft vorgeschlagen. Sie kann die spezifischen **Informationen des Mastes effizient extrahieren und so die hohe Genauigkeit der charakteristischen Parameter gewährleisten.**

3. Die 3D-Modelle der vier Masttypen werden auf Grundlage der identifizierten Masttypen und berechneten Parameter rekonstruiert, wodurch die tatsächliche Struktur der Masten genau wiedergegeben werden kann.

2. Identifizierung des Pylon-Punktwolkentyps basierend auf Ähnlichkeitsmessungen

Beim Erfassen von Mastpunktwolken mithilfe von luftgestütztem LiDAR werden die meisten Laserpunkte nicht durch vertikales Scannen der Zieloberfläche, sondern durch schrägen Einfall erfasst. Daher befinden sich einige der Mastpunktwolken nicht nur auf der Spitze des Masts, sondern auch an den Seiten des Masts. Da außerdem keine Flugrouten für unbemannte Luftfahrzeuge (UAV) und keine spezifischen räumlichen geografischen Informationen zu den Masten vorliegen, ist es unmöglich zu bestimmen, zu welcher Seite die Punkte Darüber hinaus ist die Struktur eines Mastes im Allgemeinen komplex. Würde man alle Mastpunktwolken nutzen, um den Masttyp im 3D-Raum zu bestimmen, wäre es aufgrund der oben genannten Umstände schwierig, dieses Ziel zu erreichen.

Um dieses Problem effektiver anzugehen, führt dieses Dokument Pylonformkurven ein, wodurch die Pylontypidentifizierung im 3D-Raum auf die Identifizierung von Kurven im zweidimensionalen (2D) Raum vereinfacht wird. Diese Kurve kann die Form des Pylons genau wiedergeben. Die Formen verschiedener Pylontypen weisen erhebliche Unterschiede auf, und die Formen von Pylonen, die zum selben Typ, aber zu unterschiedlichen Modellen gehören, sind im Wesentlichen ähnlich. Schließlich werden die Ähnlichkeitsmessungen kombiniert, um die Ähnlichkeit zwischen der aus den Pylonpunktwolken abgeleiteten Formkurve und den Kurven innerhalb des Datensatzes zu bestimmen, und dann wird der Typ des Pylons bestimmt.

2.1. Vorverarbeitung von Punktwolken

Da die von luftgestütztem LiDAR erfassten Pylon-Punktwolken im 3D-Raum beliebig ausgerichtet werden können, führen wir zur Vereinfachung der Berechnung der Ähnlichkeitsmessung und der anschließenden Verarbeitung eine Nullmittelwertnormalisierung der Pylon-Punktwolkendaten mithilfe von Gleichung (1) durch. Anschließend transformieren wir die Koo

Verschieben Sie sie entlang der positiven Z-Achse und erhalten Sie die aktualisierten Koordinaten (x ÿ y ÿ , ÿ z **).**

$$\begin{aligned} \ddot{y} \quad {}^{\prime}X &= x \ddot{y} x0 \\ \ddot{y} \quad j^{\ddot{y}} &= y \ddot{y} y0 \\ \ddot{y} \quad {}^{\ddot{y}}z &= z \ddot{y} z0 \ddot{y} \min(z \ddot{y} z0) \end{aligned} \tag{1}$$

Dabei stellen **x0, y0 und z0 die Koordinaten der zentralen Position der anfänglichen Punktwolken dar; und** x **und** z **stellen die Koordinaten der Punktwolken nach der Normalisierung und Verschiebung auf den Mittelwert Null dar.**

Die Pylonstruktur ist im Allgemeinen symmetrisch, und die Erfassung der Formmerkmale des Pylons aus der Vorderansicht ist effektiver. Daher ist es erforderlich, die Punktwolken des Pylons um einen bestimmten Winkel ÿ zu drehen, bevor die Pylonformkurve generiert wird, wobei die horizontale Richtung senkrecht zur X-Achse ausgerichtet wird.

Die horizontale Richtung des Pylons ist im Allgemeinen relevanter für seine obere Struktur . In dieser Studie werden Punktwolken mit Z-Koordinaten größer als H **ausgewählt und auf die XY-Ebene projiziert. Anschließend werden die Eigenwerte und Eigenvektoren der resultierenden projizierten Punktwolken mithilfe des Hauptkomponentenanalyse-Algorithmus Der erhaltene minimale Eigenwert entspricht dem Eigenvektor (v1, v2) senkrecht zur horizontalen Orientierung der Punktwolke. Und der Rotationswinkel** *ÿ* **wird berechnet durch**

Machine Translated by Google

die XY-Ebene. Anschließend werden die Eigenwerte und Eigenvektoren der resultierenden projizierten Punktwolken mithilfe des Hauptkomponentenanalyse-Algorithmus (PCA) berechnet. Der erhaltene minimale Eigenwert entspricht dem Eigenvektor (v1, v2) senkrecht zur horizontalen Ausrichtung der Punktwolke. Und der Drehwinkel $\ddot{y}$ wird mit Gleichung (2) berechnet [20]. Anschließend wird die Koordinatentransformation nach Gleichung (3) durchgeführt um die Koordinaten (x, y, z) der gedrehten Punktwolken zu erhalten. Abbildung 2 zeigt die gefundenen umgelenkten Pylone.

$$\ddot{y} = \arccos\left(\frac{\text{Version 1}}{\sqrt{v_1^2 + v_2^2}}\right) \quad (2)$$

$$\begin{aligned} x'' &= x \cos(\ddot{y}) + y \sin(\ddot{y}) \\ y'' &= -x \sin(\ddot{y}) + y \cos(\ddot{y}) \\ z'' &= z \end{aligned} \quad (3)$$

wobei x'', y'' und z'' die Koordinaten der gedrehten Punktwolken darstellen.

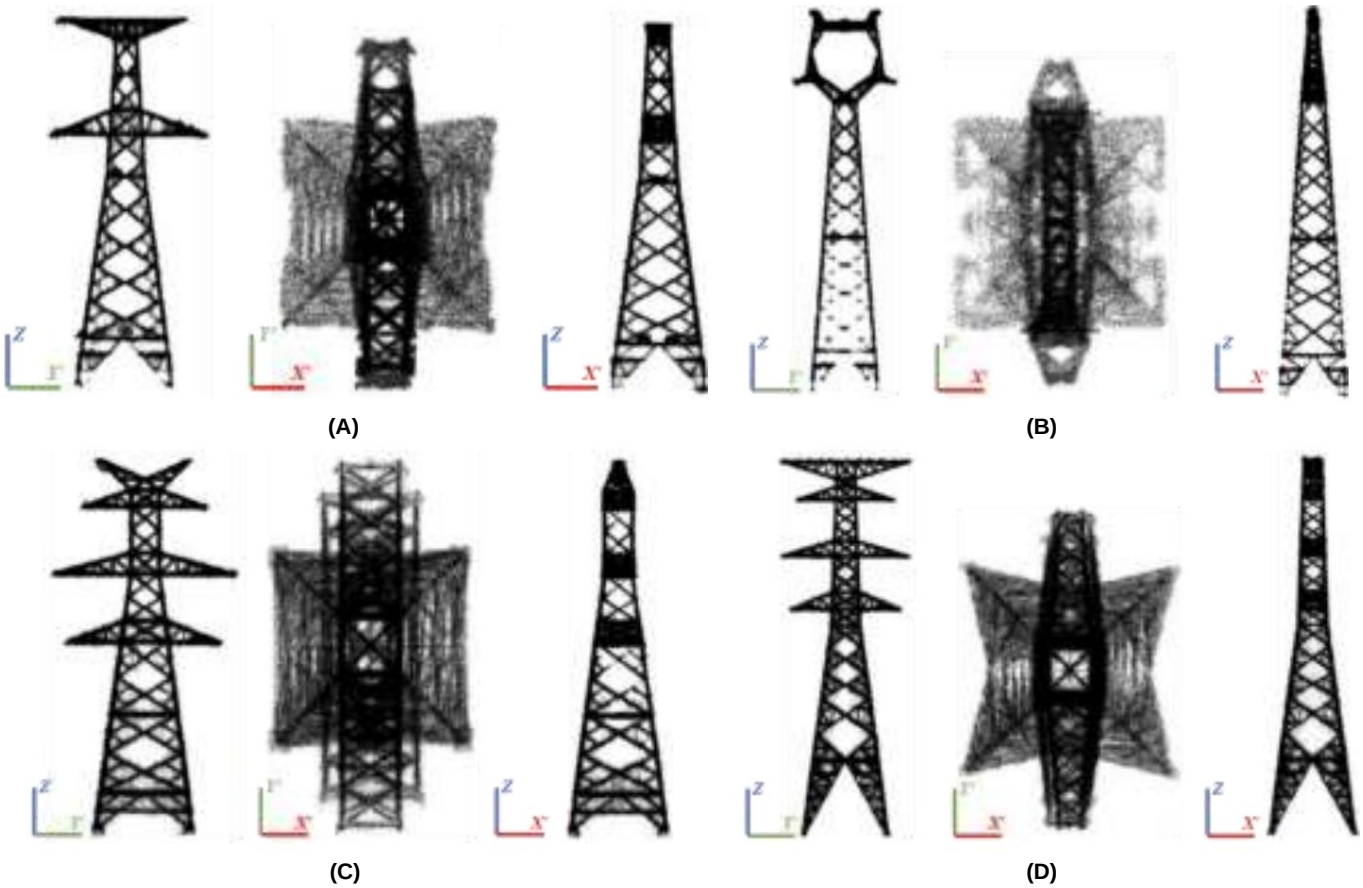


Abbildung 2. Die Projektionen der umgeleiteten Pylone auf die Ebenen YŷZ, XŷYŷ und XŷZ. (a) Pylone vom Typ a. (b) Pylon Typ B. (c) Pylon Typ C. (d) Pylon Typ D.

Abbildung 2. Die Projektionen der umgeleiteten Pylonen auf den YŷZ- , XŷYŷ und XŷZ- Ebenen. (a) Pylone vom Typ a. (b) Pylon Typ B. (c) Pylon Typ C. (d) Pylon Typ D.

2.2. Generieren der Pylon-Punktwolken-Formkurve

2.2. Generieren der Pylon-Punktwolken-Formkurve

Die umgelenkten Pylonenpunktwolken werden auf die YŷZ-Ebene projiziert und dann vertikal entlang der Z-Achse in einem bestimmten Abstand von h1 geschichtet. Die Grenzpunkte jeder Schicht entlang der Z-Achse in einem bestimmten Abstand von h1 geschichtet. Die Grenzpunkte des Pylons. Diese Punkte bilden die Kurve der Gesamtform des Pylons.

Unter Berücksichtigung der Symmetrie der Pylonstruktur kann die Hälfte der Randpunkte ausgewählt und in Folge verbunden werden. Schließlich gleichmäßig verteilte diskrete Kurven, die sind die Formkurven der Pylon-Punktwolke, die erhalten werden, wie in Abbildung 3 gezeigt.

ausgewählt und in Folge verbunden werden. Schließlich gleichmäßig verteilte diskrete Kurven, die die Formkurven der Pylon-Punktwolke, die erhalten werden, wie in Abbildung 3 gezeigt.

[20]. Anschließend wird die Koordinatentransformation nach Gleichung (3) durchgeführt um die Koordinaten (x, y, z) der gedrehten Punktwolken zu erhalten. Abbildung 2 zeigt die gefundenen umgelenkten Pylone.

wobei x'', y'' und z'' die Koordinaten der gedrehten Punktwolken darstellen.

Machine Translated by Google

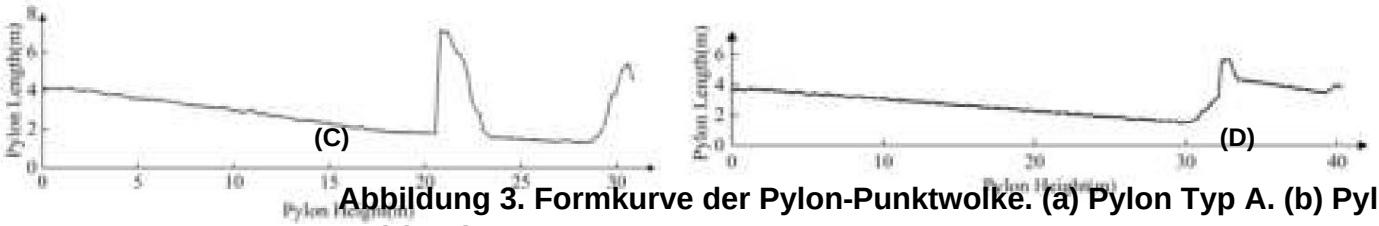


Abbildung 3. Formkurve der Pylon-Punktwolke. (a) Pylon Typ A. (b) Pylon Typ B. (c) Pylon Typ D.

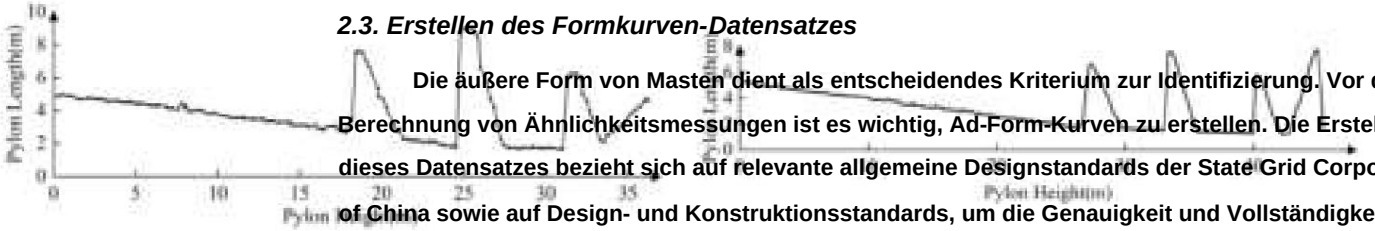


Abbildung 3. Formkurve der Pylon-Punktwolke. (a) Pylon Typ A. (b) Pylon Typ B. (c) Pylon Typ D.
(d) Pylone vom Typ d.

2.3. Erstellen des Formkurven-Datensatzes

Die äußeren Formen der Masten dienen als entscheidendes Kriterium zur Identifikation des Masttyps.

Die äußere Form eines Mastes dient als entscheidendes Kriterium zur Identifizierung des Masttyps.

Vor der Berechnung von Ähnlichkeitsmessungen ist es wichtig, einen Datensatz der Pylonformkurven zu erstellen. Die Erstellung dieses Datensatzes bezieht sich auf relevante allgemeine Designstandards der State Grid Corporation of China sowie auf Design- und Konstruktionsstandards, wodurch die Genauigkeit und Vollständigkeit der Daten sichergestellt wird.

Derzeit enthält der Datensatz die Formkurven verschiedener gängiger Entwurfs- und Konstruktionsstandards, wodurch die Genauigkeit und Vollständigkeit der Daten sichergestellt wird. Basierend auf den allgemeinen Pylonparametern wie Höhe, Länge, Querarm

Basierend auf den allgemeinen Mastparametern, die von diesen Standards bereitgestellt werden, wie z. B. Abbildung 3.4 Formkurve der Mastpunktwolke (a) Mast Typ A; (b) Mast Typ B; (c) Mast Typ D; Querarmlänge und weitere Angaben im vertikalen und horizontalen Pylon Typ D.

Es werden Formkurven verschiedener Pylone im Datensatz generiert. Anschließend werden die

2.3. Erstellen des Formkurven-Datensatzes

Die äußeren Formen der Masten dienen als entscheidendes Kriterium zur Identifikation des Masttyps.

Die äußere Form eines Mastes dient als entscheidendes Kriterium zur Identifizierung des Masttyps.

Vor der Berechnung von Ähnlichkeitsmessungen ist es wichtig, einen Datensatz der Pylonformkurven zu erstellen. Die Erstellung dieses Datensatzes bezieht sich auf relevante allgemeine Designstandards der State Grid Corporation of China sowie auf Design- und Konstruktionsstandards, wodurch die Genauigkeit und Vollständigkeit der Daten sichergestellt wird. Basierend auf den allgemeinen Pylonparametern wie Höhe, Länge,

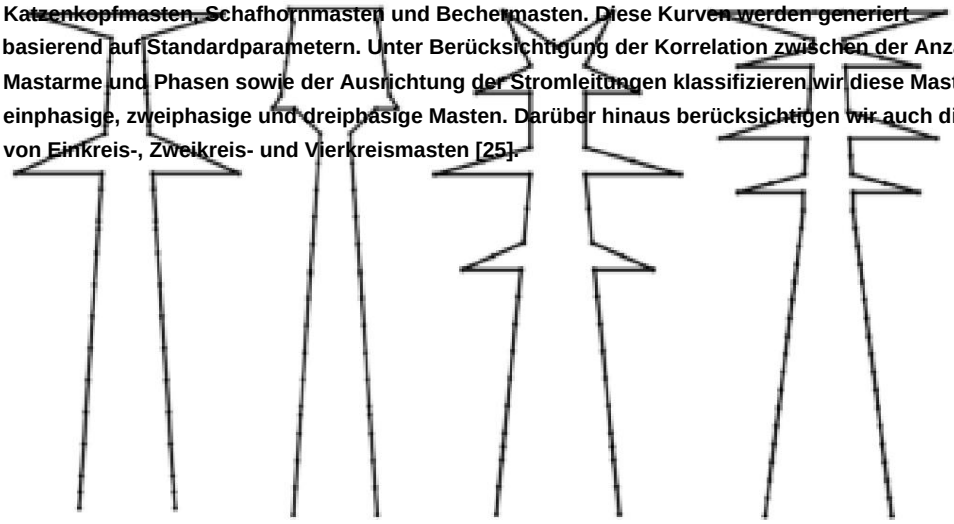
Derzeit enthält der Datensatz die Formkurven verschiedener gängiger Entwurfs- und Konstruktionsstandards, wodurch die Genauigkeit und Vollständigkeit der Daten sichergestellt wird. Basierend auf den allgemeinen Pylonparametern wie Höhe, Länge,

Die Anzahl der Mastarme und Phasen sowie die Ausrichtung der Stromleitung werden linear auf die gleiche Höhe transformiert, wie in Abbildung 4 dargestellt. Und um die

Masten in einphasige, zweiphasige und dreiphasige Masten. Weitere Quelle nicht gefunden. Abbildung 4. Und

Die Hälfte der Mastquerschnitte wird vor der Verwendung zu einer diskreten Kurve verbunden.

Derzeit enthält der Datensatz die Formkurven verschiedener gängiger Strommasten, wie Katzenkopfmasten, Schafhornmasten und Bechermasten. Diese Kurven werden generiert basierend auf Standardparametern. Unter Berücksichtigung der Korrelation zwischen der Anzahl der Mastarme und Phasen sowie der Ausrichtung der Stromleitungen klassifizieren wir diese Masten in einphasige, zweiphasige und dreiphasige Masten. Darüber hinaus berücksichtigen wir auch die Fälle von Einkreis-, Zweikreis- und Vierkreismasten [25].



Formkurven einiger Pylone im Datensatz.

Abbildung 4. Abbildung 4. Formkurven einiger Pylone im Datensatz.

Abbildung 4. Formkurven einiger Pylone im Datensatz.

Machine Translated by Google

Derzeit enthält der Datensatz die Formkurven verschiedener gängiger Strommasten, wie Katzenkopfmasten, Schafhornmasten und Bechermasten. Diese Kurven werden auf der Grundlage von Standardparametern generiert. Unter Berücksichtigung der Korrelation zwischen der Anzahl der Mastarme und Phasen sowie der Ausrichtung der Stromleitungen klassifizieren wir diese Masten außerdem in einphasige, zweiphasige und dreiphasige Masten. Darüber hinaus berücksichtigen wir auch die Fälle von Einkreis-, Doppelkreis- und Vierkreismasten [25].

2.4. Ähnlichkeitsmessungen

Die Formkurve eines Mastes kann verschiedene Informationen liefern, wie etwa die Form des Mastkopfes, die Masthöhe und die Anzahl der Querarme. Die sinnvolle Nutzung der Unterscheidungsmerkmale verschiedener Masttypen trägt dazu bei, den Masttyp genau zu bestimmen. Darüber hinaus ist die Auswahl von Ähnlichkeitsmessungen für die Identifizierung von Masttypen von entscheidender Bedeutung. Diese Studie wählt vier weit verbreitete Ähnlichkeitsbewertungsmessungen in den Bereichen Mustererkennung und künstliche Intelligenz aus: die Dynamic Time Warping (DTW)-Distanz, die FastDTW-Distanz, die diskrete Fréchet-Distanz und die Hausdorff- Distanz. Wir vergleichen ihre Wirksamkeit bei der Ide

(1) DTW-Distanz und FastDTW-Distanz: Der DTW-Algorithmus wird häufig zur Bewertung der Ähnlichkeit von Zeitsequenzen unterschiedlicher Länge verwendet und findet umfangreiche Anwendung im Bereich der Spracherkennung [26,27]. Sein Hauptansatz ist wie folgt.

Angenommen, es gibt zwei Zeitsequenzen, X **= {x1, x2, . . . ,** xi , . . . , xn} **und** Y **= {y1, y2, . . . ,** yj , . . . , **ym}. Die kumulative Distanzmatrix** D **von** X **und** Y **wird mithilfe der euklidischen Distanz erstellt. Ein Verzerrungspfad** W **in Matrix** D **wird so gefunden, dass die Summe der Elemente entlang des Pfades minimiert wird. Die minimale kumulative Distanz kann mithilfe von Gleichung (4) berechnet werden, indem sowohl die Monotonie- als auch die Kontinuitätsbeschränkungen erfüllt werde**

$$D(i, j) = d(xi , yj) + \min\{D(i \ddot{y} \mathbf{1}, j), D(i, j \ddot{y} 1), D(i \ddot{y} \mathbf{1}, j \ddot{y} 1)\} \quad \mathbf{1} \ddot{y} i \ddot{y} \mathbf{n}, \mathbf{1} \ddot{y} j \ddot{y} m \tag{4}$$

wobei d(xi , yj) **die Distanz zwischen** xi **und** yj darstellt . D(i, **j) stellt die kumulative Distanz zwischen den Zeitschritten** xi **und** yj dar .

Anhand der endgültigen minimalen kumulativen Distanz lässt sich die Ähnlichkeit der Sequenzen X und **Y** beurteilen . **Ein kleinerer Wert bedeutet, dass die beiden Sequenzen hinsichtlich ihrer Form eine größere Ähnlichkeit aufweisen; ein größerer Wert hingegen deutet auf eine geringere Ähnlich**

Wenn die Zeitsequenzen jedoch lang sind, beträgt die Rechenkomplexität zur Berechnung der DTW-Distanz zwischen den beiden Sequenzen O(nm), **was zu einer relativ geringen Algorithmuseffizienz führt. In diesem Fall wird der DTW-Algorithmus normalerweise durch Begrenzung des Pfadsuchbereichs, Datenabstraktion und Indizierung beschleunigt. FastDTW verwendet die ersten beiden Methoden, um DTW zu beschleunigen. Dieser verbesserte Algorithmus verringert die Zeitkomplexität von DTW effektiv auf** O(m). **FastDTW umfasst hauptsächlich drei Prozesse: Vergröberung, Projektion und Verfeinerung [28–30].**

(2) Diskrete Fréchet-Distanz: Die diskrete Fréchet-Distanz berücksichtigt sowohl die Form von Kurven als auch die Abfolge der Punkte entlang der Kurven. Sie ist ein Distanzmaß zur Bestimmung des Ähnlichkeitsgrades zwischen Kurven und wird in verschiedenen Bereichen eingesetzt, um die Ähnlichkeit zwischen parametrisierten Kurven zu messen [31,32]. Sie ist wie folgt definier

Liegen zwei Polygonzüge X **und** Y **vor, die aus m bzw. n Punkten bestehen, so wird zur Berechnung der diskreten Fréchet-Distanz zwischen** X **und Y zunächst die entsprechende Folge von Punktpaaren ermittelt.**

$$L = (xa1 , yb1),(xa2 , yb2), \cdot \cdot \cdot ,(xak , ybk) \tag{5}$$

Darunter ist **a1 = 1, b1 = 1,** ak **= m und** bk **= n. Und um die Reihenfolge der Punkte sicherzustellen, gilt** für **jedes** i **= 1, ...,** n **ai+1 =** ai **oder ai+1 =** ai **+ 1, bi+1 =** bi **oder bi+1 =** bi **+ 1. Dann berechnen wir die maximale Distanz zwischen entsprechenden Punktpaaren.**

$$\ddot{y}L\ddot{y} = \max_{i=1,\cdots,k} d(xai , ybi) \tag{6}$$

Machine Translated by Google

Die diskrete Fréchet-Distanz zwischen X **und** Y **ist wie folgt definiert:**

$$Df\ \mathbf{(X,Y)} = \min\{\ddot{y}L\ddot{y}\} \quad (7)$$

Mit Df (X, Y) lässt sich die Ähnlichkeit zwischen X **und** Y **beurteilen.** Mit abnehmendem Wert nimmt **die Ähnlichkeit der Formen beider Kurven zu.**

(3) Hausdorff-Distanz: Die Hausdorff-Distanz beschreibt die Ähnlichkeit zweier Teilmengen , indem sie den Abstand zwischen ihnen im Raum misst [33]. Wenn es zwei endliche Punktmengen X **und Y mit den Längen** n **bzw. m gibt, dann ist die bidirektionale Hausdorff-Distanz** Dh (X, **Y) dieser beiden Datensätze:**

$$\begin{matrix}\ddot{y} \\ \ddot{y} \\ \ddot{y}\end{matrix}\ \begin{matrix}\text{Dh } \mathbf{(X,Y)} = \max\{\text{dh (X,Y), dh (Y, X)}\}\ \text{dh } \mathbf{(X,Y)} = \\ \mathbf{maxx\ddot{y}Xminy\ddot{y}Y}\ \ddot{y}\ x\ \ddot{y}\ y\ \ddot{y}\ \text{dh (Y, X)} = \mathbf{maxy\ddot{y}Yminx} \\ \mathbf{\ddot{y}X}\ \ddot{y}\ y\ \ddot{y}\ x\ \ddot{y}\end{matrix} \quad (8)$$

wobei Dh (X, **Y) den Maximalwert zwischen** dh (X, **Y) und** dh (Y, **X) annimmt.** ÿ·ÿ **stellt die euklidische Distanz zwischen den Punktmengen** X **und Y** dar. dh (X, **Y) ist die maximal kürzeste Distanz vom Punkt in** X **zur** Y **-Menge.** dh (Y, X) stellt die maximal kürzeste Distanz vom **Punkt in** Y **zur** X- Menge **dar .**

Die Hausdorff-Distanz misst die Unähnlichkeit zweier Punktsätze und kann verwendet werden, um die Ähnlichkeit von X **und Y zu beurteilen. Ein kleinerer Wert von** Dh (X, **Y) weist auf eine größere Ähnlichkeit in der Form zwischen** X **und Y** hin.

3. Berechnung der charakteristischen Parameter von Pylon-Punktwolken In diesem

Abschnitt werden zunächst die Verteilungseigenschaften der Pylon-Punktwolken einschließlich Anzahl, Dichte und horizontaler Füllrate berechnet. Anschließend werden diese Eigenschaften und der Pylon-Typ verwendet, um die wichtigsten Segmentierungspositionen zu identifizieren. Schließlich werden basierend auf diesen Positionen Merkmalspunkte aus den Punktwolken abgeleitet, um die charakteristischen Parameter zu berechnen.

3.1. Verteilungseigenschaften von Pylon-Punktwolken Die auf die YÿZ

-Ebene **projizierten Pylon-Punktwolken werden vertikal geschichtet, um Histogramme der Verteilungseigenschaften zu erzeugen. Anschließend wird ein gleitendes Fenster verwendet, um Schichten zu identifizieren, die gleichzeitig sowohl die lokale maximale Anzahl von Punktwolken als auch die lokale maximale Punktwolkendichte erfüllen. Und die horizontale Füllrate dieser Schichten** wird berechnet. Schließlich werden aus den Schichten mit der **höchsten Füllrate wichtige Segmentierungspositionen identifiziert .**

Die Punktwolken werden entlang der Z-Achse mit einem festen Intervall h1 geschichtet. Die Anzahl der Pylonpunkte und die räumliche Größe jeder Schicht werden berechnet und verwendet, um den Verteilungsgrad der Punktwolken in jeder Schicht zu ermitteln, d. h. die Punktwolkendichte. Um die Anzahl der Pylonpunkte und die Punktwolkendichte vollständig zu berücksichtigen, werden die beiden Parameter standardisiert. Die nach der Verarbeitung erhaltene Summe De **der beiden Parameter wird als Verteilungskennwert der Pylonpunktwolken verwendet. Dann gleitet ein Fenster mit der Höhe h2** in einem festen Intervall **h1 vom Boden des Pylons nach oben . Wenn der** De- **Wert der Schicht in der Mitte des Fensters während des Gleitvorgangs größer ist als der** De **-Wert anderer Schichten im Fenster, wird die Schicht als diejenige betrachtet, die gleichzeitig sowohl die lokale maximale Anzahl von Punktwolken als auch die lokale maximale Punktwolkendichte erfüllt ,** wie in A

In der praktischen Analyse sind nicht alle Schichten mit den beiden oben genannten Eigenschaften Schlüsselschichten, die für nachfolgende Berechnungen erforderlich sind. Die obersten N1- Maximalwerte werden ausgewählt, um die Schichten hier weiter zu filtern. Anschließend können wichtige Schichten weiter bestimmt werden, indem die Füllrate jeder Schicht berechnet wird. Der konkrete Berechnungsprozess der Füllrate ist wie folgt: Wichtige Schichten werden in **N2 Gitter mit einem festen Abstand L1 entlang der** Yÿ- Achse **aufgeteilt , und der Anteil der Gitter mit den Punkten** n **an der Gesamtzahl der N2** Gitter wird **als Füllrate f definiert ,** wie in Abbildung **6 da**

Machine Translated by Google

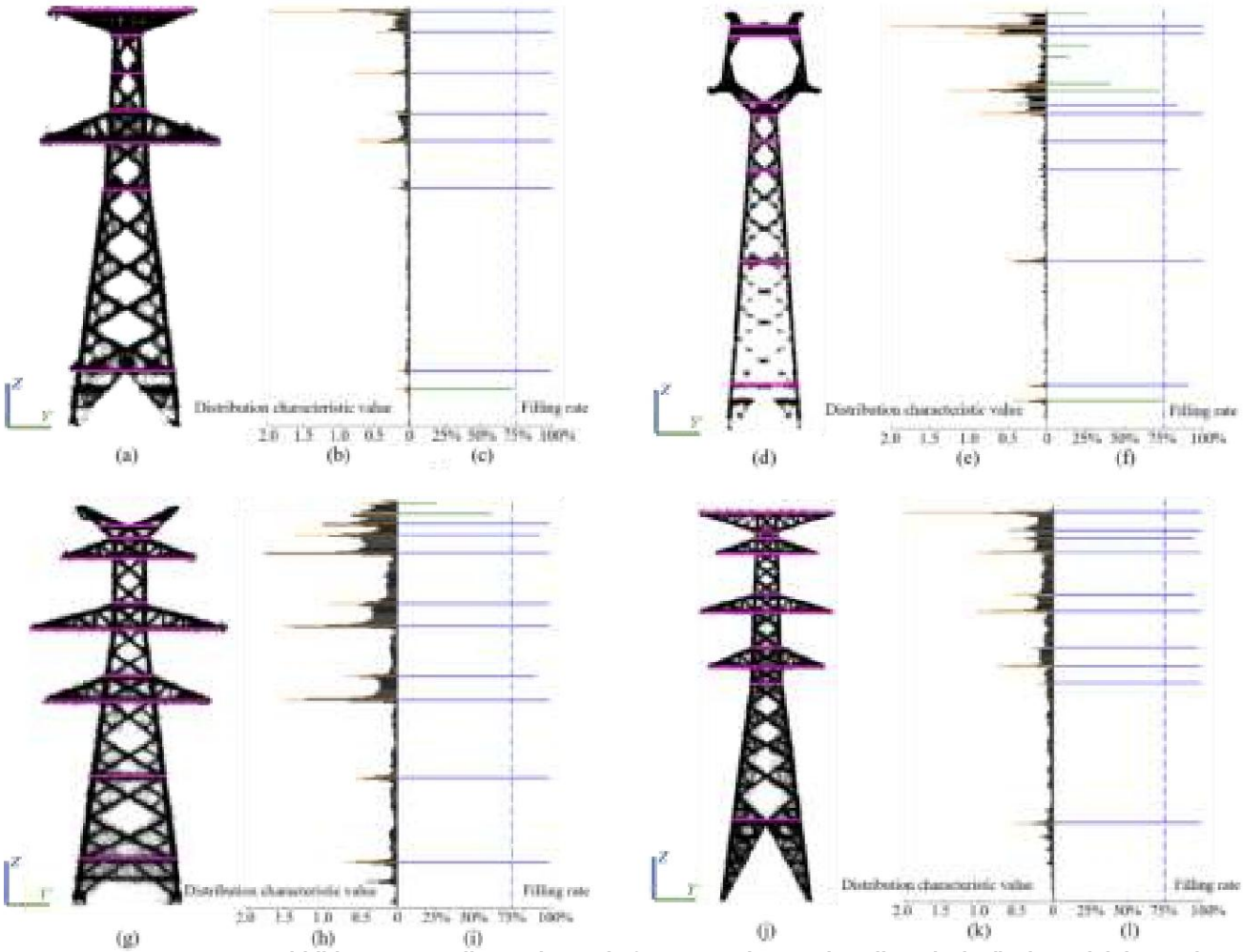


Abbildung 5. Verteilungseigenschaften von Pylon-Punktwolken. (a,d,g,j) Die Projektionen des Pylons auf der YŷZ-Ebene. (b,e,h,k) Verteilungseigenschaften von Histogrammen, und gelbe Linien sind Ebenen mit der gleichen Anzahl der lokalen Punktwolken. (c) die Dichte. (c,f,i,l) Füllratenhistogramme und blaue Linien sind die wichtigsten Segmentierungspositionen.

Abbildung 5. Verteilungseigenschaften von Pylon-Punktwolken. (a,d,g,j) Die Projektionen des Pylons auf der YŷZ-Ebene. (b,e,h,k) Verteilungseigenschaften von Histogrammen, und gelbe Linien sind Ebenen mit der gleichen Anzahl der lokalen Punktwolken. (c) die Dichte. (c,f,i,l) Füllratenhistogramme und blaue Linien sind die wichtigsten Segmentierungspositionen.

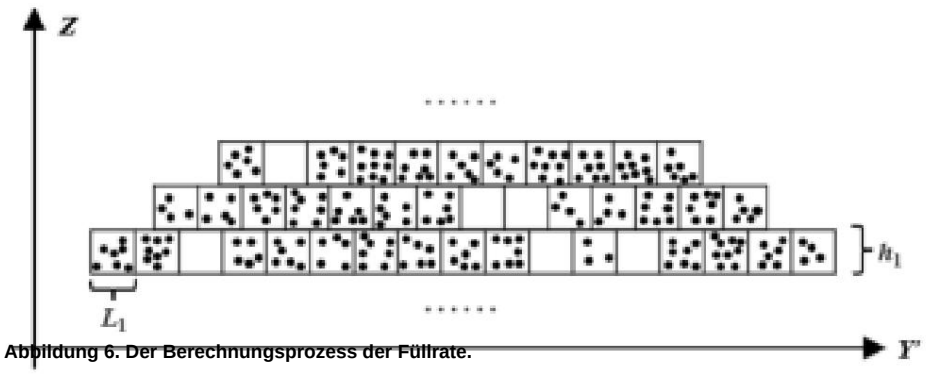


Abbildung 6. Der Berechnungsprozess der Füllrate.

Abbildung 6. Der Berechnungsprozess der Füllrate.

3.2. Identifizierung der Schlüsselsegmentierungsposition

Die Füllrate *f* der erwarteten Schlüsselsegmentierungsposition (gelbe Linien) wird berechnet 3.2. Identifizierung der Schlüsselsegmentierungsposition **im vorherigen Abschnitt. Wenn die Füllrate einer Schicht den vordefinierten Schwellenwert überschreitet *3.2. Schlüsselsegmentierungspositionsidentifikation*** Schlüsselsegmentierungsposition (gelbe Linien) wird berechnet **Füllrate *Tf*, die Schlüsselsegmentierungsposition *S* kann als die durchschnittliche *Z*- Koordinate aller im vorherigen** Abschnitt beschriebenen Elemente beschrieben werden. Wenn die Füllrate einer Schicht den vordefinierten Schwellenwert **innerhalb der Schicht, wie in Abbildung 5 dargestellt, wie in Abbildung 5 (blaue Linien) gefunden.** (blau Die Schlüsselsegmentierungsposition S kann als die durchschnittliche Z- **Koordinate aller bes** Tf, Linien). Für Masten vom Typ A und Masten vom Typ D können alle wichtigen Segmentierungspositionen im vorherigen Abschnitt liegen. Wenn die Füllrate einer Schicht die **vordefinierten Punkte innerhalb der Schicht beschrieben werden, wie in Abbildung 5 (blaue Linien) dargeste direkt mit der oben genannten Methode identifiziert. Für die anderen beiden Typen von *Tf* kann die Schlüsselsegmentierungsposition *S* jedoch als die durchschnittlichen *Z*-Koordinatenpunkte innerhalb der Schicht beschrieben werden, wie in Abbildung 5 (Fehler! Referenzquelle nicht für Linien) dargestellt. Für Pylone vom Typ A und Pylone vom Typ D können alle Schlüsselsegmentierungspositionen direkt mit der oben genannten Methode id**

Machine Translated by Google

Pylonen sind neben den wie oben beschrieben ermittelten Positionen zusätzlich noch die Ermittlung weiterer wichtiger Segmentierungspositionen erforderlich.

Für den Mast vom Typ B betrachten wir zunächst die Außenkontur. Aufgrund des Hohlprofils in der Schicht, die die Verbindungsisolatorpositionen der Masten vom Typ D enthält, können alle wichtigen Segmentierungspositionen direkt identifiziert werden.

Pylone, die Füllrate dieses Teils ist relativ gering und somit nicht die oben genannte Methode. Für die anderen beiden Pylontypen jedoch, neben den Positionen

Segmentierungsposition. In Anbetracht der Tatsache, dass die projizierte Form der Struktur wie oben beschrieben erhalten würde, sind zusätzliche Ansätze erforderlich, um andere Schlüssel zu bes Segmentierungspositionen.

Pylonen können dieser Art von Pylonen variieren in der Höhe in der XÿYÿ-Ebene, den SH-

***Gi* wird hier eingeführt, um wichtige Segmentierungspositionen besser identifizieren zu können.**

Für den Pylon Typ B betrachten wir zunächst die Außenkontur. Aufgrund der Anwesenheit von

Dieser Teil ist ein Hohlabschnitt in der Schicht, der die Position des Verbindungsisolators für den Typ enthält. Das Verhältnis der maximalen Projektionslänge der Punktwolken auf den Xÿ- und B-Pylonen, die Füllrate dieses Teils ist relativ gering und wird daher nicht als Schlüssel betracht Minimalwert. Die Summe aus G1 und der Fehlerkonstante Ce dient als t- Segmentierungsposition.

Wenn man bedenkt, dass die projizierte Form der Struktur über dem

Pylonkopf dieser Art von Pylon variiert mit der Höhe in der XÿY- **Ebenenparallelen t Ebene, der**

y-**der Formparameter. Ausgehend von G1** wird **jeder *Gi* nacheinander verglichen**

Gi **wird hier eingeführt, um wichtige Segmentierungspositionen besser zu identifizieren. Dieser Parameter definiert**

***TG*. Wenn *Gi* größer als *TG ist*,** ist **die Punktwolkenebene, die Gi-1** entspricht, **das Segment das** **Verhältnis der maximalen Projektionslänge der Punktwolken auf der** Xÿ- **und** Yÿ- **Achse zur** **Si-1 zwischen Pylonkopf und Pylonkörper, wie in Abbildung 7 dargestellt**Fehler! Re **Mindestwert. Die** **nicht gefunden. (rote Linie). Durch Beobachtung kann festgestellt werden, dass tw** TG. **Wenn** Gi TG **überschreitet ,** ist **die Punktwolkenebene, die Gi-1** entspricht , **die Segmentierung** **Positionen mit geringerer Füllrate oberhalb der Segmentierungsposition Si -1 zwischen Pylonkopf und Pylonkörper, wie in Abbildung 7 (rote Linie) dargestellt.**

Segmentierungspositionen, die die Verbindungslinien unterstützen. Ergänzung Durch Beobachtung kann festgestellt werden, dass die beiden Segmentierungspositionen mit geringerer Füllrate

Die Raten oberhalb der Segmentierungsposition sind die wichtigsten Segmentierungspositionen, die die Füllrate. Zusätzlich werden die beiden Segmentierungspositionen unterhalb der Segmentierungspos Segmentierungspositionen unterhalb der Segmentierungsposition mit Verbindungslinien zur

Anforderungen müssen nicht berücksichtigt werden, und diese Verarbeitung hat keinen Einfluss auf die Position mit Füllraten, die die Anforderungen erfüllen, müssen nicht berücksichtigt werden, u Parameterberechnung.

Die Verarbeitung hat keinen Einfluss auf die nachfolgende Parameterberechnung.

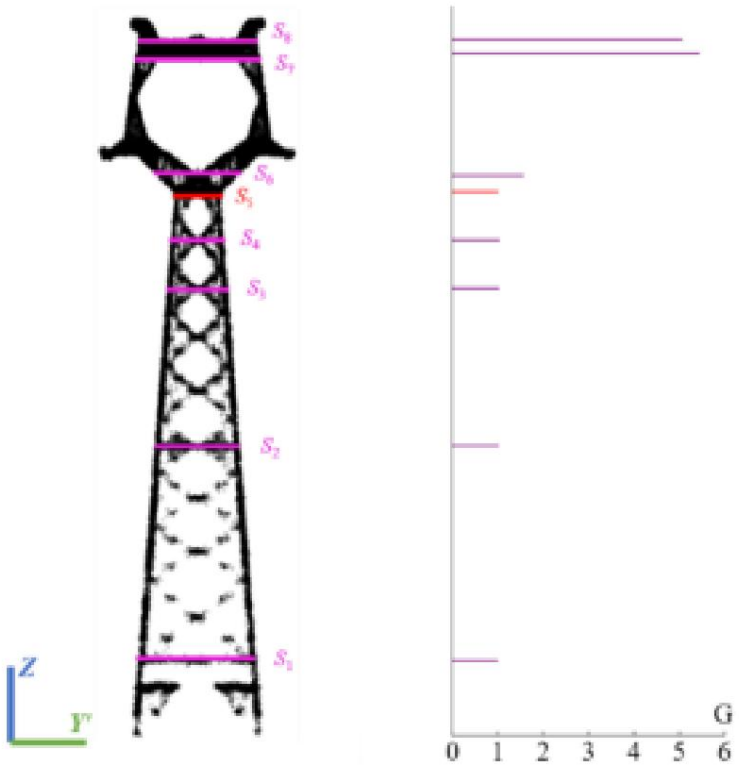


Abbildung 7. Formparameter des Pylons vom Typ B.

Abbildung 7. Formparameter des Pylons vom Typ B.

Vor der Modellierung der Masten des Typs B ist es wichtig, die Abmessungen ihrer

innere Hohlstrukturen. Aus Abbildung 6 ist ersichtlich, dass die Hohlstelle zwischen

Vor der Modellierung der Pylone vom Typ B ist es wichtig, die Schlüsselsegmentierungspositio S5 des Pylonkörpers und die Schlüsselsegmentierungsposition zu erhalten

innere Hohlstrukturen. Aus Abbildung 6S7 des Pylonenkopfes. Darüber hinaus gibt es an den Schlüsselsegmentierungspositionen innerhalb der Hohlstruktur im Vergleich zu den

vorgenannten Positionen relativ geringere Füllraten.

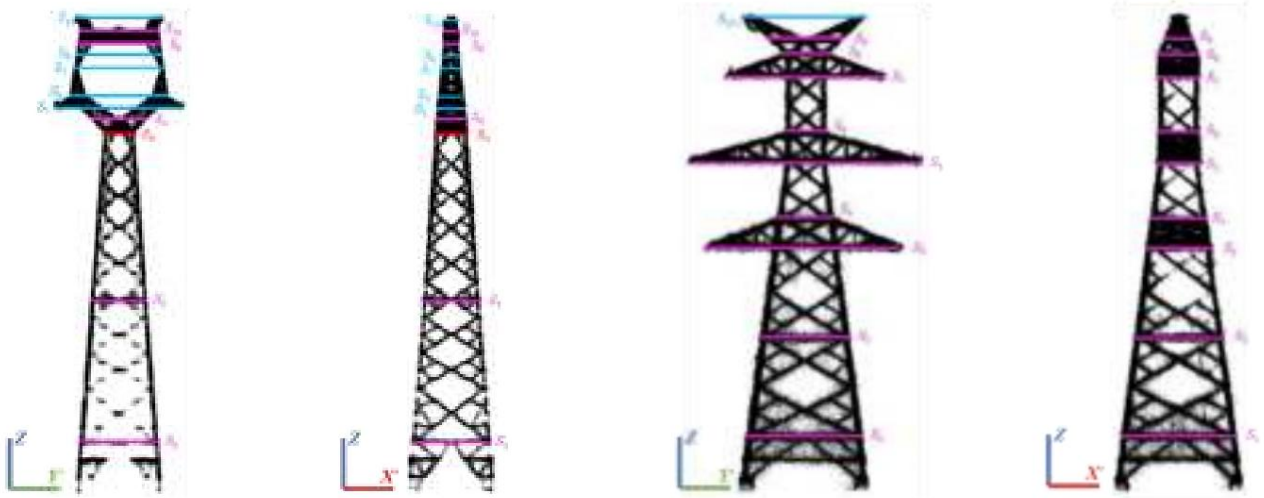
Durch die Kombination der Positionsbeziehung und der Verteilungseigenschaften der Punktwolken wird der Schlüssel und die Schlüsselsegmentierungsposition S7 des Pylonenkopfes. Darüber hinaus können weitere Segmentierungspositionen in der Hohlstruktur identifiziert werden.

niedrigere Füllraten an den wichtigsten Segmentierungspositionen innerhalb der Hohlstruktur als an den zuvor genannten Positionen. Durch die Kombination der Positionsbeziehung mit Verteilungseigenschaften werden die wichtigsten Segmentierungspositionen in der Hohlstruktur identifiziert.

Für den Pylon Typ B und den Pylon Typ C kann festgestellt werden, dass

Machine Translated by Google

Für den Pylon vom Typ B und den Pylon vom Typ C kann festgestellt werden, dass die Segmentierung
Die Position an der Spitze des Pylons konnte aufgrund der geringen Füllrate des
Punktwolken durch Betrachtung von Abbildung 5, und es muss separat betrachtet werden. Die höchste
unter den erwarteten Schlüsselsegmentierungspositionen, die den Schwellenwert für die Füllrate nicht erreichen
Anforderung kann als eine der Schlüsselsegmentierungspositionen gewählt werden. Schließlich werden alle Schlüssel
und der Pylon vom Typ C kann erhalten werden wie in Abbildung 8 gezeigt Fehler! Referenzquelle nicht
Segmentierungspositionen des Pylons vom Typ B und des Pylons vom Typ C können erhalten werden, wie gezeigt
gefunden.
in Abbildung 8.



(A) (B)

Abbildung 8. Die wichtigsten Segmentierungspositionen des Pylons. (a) Pylon Typ B. (b) Pylon Typ C.

Abbildung 8. Die wichtigsten Segmentierungspositionen des Pylons. (a) Pylon Typ B. (b) Pylon Typ C.

3.3. Bestimmung der Merkmalspunkte

3.3. Bestimmung der Merkmalspunkte

Der Hauptzweck der Bestimmung von Pylonmerkmalspunkten besteht darin, verschiedene Parameter unbekannter
Der Hauptzweck der Bestimmung von Pylon-Merkmalspunkten besteht darin, verschiedene
Pylone besser berechnen zu können, wie z. B. Pylonhöhe, Querarmlänge, Pylonbeinabstand
Parameter unbekannter Pylone, wie Pylonhöhe, Querarmlänge, Pylonbeinabstand usw. Diese Studie wählt Punkte an
den Grenzen der wichtigsten Segmentierungspositionen der Pylone als
usw. Diese Studie wählt Punkte an den Grenzen der wichtigsten Segmentierungspositionen der Pylone als
Merkmalspunkte. Die vier in dieser Studie betrachteten Pylontypen haben den gleichen Pylonkörper
Merkmalspunkte. Die vier in dieser Studie betrachteten Masttypen haben die gleiche Mast- und Mastbeinstruktur. Diese
Merkmalspunkte reichen aus, um die erforderlichen Parameter zu berechnen
Körper- und Pylonbeinstrukturen. Diese Merkmalspunkte reichen aus, um die für diese beiden Strukturen
erforderlichen zu berechnen. Jeder Pylontyp hat jedoch eine komplexe und einzigartige Pylonstruktur.
Parameter für diese beiden Strukturen. Allerdings hat jeder Pylontyp eine komplexe und kopfförmige Struktur, die
die Identifizierung zusätzlicher Merkmalspunkte erfordert, um die
Einzigartige Pylonenkopfstruktur, die die Identifizierung zusätzlicher Merkmalspunkte zu relevanten Parametern des
Pylonenkopfes erfordert.
Berechnen Sie die relevanten Parameter des Pylonkopfes.

Am Beispiel des Pylons vom Typ B wird die Strategie zur Auswahl von Merkmalspunkten vorgestellt.
Am Beispiel des Pylons vom Typ B wird die Auswahl der Merkmalspunkte vorgestellt. Die ausgewählte
Position der Merkmalspunkte des Pylons vom Typ B ist in Abbildung 9 dargestellt.
Strategie, die ausgewählte Position der Merkmalspunkte des Typ-B-Pylons ist in Abbildung dargestellt, wobei die
orangefarbenen Punkte die Merkmalspunkte darstellen, die in Verbindung mit dem Schlüssel erhalten wurden
darstellen. Punkt 25, wie in der Abbildung angegeben, stellt den Verbindungspunkt dar
9Fehler! Referenzquelle nicht gefunden., wobei die orangefarbenen Punkte die Positionen der Merkmalssegmentierung
dem Querarm und dem Pylonkörper angegeben. Die Z-Achsenkoordinate kann bestimmt werden durch
Punkte, die in Verbindung mit den wichtigsten Segmentierungspositionen erhalten wurden. Punkt 25 wie zwischen
stellt die Abbildung Position Verbindungspunkt zwischen Querarm und Pylonkörper dar. Punkt 27 kann als
Die Z-Achsenkoordinate kann durch Identifizierung ihrer charakteristischen Position bestimmt werden. Auf diese Weise
Konnten Punkte 11, 12 und 26 identifiziert werden. Bezogen auf Punkte 7 und Punkte
können Punkte 11, 12 und 26 sowie der Punkt mit der größten Z-Achsenkoordinate in jeder Schicht ausgewählt werden, um die 8 zu berechnen. Die
Punktwolken zwischen Punkt 1 und Punkt 2 sind entlang der Yÿ-Achse mit einer bestimmten
Yÿ-Achsenkoordinate. Ebenso werden Koordinateninformationen für die Intervalle von Punkt 11,
Punkt 12 und Punkt 26 sowie der Punkt mit der größten Z-Achsenkoordinate in jeder Schicht ausgewäh
kann auf diese Weise erhalten werden. Bezüglich Punkt 7 und Punkt 8 bewegen sich die Punktwolken zwischen Punkt 1
1 und Punkt 2 werden entlang der Yÿ-Achse mit einem bestimmten Intervall geschichtet, und der Punkt mit
den Z -Achsen-Koordinaten, der sich an der Schlüsselsegmentierungsposition S10 befindet , wird gefunden. Diese
Die maximale Z-Achsenkoordinate in jeder Ebene wird ausgewählt. Dann werden die Punkte, die von Punkt 1 und Punkt
2 ausgehen, als Punkt 7 und Punkt 8 identifiziert, wie durch die roten Markierungen in Abbildung 10 angezeigt.
2 in Richtung Mitte durch ein Schiebefenster, die ersten Punkte mit Z-Achsen-Koordinaten befinden sich Insgesamt
müssen 36 Merkmalspunkte für den Pylon vom Typ A bestimmt werden, während der Pylon vom Typ B,
an der Schlüsselsegmentierungsposition S10 gefunden. Diese beiden Punkte werden als Punkt 7 und Punkt 8 identifiziert,
Pylone vom Typ C und Typ D erfordern jeweils 63, 51 und 52 Merkmalspunkte.
wie durch die roten Markierungen in Abbildung 10 angezeigt. Fehler! Referenzquelle nicht gefunden.

Insgesamt müssen für den Pylon Typ A 36 Merkmalspunkte bestimmt werden, während für den Pylon Typ
Pylone vom Typ b, Typ c und Typ d erfordern jeweils 63, 51 und 52 Merkmalspunkte.

Machine Translated by Google

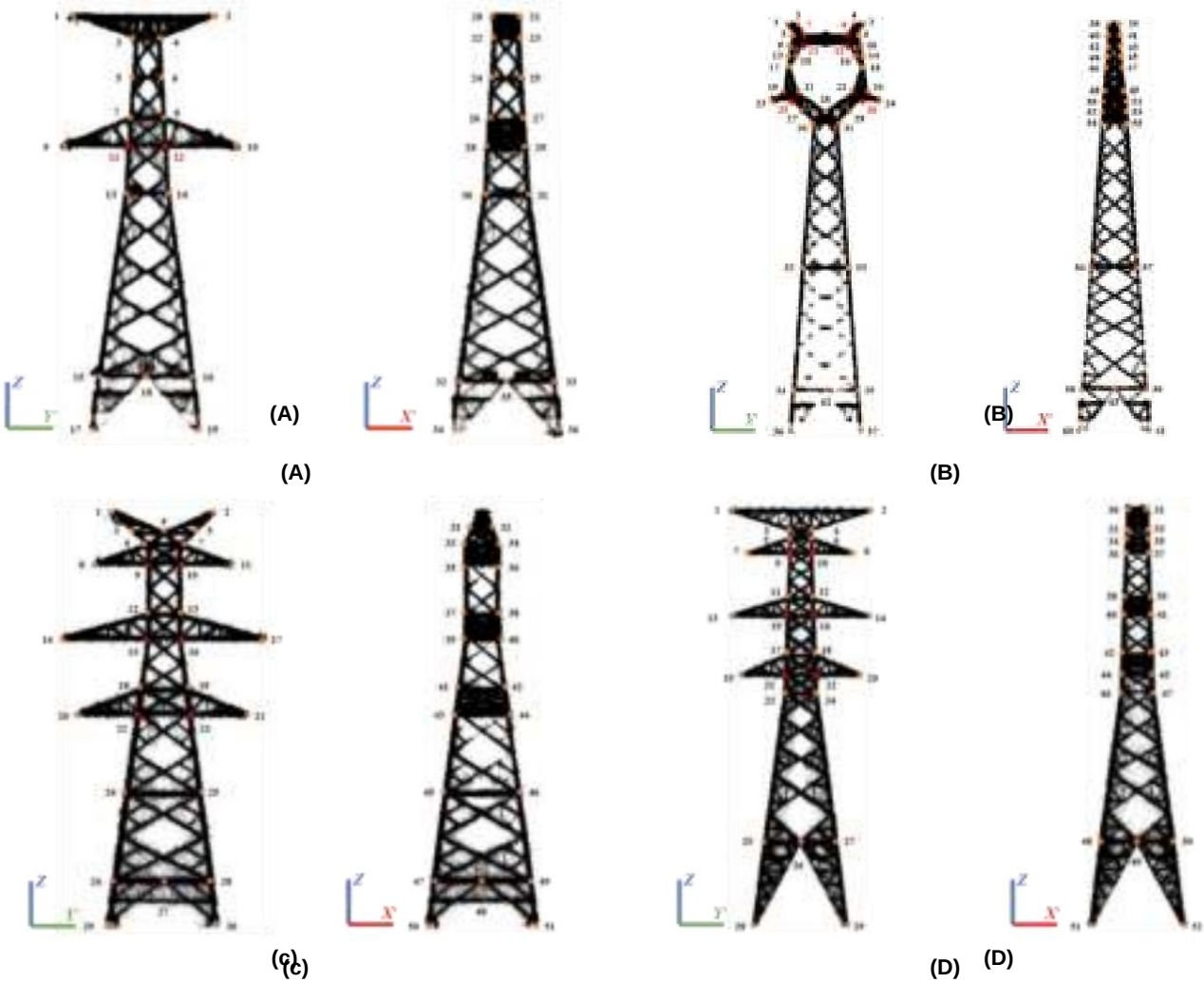


Abbildung 9. Pylon-Merkmalspunkte und Nummern, und rote Punkte sind die Merkmalspunkte, die durch Abbildung 9. Pylon-Merkmalspunkte und -Nummern, und rote Punkte sind die Merkmalspunkte, die durch Kombination mit den umliegenden orangefarbenen Punkten. (a) Typ-A-Mast. (b) Typ-B-Mast. (c) Typ-C- Kombination mit den umliegenden orangefarbenen Punkten. (a) Typ-A-Mast. (b) Typ-B-Mast. (c) Typ-Pylon. (d) Pylon Typ d.
(d) Pylone vom Typ d.

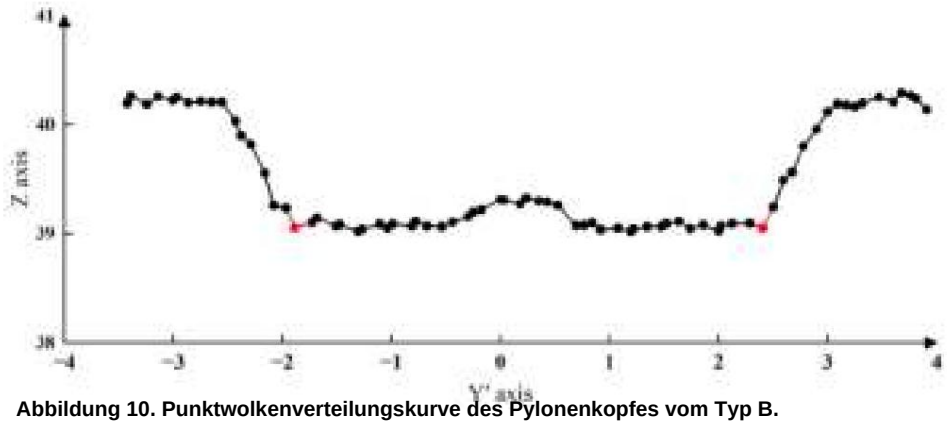


Abbildung 10. Punktwolkenverteilungskurve des Pylonenkopfes vom Typ B.

Abbildung 10. Punktwolkenverteilungskurve des Pylonenkopfes vom Typ B.

Abbildung 10. Punktwolkenverteilungskurve des Pylonenkopfes vom Typ B.

3.4. Berechnung charakteristischer Parameter

3.4. Berechnung charakteristischer Parameter

Basierend auf den Koordinateninformationen der Merkmalspunkte werden charakteristische Parameter

Parameter *3.4. Berechnung der charakteristischen Parameter*

Basierend auf den Koordinateninformationen der Merkmalspunkte werden charakteristische Parameter

Basierend auf den Koordinateninformationen der Merkmalspunkte werden charakteristische

werden aus den Perspektiven Höhe, Länge und Breite berechnet. Diese Studie untersucht die

Parameter aus den Perspektiven Höhe, Länge und Breite berechnet. Diese Studie untersucht die charakteristische

Die zu berechnenden Parameter werden am Beispiel eines Pylons vom Typ A vorgestellt. Die

Parameter. Am Beispiel des Pylons vom Typ A werden die zu berechnenden Parameter vorgestellt. Die

In Bezug auf die Höhe stellt der vertikale Koordinatenunterschied zwischen Punkt 20 und Punkt 30 dar. In Bezug auf die Höhe stellt der vertikale Koordinatenunterschied zwischen Punkt 20 und Punkt 3 dar. Der vertikale Koordinatenunterschied zwischen Punkt 30 und Punkt 3 stellt die Pylonkopfhöhe dar.

In Bezug auf die Höhe stellt der vertikale Koordinatenunterschied zwischen Punkt 20 und Punkt 30 dar. In Bezug auf die Höhe stellt der vertikale Koordinatenunterschied zwischen Punkt 20 und Punkt 3 dar. Der vertikale Koordinatenunterschied zwischen Punkt 30 und Punkt 3 stellt die Pylonkopfhöhe dar.

Machine Translated by Google

die Höhe des Pylonkopfes. Der vertikale Koordinatenunterschied zwischen Punkt 30 und Punkt 32 stellt die Pylonenkörperhöhe dar, und die Differenz zwischen Punkt 32 und Punkt 34 stellt die Höhe des Pylonbeins dar. Was die Länge betrifft, kann die Grundlinien-Querarmlänge erhält man durch Messen der horizontalen Koordinatendifferenz zwischen Punkt 1 und Punkt 2, während die Querarmlänge durch die Differenz zwischen Punkt 9 und Punkt bestimmt wird. 10. Zusätzlich ist der Unterschied in den horizontalen Koordinaten zwischen Punkt 17 und Punkt 19 stellt den Abstand der Pylonbeine dar. Bezüglich der Breite gilt für Punkte, die auf die XÿZ- Ebene projiziert werden, Die Berechnung der horizontalen Koordinatendifferenz der Punkte auf gleicher Höhe kann ergeben die erforderlichen Breitenparameter.

4. Experimente

In diesem Abschnitt werden zunächst die im Experiment verwendeten Daten und Parameter vorgestellt. Abschnitt 4.1, und dann die Genauigkeit der Turmtypidentifikation und Merkmalsparameter Berechnung ist in den Abschnitten 4.2 und 4.3 aufgeführt . Schließlich wird die rekonstruierte 3D Modelle der Pylone sind in Abschnitt 4.4 dargestellt.

4.1. Einführung experimenteller Daten und Parameter

Diese Studie erfasste Punktwolkendaten mithilfe der Flugplattform DJI M300RTK mit dem integriertes Livox L1 LiDAR-Modul. Die grundlegenden Parameter der LiDAR-Daten werden angezeigt in Tabelle 1. Die gesammelten LiDAR-Daten umfassen nicht nur die 3D-Koordinaten der Punkte sondern auch deren RGB-Farbinformationen. Die Daten umfassen Stromübertragungskorridore mit Spannungsebenen von 110 kV und 220 kV in den Provinzen Hubei und Sichuan. Um den Fortschritt dieser Studie zu erleichtern, wurden die Daten mithilfe der Open-Source-Software CloudCompare 2.13 zum Abrufen der Pylon-Punktwolkendaten.

Tabelle 1. Grundlegende Parameter der LiDAR-Daten.

Punktdichte	Horizontale Genauigkeit	Vertikale Genauigkeit
>100 Punkte/m2	10 cm	5 cm

Die Programme zur Masttypidentifikation und Parameterberechnung wurden in Python und auf einem Laptop ausführen. Die Konfigurationsinformationen des Laptops sind in Tabelle 2 aufgeführt. Die durch Segmentierung gewonnenen Pylon-Punktwolkendaten wurden mit der Methode in diesem Papier vorgeschlagen. Die an der Verarbeitung beteiligten Parameter sind in Tabelle 3 aufgeführt.

Tabelle 2. Informationen zur Laptop-Konfiguration.

Laptop	CPU	Grafikkarte	RAM
Lenovo Legion R9000P 2023 AMD Ryzen 9 7945HX NVIDIA RTX 4060			16 Gigabyte

Tabelle 3. Parametereinstellungen.

Parameter	Bedeutung	Werte
H	Mindesthöhe der Punktwolken für die Umleitung	(3/4) × die Pylonhöhe
h1	Der Schichtabstand entlang der Z-Achsenrichtung	0,1 m
h2	Die Höhe des Schiebefensters	1,1 m
L1	Das Gitterintervall entlang der Yÿ -Achsenrichtung	0,1 m
Tf	Die Schwelle der Füllrate	75 %
Ce	Fehlerkonstante	0,5
N1 (Typ A)	Die gewählte Anzahl an Schlüsselebenen	8
N1 (Typ B)		14
N1 (Typ C)		11
N1 (Typ D)		10

Machine Translated by Google

4.2. Genauigkeit der Pylonentypidentifizierung

Zur Identifizierung des Pylonentyps werden vier verschiedene Ähnlichkeitsbewertungsmessungen verwendet, nämlich DTW-Distanz, FastDTW- Distanz, diskrete Fréchet-Distanz und Hausdorff-Distanz . **Dabei wird die im vorherigen Abschnitt beschriebene Methode verwendet. In diesem Abschnitt werden zwei Arten von Formkurven-Datensätzen verwendet, um die Experimente durchzuführen. Einer ist der Datensatz, der direkt auf Grundlage von Zeichnungsinformationen gewonnen wurde, und der andere ist der Datensatz, der durch lineare Skalierung des obigen Datensatzes auf dieselbe Höhe gewonnen wurde. Schließlich wurden die Ergebnisse der Pylonentypidentifizierung der vier Ähnlichkeitsbewertungsmessun**

Um die Identifikationsleistung der verschiedenen Ähnlichkeitsmessungen umfassend zu bewerten , wurden in dieser Studie die **experimentellen Ergebnisse anhand der drei Indizes Präzision, Rückruf und F-Score analysiert [25].**

Mit Präzision ist der Anteil der Pylonen gemeint, bei denen eine Zugehörigkeit zu einem bestimmten Typ vorhergesagt wurde und die im Experiment auch tatsächlich diesem Typ angehören.

$$P = \times \frac{TP}{TP + FP} 100\ \% \quad (9)$$

Mit Recall wird der Anteil der Masten eines bestimmten Typs bezeichnet, bei denen letztlich die Zugehörigkeit zu diesem Typ vorhergesagt wird.

$$R = \times \frac{TP}{TP + FN} 100\ \% \quad (10)$$

Dabei ist TP die Anzahl der Proben, die einem bestimmten Masttyp zugeordnet werden und tatsächlich zu diesem Typ gehören. FP ist die Anzahl der Proben, die einem bestimmten Masttyp zugeordnet werden, tatsächlich aber zu anderen Typen gehören. FN ist die Anzahl der Proben eines bestimmten Masttyps, die anderen Typen zugeordnet werden.

Der F-Score bezieht sich auf das harmonische Mittel aus Präzision und Rückruf und wurde als Gesamtbewertungsindex der Pylonidentifizierungsleistung in dieser Studie.

$$F = \frac{2P \times R}{P + R} \times 100\ \% \quad (11)$$

Je höher die Werte dieser drei Bewertungsindizes, desto besser ist die Identifizierungsleistung der Ähnlichkeitsmessung. Für die vier Masttypen dieser Studie werden die Ergebnisse der Masttypidentifizierung unter Verwendung unterschiedlicher Ähnlichkeitsbewertungsmethoden in Abbildung 11 dargestellt.

Bei Verwendung des Originaldatensatzes zur Berechnung wiesen die DTW-Distanz und die FastDTW-Distanz sowohl für den Mast vom Typ A als auch für den Mast vom Typ B hinsichtlich der Präzision die höchste Präzision auf. Für den Mast vom Typ C betrug die Präzision der diskreten Fréchet-Distanz 89,7 % und war damit deutlich höher als bei den anderen drei Ähnlichkeitsmessungen. Für den Mast vom Typ D waren alle vier Ähnlichkeitsmessungen im Wesentlichen gleichwertig. Hinsichtlich der Trefferquote ergaben die Hausdorff-Distanz und die diskrete Fréchet-Distanz für den Mast vom Typ A und den Mast vom Typ B leicht höhere Ergebnisse. Für den Mast vom Typ C lag die Trefferquote aller vier Ähnlichkeitsmessungen bei etwa 76 %. Für den Mast vom Typ D waren die Ergebnisse aller vier Ä Betrachtet man den Gesamtbewertungsindex der Algorithmusleistung, so ergaben alle vier Ähnlichkeitsmessungen für den Typ-A-Pylon ähnliche Ergebnisse, wobei die Hausdorff-Distanz etwas schlechter war. Für den Typ-B-Pylon lagen die F-Scores aller vier Ähnlichkeitsmessungen unter 66 %, was auf eine schlechte Identifizierungsleistung hindeutet. Für den Typ-C-Pylon betrug der F-Score der diskreten Fréchet-Distanz 83,8 %, was besser war als bei anderen Ähnlichkeitsmessungen . Für den Typ-D-Pylon ergaben die DTW-Distanz und die FastDTW-Distanz etwas höhere F-Scores.

Machine Translated by Google

Der Wert der Hausdorff-Distanz war höher als bei den anderen drei Messungen.

Zusammenfassend lässt sich sagen, dass bei der Identifizierung des Typs der Pylon-Punktwolken unter Verwendung des normalen Datensatzes zur Berechnung die diskrete Fréchet-Distanz mit einem durchschnittlichen F-Score von 76,4 % den besten Gesamtbewertungsindex aufwies. Unter Verwendung des durch lineare Skalierung transformierten Datensatzes zur Berechnung war der Gesamtbewertungsindex der Hausdorff-Distanz mit einem durchschnittlichen F-Score von 86,4 % der beste.

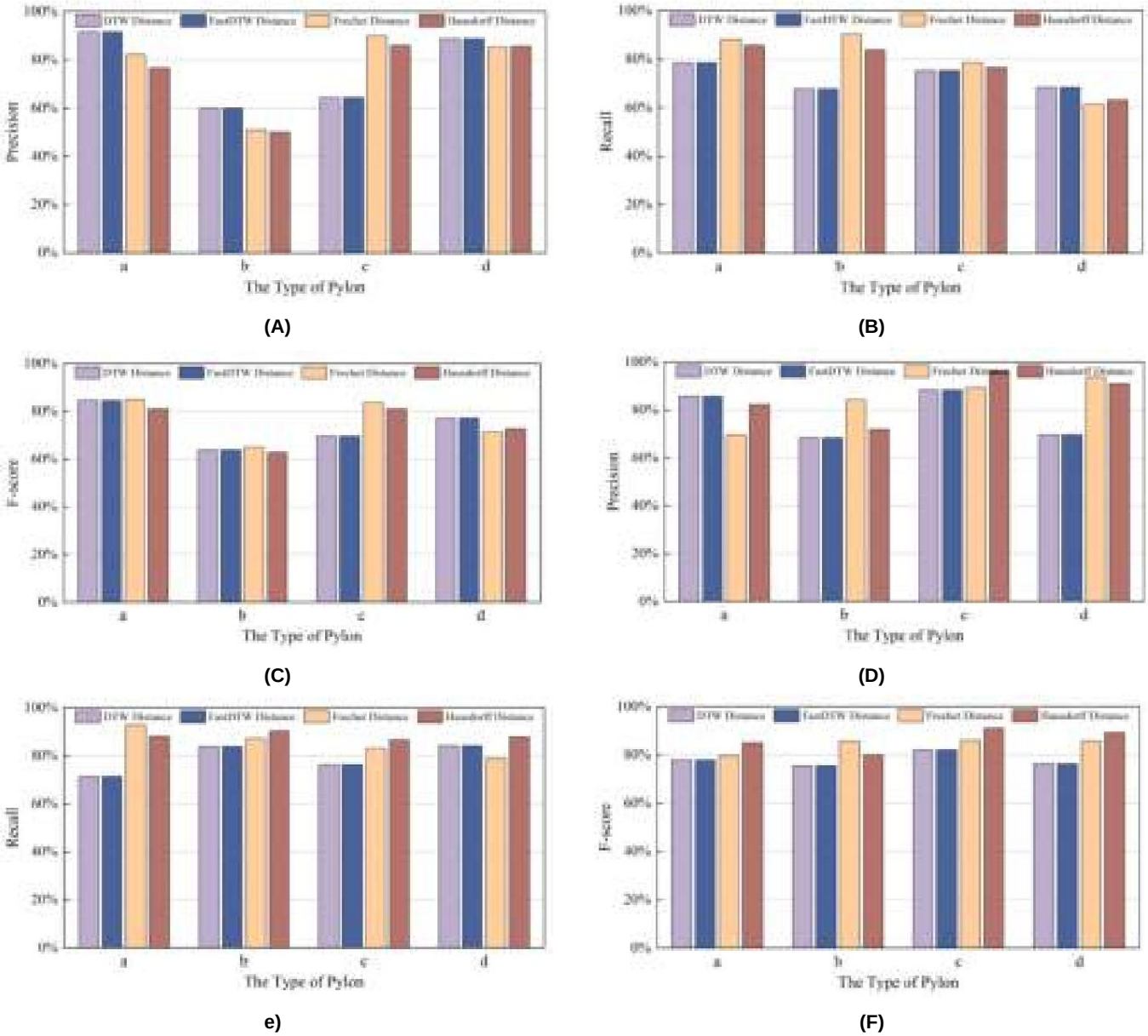


Abbildung 11. Identifikationsergebnisse von Masttypen unter Verwendung verschiedener Ähnlichkeitsmessungen. (a) Die Genauigkeit von vier Ähnlichkeitsmessungen (ursprünglicher Datensatz). (b) Der Abruf von vier Ähnlichkeitsmessungen (ursprünglicher Datensatz). (c) Der F-Score von vier Ähnlichkeitsmessungen (ursprünglicher Datensatz). (d) Die Präzision von vier Ähnlichkeitsmessungen (linear transformierter Datensatz). (e) Die Wiedererkennung von vier Ähnlichkeitsmessungen (linear transformierter Datensatz). (f) Der F-Score von vier Ähnlichkeitsmessungen (linear transformierter Datensatz).

Abbildung 11. Identifikationsergebnisse von Masttypen unter Verwendung verschiedener Ähnlichkeitsmessungen. (a) Die Genauigkeit von vier Ähnlichkeitsmessungen (ursprünglicher Datensatz). (b) Der Abruf von vier Ähnlichkeitsmessungen (ursprünglicher Datensatz). (c) Der F-Score von vier Ähnlichkeitsmessungen (ursprünglicher Datensatz). (d) Die Präzision von vier Ähnlichkeitsmessungen (linear transformierter Datensatz). (e) Die Wiedererkennung von vier Ähnlichkeitsmessungen (linear transformierter Datensatz). (f) Der F-Score von vier Ähnlichkeitsmessungen (linear transformierter Datensatz).

Unter Verwendung des durch lineare Skalierung transformierten Datensatzes für die Berechnung, alle vier Ähnlichkeit Messungen zeigten eine signifikante Verbesserung der Genauigkeit bei der Identifizierung des Typ-B Pylon. Für den Typ-C-Pylon und den Typ-D-Pylon beträgt die Genauigkeit der Hausdorff-Distanz könnte 90% übersteigen. Was den Rückruf betrifft, so ist für den Typ-A-Pylon die diskrete Fréchet-Distanz ergaben höhere Ergebnisse. Für die anderen drei Masttypen waren die Rückrufergebnisse der Hausdorff Die Distanz betrug 90,3 %, 86,5 % bzw. 87,7 %, also deutlich mehr als die anderen Ähnlichkeitsmessungen. Betrachtet man den Gesamtbewertungsindex der Algorithmusleistung , so haben sich die F-Scores aller vier Ähnlichkeitsmessungen deutlich verbess für den Typ-b-Pylon. Für die anderen drei Pylontypen beträgt der F-Score des Hausdorff-Die Distanz war größer als bei den anderen drei Messungen.

Machine Translated by Google

Zusammenfassend lässt sich sagen, dass bei der Identifizierung des Typs von Pylon-Punktwolken die Verwendung der gewöhnlichen Als Berechnungsdatensatz zeigte die diskrete Fréchet-Distanz die beste Gesamtbewertung Index mit einem durchschnittlichen F-Score von 76,4 %. Unter Verwendung des durch lineare Skalierung transformierten Datensatzes Für die Berechnung war der Gesamtbewertungsindex der Hausdorff-Distanz der beste, mit einem durchschnittlicher F-Score von 86,4 %.

4.3. Berechnungsgenauigkeit der charakteristischen Parameter von Masten

In diesem Abschnitt wurde die Genauigkeit der Berechnung der Schlüsselparameter für vier Arten von Masten anhand praktischer Beispiele. Als typischer Fall wurde hier der Mast Typ A betrachtet. Der Die berechneten Werte dieses Pylons sind in Tabelle 4 aufgeführt und wurden mit dem Handbuch verglichen. Messwerte. Abschließend wurde der relative Fehler ermittelt.

Tabelle 4. Ergebnisse der Parameterberechnung des Pylons vom Typ A.

Seriell Nummer	Wichtige Punkte Verbindung	Positionsbeschreibung	Berechnet Wert/m	Manuelle Messung Wert/m	Relativ Fehler
1	1–2 Abszissendifferenz	Grundlinienquerarmlänge 3–4	10,2897	10.5	ÿ2,00 %
2	Abszissendifferenz / 5–6	Abszissendifferenz / 7–8	1,6932	1.8	ÿ5,93 %
3	Abszissendifferenz / 9–10	Abszissendifferenz	2,0760	2,1	ÿ1,14 %
4			2,4361	2,4	1,50 %
5		Querarmlänge 11–12	12,5012	13	ÿ3,84 %
6	Abszissendifferenz / 13–14	Abszissendifferenz / 15–	2,6784	2,7	ÿ0,80 %
7	16 Abszissendifferenz / 17–19	Abszissendifferenz	3,1005	3,2	ÿ3,11 %
8			6,8951		ÿ1,50 %
9		Abstand der	8,0340		0,43 %
10	Pylonbeine 20–21 Abszissendifferenz	Grundlinien-Querarmbreite 22–	1,5211	7 8	ÿ4,93 %
11	23 Abszissendifferenz / 24–25	Abszissendifferenz /	1,6932	1,6	ÿ5,93 %
12	26–27 Abszissendifferenz /		2,1018	1,8 2,1	0,09 %
13			2,4824	2.4	3,43 %
14	28–29 Abszissendifferenz	Traversenbreite	2,7438	2.8	ÿ2,01 %
15	30–31 Abszissendifferenz		3.1937	3.2	ÿ0,20 %
16	32–33 Abszissendifferenz	/ /	6,9042	7	ÿ1,37 %
17	34–36 Abszissendifferenz	Pylonenbeinabstand	8,0179	8	0,22 %
18	32–34 Ordinatendifferenz	Pylonenbeinhöhe	4,0512	4	1,28 %
19	30–32 Ordinatendifferenz	Pylonenkörperhöhe / / / / /	13,4974	13.5	ÿ0,02 %
20	28–30 Ordinatendifferenz 26–		3,4989	3.5	ÿ0,03 %
21	28 Ordinatendifferenz 24–26		2,0084	2	0,42 %
22	Ordinatendifferenz 22–24		2,9849	3	ÿ0,50 %
23	Ordinatendifferenz 20–22		3,0140	3	0,47 %
24	Ordinatendifferenz		1,5022	1.5	0,15 %

Schließlich wurde festgestellt, dass der relative Gesamtfehler 5 % nicht überschritt, indem man berechnete 80 Mastproben und der Vergleich ihrer maximalen relativen Fehler, wodurch die Durchführbarkeit des Algorithmus. Die berechneten Ergebnisse für die Pylone sind in Tabelle 5 dargestellt.

Tabelle 5. Berechnungsfehler.

Pylonentyp	Menge	Maximaler relativer Fehler	Durchschnittlicher Fehler
A	20	3,28 %	2,36 %
B	20	4,96 %	2,71 %
C	20	4,62 %	3,05 %
D	20	4,37 %	2,93 %

Machine Translated by Google

4.4. Wiederaufbau des Pylons

Basierend auf den im vorherigen Abschnitt berechneten charakteristischen Parametern werden 3D-Modelle der vier Masttypen wurden rekonstruiert. Die endgültigen Modelle sind in Abbildung 12 dargestellt.

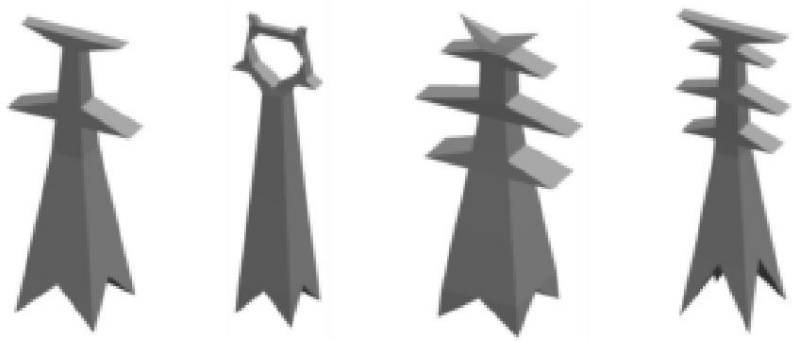


Abbildung 12. 3D-Modelle von Pylonen.

Abbildung 12. 3D-Modelle von Pylonen.

5. Diskussion

5. Diskussion

In diesem Abschnitt werden hauptsächlich die Fehler bei der Identifizierung des Pylontyps behandelt, die Einfluss der Rauschpunkte der LiDAR-Daten auf die Berechnung charakteristischer Parameter und der Einfluss von Datensparlichkeit.

In diesem Abschnitt werden hauptsächlich die Fehler erörtert, die bei der Pylontyp-Identifizierung auftreten und sich auf die Berechnung von Zeichen 5.1 auswirken. Fehler bei der Pylontyp-Identifizierung

der Einfluss von Datensparlichkeit.

Bei Verwendung des normalen Datensatzes tendieren die DTW-Distanz und die FastDTW-Distanz den Typ-B-Pylon und den Typ-D-Pylon fälschlicherweise als Typ-C-Pylon zu klassifizieren. Dies kann Dies ist darauf zurückzuführen, dass DTW eine lokale Matching-Methode ist, die unempfindlich gegenüber globalen Formen ist.

Änderungen, während Pylonformen globale Änderungen mit sich bringen, was zur Fehlidentifikation von bestimmte Pylonentypen. Die diskrete Fréchet-Distanz und die Hausdorff-Distanz weisen mehr

5.1. Fehler bei der Pylontyp-Identifizierung

Identifizierungsfehler auf. Bei der Pylon-Rekonstruktion vom Typ D, die mit der DTW-Distanz zusammenhängen können, in der Höhe zwischen dem Mast des Typs d und anderen Masttypen im Datensatz. Nach der Verwendung der

Wenn der Typ-B-Pylon und der Typ-D-Pylon als linear transformierter Datensatz ty

falsch klassifiziert werden, werden die Fehler bei der Identifizierung des Pylontyps red

Dies ist darauf zurückzuführen, dass DTW eine lokale Matching-Methode ist. Diese Die Forschung kann sich auf die Entwicklung besser geeigneter Identifikationsalgorithmen konzentrieren, basierend auf den dieser Studie ihre Allgemeingültigkeit verbessern.

Änderungen, während Pylonformen globale Änderungen mit sich bringen, was dazu führt, dass die Ergebnisse

bestimmte Masttypen. Die diskrete Fréchet-Distanz und die Hausdo-Fehler bei

der Identifizierung der Mastpunktwolken des Typs D, die in der Höhe zwischen

dem Mast des Typs D und anderen Masttypen liegen können, sind im linear

transformierten Datensatz zu finden. Die Fehler bei der Identifizierung des

Masttyps sind derzeit nur auf die Identifizierung bestehender Masttypen

anwenden. Zukünftige Forschung kann sich auf die Entwicklung geeigneterer

Identifizierungsmethoden konzentrieren, um die Ergebnisse dieser Studie zu verb

5.2. Der Einfluss von Störpunkten auf die Berechnung charakteristischer Kennwerte

Die Störpunkte verteilen sich hauptsächlich auf zwei Bereiche des Pylons. Einer der Bereiche befindet sich an der Verbindungsstelle zwischen Pylonkörper und Querträger, wie durch den blauen Kreis in Abbildung 13. Diese Rauschpunkte bewirken eine Verschiebung der ausgewählten Merkmalspunkte auf dem Pylonenkörper nach außen, was zu Fehlern bei der Berechnung der charakteristischen Parameter führt. In dieser Studie Die wahren Merkmalspunkte können durch die Dichtemerkmale der Punkte nach unterschieden werden

Punktwolken-Schlüsselsegmentierung positioniert. Daher können solche Störpunkte durch die

die Schlüsselsegmentierung positioniert Schichtung, wodurch die Störungen durch solche Geräusche

Isolator-String-Punkte und Leitungspunkte, wie durch den roten Kreis in Abbildung 13 dargestellt. Diese Art des Rauschpunktes hat den größten Einfluss auf die Berechnung der charakteristischen Parameter.

solche Punkte vorhanden sind, wird es schwierig sein, die Pylonparameter mit dem

Methode, die in diesem Papier vorgeschlagen wird. Daher ist es notwendig, solches Rauschen manuell zu entfernen

5.2. Der Einfluss von Störpunkten auf die Berechnung charakteristischer **Punkte vor der Bestimmung des Masttyps.**

Rauschpunkte sind hauptsächlich in zwei Bereichen des Pylonkörpers verteilt, an der Verbindung von Pylonkörper und Querarm, wie in Abbildung 13 angegeben. Fehler! Referenzquelle nicht gefunden. Diese Rauschpunkte verschieben sich nach außen, was zu einer übermäßigen Parameterberechnung führt. In dieser Studie können die wahren Merkmalspunkte Dichteeigenschaften von Punkten nach der Schlüsselsegmentierung p eliminieren, indem die Interferenz solcher Rauschpunkte an Merkmalspunkten des Rauschpunkts rimaril aus Isolatorsträngen besteht

Machine Translated by Google

hat den größten Einfluss auf die Berechnung der charakteristischen Parameter. Es wird schwierig sein, die Pylonparameter in diesem Dokument genau zu berechnen. Daher ist es notwendig, die Identifizierung des Pylontyps manuell zu entfernen.

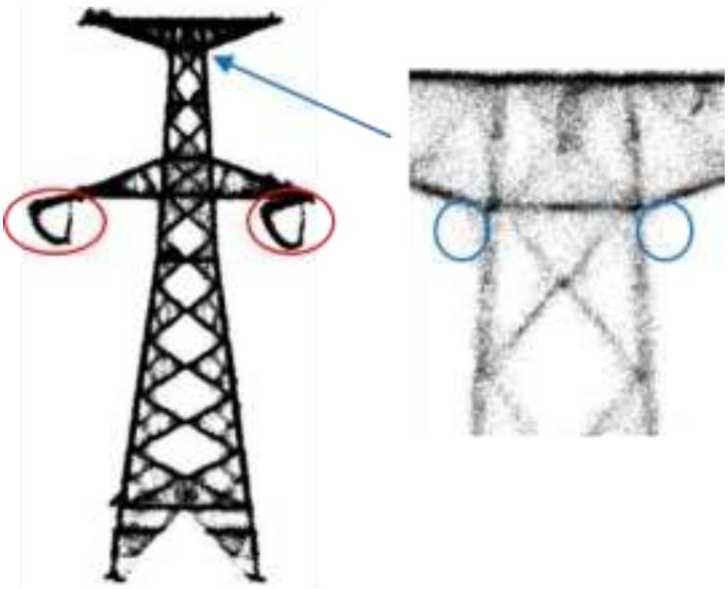


Abbildung 13. Pylon-Punktwolken mit Interferenzpunkten.

5.3. Der Einfluss von Datensparsity

Im Prozess der Leistungsinspektion, die Änderung der Flughöhe und Geschwindigkeit des UAV führt zu einem Unterschied in der Punktdichte. Wenn die Flughöhe höher ist und der Flug Geschwindigkeit ist schneller, die erreichte Punktdichte ist geringer. Um den Einfluss der Datensparsity zu analysieren zur Identifizierung des Masttyps und zur Berechnung der charakteristischen Parameter, der ursprüngliche Mast Punktwolkendaten werden mit der Voxel-Sampling-Methode abgetastet. Die Anzahl der Pylonpunkte Die bei verschiedenen Abtastentfernungen erhaltenen Werte sind in Tabelle 6 aufgeführt. Wenn die Abtastentfernung weniger als 0,4 m beträgt, können die Methoden in diesem Dokument verwendet werden, um die Art der Pylon und die berechneten Pylonparameter sind relativ genau. Wenn jedoch die Wenn die Abtastentfernung größer als 0,4 m ist, werden die Punktwolken dünner und die Verteilung Parameter der Punktwolken können nicht die Eigenschaften der Schlüsselsegmentierung widerspiegeln Positionen der Pylone. Die Umleitungsverarbeitung wird auch beeinflusst, wenn die Probenahme Die Entfernung überschreitet 0,6 m. Die Parameter in Tabelle 3 können die horizontale Richtung des der Pylone senkrecht zur X-Achse. In diesem Fall müssen die Parameter neu eingestellt werden, um dieses Problem.

Tabelle 6. Die Anzahl der mit unterschiedlichen Abständen abgetasteten Pylonenpunkte.

Pylonentyp	Die Anzahl der Punkte				
	Original Punktwolke	Abtastweite			
		0,1 m	0,2 m	0,3 m	0,4 m
A	203.792	67.631	20.809	9528	5492
B	102.429	47.194	16.360	7786	4706
C	239.345	128.429	42.751	19.832	11.063
D	196.601	112.847	42.108	19.851	11.166

6. Schlussfolgerungen

Diese Studie schlägt eine verbesserte Methode zur Identifizierung von Strommasttypen vor und eine neuartige Methode zur automatischen Berechnung charakteristischer Parameter mit dem Ziel, das Probleme der komplexen Berechnung und der geringen Effizienz bei bestehenden Methoden. Sie können die notwendige Datenunterstützung für die Rekonstruktion von 3D-Modellen von Masten. Die vorgeschlagene Methode in diesem Papier weist mehrere Merkmale auf und zeigt großes Potenzial bei der Nutzung

Machine Translated by Google

luftgestützte LiDAR-Daten, um grundlegende Informationen über Masten zu erhalten. Die Forschungsergebnisse dieses Artikels können wie folgt zusammengefasst werden.

(1) Dieser Artikel stellt eine Methode zur Generierung von Pylonenformkurven auf der Grundlage von Punktwolkendaten vor . Auf dieser Grundlage wird eine verbesserte Methode zur Identifizierung von Punktwolkentypen auf der Grundlage von Ähnlichkeitsmessungen und einem linear transformierten Datensatz vorgeschlagen. Diese Methode vereinfacht das Problem der Pylonentypidentifizierung im 3D-Raum zur Kurvenidentifizierung im 2D- Raum. Sie kann eine Vielzahl von Pylonentypen effektiv identifizieren und Informationsunterstützung für die Parameterberechnung von Pylonenpunktwolken bieten.

(2) In dieser Studie wurden die Identifikationseffekte von vier Ähnlichkeitsmessungen verglichen: DTW-Distanz, FastDTW-Distanz, diskrete Fréchet-Distanz und Hausdorff-Distanz. In Bezug auf den Gesamtbewertungsindex (F-Score) ergab die diskrete Fréchet-Distanz als Ähnlichkeitsmessung bei Verwendung des gewöhnlichen Datensatzes den optimalen Gesamtbewertungsindex von 76,4 %. Unterdessen erzielte die Hausdorff-Distanz als Ähnlichkeitsmessung bei Verwendung des Datensatzes nach linearer Transformation mit einem durchschnittlichen F-Score von 86,4 % die beste Leistung.

(3) Es wird eine neuartige Methode zur Berechnung von Mastparametern auf Grundlage der Verteilungseigenschaften von Punktwolken vorgeschlagen. Mit dieser Methode lassen sich punktwolkenspezifische Informationen effektiv extrahieren und die Genauigkeit der Parameterberechnung sicherstellen. Durch die Berechnung und Analyse von 80 Mastgruppen konnte festgestellt werden, dass der maximale relative Fehler, der durch diesen Algorithmus erzeugt wird, 5 % nicht überschreitet, was die Durchführbarkeit des Algorithmus bestätigt.

Obwohl die in dieser Studie vorgeschlagene Methode relativ genaue Ergebnisse **bei der Identifizierung von Masttypen und der Berechnung charakteristischer Parameter liefern kann ,** gibt es noch einige **Aspekte, die in zukünftigen Untersuchungen optimiert werden müssen. (1) Die in diesem Artikel betrachteten Masttypen sind durch den Datensatz begrenzt. Daher ist es notwendig, den Datensatz in zukünftigen Studien zu erweitern. (2) Die Auswahlmethode für Merkmalspunkte muss kontinuierlich optimiert werden, um Berechnungsfehler in den charakteristischen Parametern zu reduzieren. (3) In zukünftigen Untersuchungen werden auch Masten**

Beiträge der Autoren: Konzeptualisierung, SZ und LW; Methodik, SZ; Validierung, SZ; Ressourcen, LW und BS; Datenkuratierung, SZ, SW und GW; Schreiben – Vorbereitung des Originalentwurfs, SZ; Schreiben – Überprüfung und Bearbeitung, SZ; Visualisierung, SZ. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Forschung erhielt keine externe Finanzierung.

Datenverfügbarkeitserklärung: Die in dieser Studie präsentierten Daten sind auf Anfrage beim entsprechenden Autor erhältlich. Aus Datenschutzgründen sind die Daten nicht öffentlich verfügbar.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Verweise

1. **Lu, Z.; Gong, H.; Jin, Q.; Hu, Q.; Wang, S. Ein auf dichten Punktwolken basierender Ansatz zur Bewertung des Neigungszustands von Sendemasten von UAV-basiertem LiDAR.** Remote Sens. **2022,** 14, **408. [CrossRef]**
2. **Li, W.; Luo, Z.; Xiao, Z.; Chen, Y.; Wang, C.; Li, J. Eine GCN-basierte Methode zum Extrahieren von Stromleitungen und Masten aus luftgestützten LiDAR-Daten.** IEEE Trans. Geosci. Remote Sens. **2022,** 60, **1–14. [CrossRef]**
3. **Chen, C.; Jin, A.; Yang, B.; Ma, R.; Sun, S.; Wang, Z.; Zong, Z.; Zhang, F. DCPLD-Net: Ein diffusionsgekoppeltes Faltungsneuronales Netzwerk zur Echtzeiterkennung von Stromübertragungsleitungen anhand von UAV-gestützten LiDAR-Daten.** Int. J. Appl. Earth Obs. Geoinf. **2022,** 112, **102960. [CrossRef]**
4. **Yang, J.; Kang, Z. Voxelbasierte Extraktion von Übertragungsleitungen aus luftgestützten LiDAR-Punktwolkendaten.** IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **2018,** 11, **3892–3904. [CrossRef]**
5. **Xie, X.; Liu, Z.; Xu, C.; Zhang, Y. Eine Methode mit mehreren Sensoren zur Inspektion von Stromleitungen auf Basis eines großen unbemannten Hubschraubers.** Sensors **2017,** 17, **1222. [CrossRef] [PubMed]**
6. **Jiang, S.; Jiang, W.; Huang, W.; Yang, L. UAV-basierte Schrägphotogrammetrie zur Datenerfassung im Freien und zur externen visuellen Inspektion von Übertragungsleitungen.** Remote Sens. **2017, 9, 278. [CrossRef]**
7. **Yang, L.; Fan, J.; Liu, Y.; Li, E.; Peng, J.; Liang, Z. Ein Überblick über modernste Techniken zur Inspektion von Stromleitungen.** IEEE Trans. Instrument. Mess. **2020,** 69, **9350–9365. [CrossRef]**
8. **Zhou, R.; Jiang, W.; Jiang, S. Eine neuartige Methode zur Rekonstruktion von Hochspannungsbündelleitern aus luftgestützten LIDAR-Daten.** Remote Sens. **2018,** 10, **2051. [CrossRef]**

Machine Translated by Google

9. **Zhao, P.; Yang, W.; Feng, Y.; Li, F.; Huang, X. Konstruktion einer 3D-Szene eines Übertragungsleitungskorridors basierend auf GIM und 3D-GIS. 2005 ,** 012083. [CrossRef]
10. **Li, Q.; Chen, Z.; Hu, Q. Ein modellbasierter Ansatz zur 3D-Modellierung von Pylonen aus luftgestützten LiDAR-Daten.** Remote Sens. **2015, 7, 11501–11524.** [CrossRef]
11. **Tang, Q.; Zhang, L.; Lan, G.; Shi, X.; Duanmu, X.; Chen, K. Eine Klassifizierungsmethode für Punktwolken von Übertragungsleitungskorridoren basierend auf verbesserten Random-Forest- und Multi-Scale-Funktionen.** Sensors **2023,** 23, **1320.** [CrossRef] [PubMed]
12. **Zhu, S.; Li, Q.; Zhao, J.; Zhang, C.; Zhao, G.; Li, L.; Chen, Z.; Chen, Y. Eine auf Deep Learning basierende Methode zum Extrahieren einer beliebigen Anzahl einzelner Stromleitungen aus auf UAVs montierten Laserscan-Punktwolken.** Remote Sens. **2024,** 16, **393.** [CrossRef]
13. **Guo, B.; Huang, X.; Li, Q.; Zhang, F.; Zhu, J.; Wang, C. Eine stochastische Geometriemethode für die Rekonstruktion von Pylonen aus der Luft LiDAR-Daten.** Remote Sens. **2016, 8, 243.** [CrossRef]
14. **Yang, L.; Kong, S.; Deng, J.; Li, H.; Liu, Y. DRA-Net: Ein Dual-Branch Residual Attention Network für Pixelwise Power Line Erkennung.** IEEE Trans. Instrument. Mess. **2023,** 72, **5010813.** [CrossRef]
15. **Wang, G.; Wang, L.; Wu, S.; Zu, S.; Song, B. Semantische Segmentierung von 3D-Punktwolken von Übertragungskorridoren basierend auf CA-PointNet++.** Electronics **2023,** 12, **2829.** [CrossRef]
16. **Shen, Y.; Yang, Y.; Jiang, J.; Wang, J.; Huang, J.; Ferreira, V.; Chen, Y. Eine neuartige Methode zum Segmentieren einzelner Drähte aus Bündeln Leiter unter Verwendung von UAV-LiDAR-Punktwolkendaten.** Measurement **2023,** 211, **112603.** [CrossRef]
17. **Han, W. Dreidimensionale Modellierung von Kraftwerkstürmen mit luftgestützten LiDAR-Daten.** J. Yangtze River Sci. Res. Inst. **2012,** 29, **122–126.** [Querverweis]
18. **Zhou, R.; Jiang, W.; Huang, W.; Xu, B.; Jiang, S. Eine heuristische Methode zur Rekonstruktion von Strommasten anhand von luftgestützten LiDAR-Daten.** Remote Sens. **2017, 9, 1172.** [CrossRef]
19. **Camuffo, E.; Mari, D.; Milani, S. Jüngste Fortschritte bei Lernalgorithmen für Punktwolken: Ein aktualisierter Überblick.** Sensoren **2022,** 22, **1357.** [CrossRef]
20. **Qiao, Y.; Xi, X.; Nie, S.; Wang, P.; Guo, H.; Wang, C. Rekonstruktion von Strommasten aus luftgestützten LiDAR-Daten basierend auf Komponenten Segmentierung und Modellabgleich.** Remote Sens. **2022,** 14, **4905.** [CrossRef]
21. **Chen, S.; Wang, C.; Dai, H.; Zhang, H.; Pan, F.; Xi, X.; Yan, Y.; Wang, P.; Yang, X.; Zhu, X.; et al. Wiederaufbau von Strommasten basierend über abstrakte Musterstrukturen unter Verwendung von luftgestützten LiDAR-Daten.** Remote Sens. **2019,** 11, **1579.** [CrossRef]
22. **Silva, F.; Amaro, N. Klassifizierung von Sendemasten anhand der Ähnlichkeit von Punktwolken. In Proceedings of the APCA International Konferenz zu automatischer Steuerung und Soft Computing, Caparica, Portugal, 6.–8. Juli 2022; S. 609–618.**
23. **Chen, Z.; Lan, Z.; Long, H.; Hu, Q. 3D-Modellierung eines Pylons anhand von LiDAR-Daten aus der Luft. In Proceedings of the Remote Sensing of the Environment: 18. Nationales Symposium zur Fernerkundung Chinas, Wuhan, China, 14. Mai 2014; S. 915807.**
24. **Wang, H.; Hu, T.; Wang, Z.; Kang, Z.; Akwensi, PH; Yang, J. Rekonstruktion von Strommasten aus LiDAR-Punktwolken basierend auf Struktursegmentierung und Parameterschätzung.** IEEE Geosci. Remote Sens. Lett. **2022,** 19, **1–5.** [CrossRef]
25. **Zhang, M.; Su, X.; Xu, H.; Li, H.; Wang, D. Identifizierung von Sendemastenkategorien anhand von luftgestützten LiDAR-Punktwolken basierend auf Formkurven. In Proceedings der 3. Internationalen Konferenz für Mechatronik, Automatisierung und intelligente Steuerung, Guilin, China, 15.–17. September 2023; S. 012026.** [CrossRef]
26. **Principi, E.; Squartini, S.; Cambria, E.; Piazza, F. Akustisches Template-Matching zur automatischen Erkennung von Notfällen: Ein ELM Basierter Algorithmus.** Neurocomputing **2015,** 149, **426–434.** [CrossRef]
27. **Obaid, M.; Hodrob, R.; Abu Mwais, A.; Aldababsa, M. Automatische Spracherkennung für isolierte Wörter mit kleinem Wortschatz für Einwortbefehle im gesprochenen Arabisch.** Soft Comput. **2023, 1–14.** [CrossRef]
28. **Salvador, S.; Chan, P. Auf dem Weg zu präziser dynamischer Zeitverzerrung in linearer Zeit und Raum.** IDA **2007,** 11, **561–580.** [CrossRef]
29. **Gao, Y.; Yang, Y.; Ma, Y.; Xu, W. Studie zur intelligenten Diagnose von Eisenbahnweichen basierend auf verbessertem FastDTW und Zeitreihensegmentierung unter Big Data-Überwachung.** Math. Probl. Eng. **2022,** 2022, **7048813.** [CrossRef]
30. **Yeo, K.; Yin, OS; Han, PY; Kwee, WK Echtzeit-Mobilanwendung einer In-Air-Signatur mit Fast Dynamic Time Warping (FastDTW). In Proceedings der 2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Kuala Lumpur, Malaysia, 19.–21. Oktober 2015; S. 315–320.**
31. **Barbay, J. Adaptive Berechnung der diskreten Fréchet-Distanz. In Proceedings of the String Processing and Information Retrieval, Lima, Peru, 14. September 2018; S. 50–60.**
32. **Avraham, RB; Filtser, O.; Kaplan, H.; Katz, MJ; Sharir, M. Die diskrete und semikontinuierliche Fréchet-Distanz mit Abkürzungen über Ungefähre Distanzzählung und -auswahl.** ACM Trans. Algorithms **2015,** 11, **1–29.** [CrossRef]
33. **Ali, M.; Hussain, Z.; Yang, M.-S. Hausdorff-Distanz- und Ähnlichkeitsmaße für einwertige neutrosophische Mengen mit Anwendung in der multikriteriellen Entscheidungsfindung.** Electronics **2022,** 12, **201.** [CrossRef]

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultieren.

Machine Translated by Google

Elektronik

MDPI

Artikel

Algorithmus zur Erkennung und Verfolgung mehrerer Objekte auf Basis von YOLOv8 und ByteTrack

Longxiang You, Yajun Chen *, Ci Xiao, Chaoyue Sun und Rongzhen Li

Fakultät für elektronische Informationstechnik, China West Normal University, Nanchong 637009, China; ylxzxx@stu.cwnu.edu.cn (LY); xiaoci035560@stu.cwnu.edu.cn (CX); sunchaoyue@stu.cwnu.edu.cn (CS); lee1225rz@stu.cwnu.edu.cn (RL)
*** Korrespondenz: scnccyj@cwnu.edu.cn**

Zusammenfassung: Die Technologie zur Fahrzeugerkennung und -verfolgung spielt in intelligenten Transportsystemen eine entscheidende Rolle . Aufgrund von Faktoren wie komplexen Szenarien, unterschiedlichen Maßstäben und Verdeckungen treten jedoch häufig Probleme wie falsche Erkennungen, verpasste Erkennungen und Identitätswechsel auf. Um diese Probleme zu lösen, schlägt dieses Dokument einen Algorithmus zur Fahrzeugerkennung und -verfolgung für mehrere Objekte vor, der auf CDS-YOLOv8 und verbessertem ByteTrack basiert. Zur Fahrzeugerkennung wird während des Downsampling-Prozesses das Context-Guided (CG)-Modul eingeführt, um die Funktionen zur Merkmalsextrak Der Dilated Reparam Block (DRB) wird neu aufgebaut, um Probleme auf mehreren Skalen zu lösen, und Soft-NMS ersetzt das herkömmliche NMS, um die Leistung in dicht besiedelten Fahrzeug-Szenarien zu verbessern. Für die Fahrzeugverfolgung werden der Zustandsvektor und die Kovarianzmatrix des Kalman-Filters verbessert, um die nichtlineare Bewegung von Fahrzeugen besser handhaben zu können, und die Gaussian Smoothed Interpolation (GSI) wird eingeführt, um durch Erkennungsfehler verursachte Lücken in der Flugbahn zu füllen. Experimente, die mit dem UA-DETRAC-Datensatz durchgeführt wurden, zeigen, dass der verbesserte Algorithmus die Erkennungsleistung steigert, wobei sich mAP@0,5 und mAP@0,5:0,95 um 9 % bzw. 8,8 % verbessern. In Bezug auf die Verfolgungsleistung verbessert sich mMOTA um 6,7 %. Darüber hinaus zeigen Vergleichsexperimente mit Mainstream-Erkennungs- und zweistufigen Verfolgungsalgorithmen die überl

Schlüsselwörter: Computer Vision; intelligente Transportsysteme; Fahrzeugerkennung; Fahrzeugverfolgung

check for updates

Zitat: You, L.; Chen, Y.; Xiao, C.; Sun, C.; Li, R. Algorithmus zur Erkennung und Verfolgung von Mehrobjektfahrzeugen basierend auf verbessertem YOLOv8 und ByteTrack. Electronics **2024,** 13, **3033. https://doi.org/10.3390/ Elektronik13153033**

Wissenschaftlicher Herausgeber: Fernando De la Prieta Pintado

Empfangen: 11. Juli 2024
Überarbeitet: 29. Juli 2024
Akzeptiert: 30. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

Die Technologie zur Fahrzeugerkennung und -verfolgung spielt eine Schlüsselrolle in intelligenten Transportsystemen [1]. Der Einsatz dieser Technologie ermöglicht die Echtzeitüberwachung des Straßenverkehrsflusses und unterstützt Verkehrsmanagementbehörden dabei, fundierte Entscheidungen zu treffen, um Verkehrsstaus effektiv zu entschärfen. Darüber hinaus kann die Technologie zur Fahrzeugerkennung und -verfolgung Verkehrsverstöße identifizieren und so die Wahrscheinlichkeit von Verkehrsunfällen verringern. Diese Technologie ermöglicht auch eine intelligente Wahrnehmung von Fahrzeugen und Straßenumgebungen und bietet intelligentere und effizientere Lösungen für das Verkehrsmanagement. In den letzten Jahren hat die KI-Technologie eine bedeutende Rolle bei der Verbesserung der Fähigkeiten von Fahrzeugnetzwerken gespielt. Beispielsweise haben Y Ju et al. **[2]** im Bereich der Fahrzeug **-Edge-Network-Computing** Forschungen zum sicheren Offloading **mithilfe der Technologie für nichtorthogonalen Mehrfachzugriff (NOMA) und asynchronem Deep Reinforcement** Learning **durchgeführt . Dies verbessert nicht nur die Sicherheit und Effizienz der Datenübertragung, sondern un**

Bei der Fahrzeugerkennung und -verfolgung treten jedoch häufig mehrere Herausforderungen auf. Faktoren wie komplexe Lichtverhältnisse, Wettereinflüsse, die Vielfalt der Fahrzeugtypen und -größen sowie dichte Verkehrssituationen können die Effektivität der Erkennung und Verfolgung erheblich beeinträchtigen. Als Reaktion auf diese Herausforderungen wurden zahlreiche Forschungsanstrengungen unternommen. Beispielsweise verbesserten D. Roy et al. [3] die Leistung **der Fahrzeugerkennung und -verfolgung durch die Einführung multimodaler Informationen. M. Humayun et al. [4] berücksichtigten basierend auf YOLOv4 die Variationen bei verschiedenen Wetterbedingungen, indem sie einen Spatial Pyramid Pooling Block hinzufügten. P. Deshmukh et al. [5] ver**

Machine Translated by Google

Multiskalen-Features und nutzten ein Bidirektionales Feature-Pyramiden-Netzwerk zur Handhabung der Vielfalt der Fahrzeugmaßstäbe. J. Wang et al. [6] führten auf Basis von YOLOv5 einen Aufmerksamkeitsmechanismus zur Verbesserung der Erkennungsleistung kleiner Fahrzeuge ein und setzten den JDE-Algorithmus zur Fahrzeugverfolgung ein. T. Bui et al. [7] führten ebenfalls auf Basis von YOLOv5 ein Aufmerksamkeitsbasiertes Intra-Scale-Feature-Interaction-Modul zur Beschleunigung der Konvergenz des Erkennungsnetzwerks ein und verbesserten ResNet36 als DeepSORT-Feature-Extraktionsnetzwerk zur Fahrzeugverfolgung. Neben der Verbesserung von Objekterkennungsmodellen und Fahrzeugverfolgungsmethoden auf Basis des Kalman-Filters wurden mehrere technische Lösungen zur Bewältigung dieser Herausforderungen vorgeschlagen. Y Wu et al. [8] schlugen eine Technik unter Verwendung eines Generative Adversarial Network (GAN) vor, um Tagesfahrzeugdatensätze in Nachtdatensätze umzuwandeln und so die nächtliche Erkennungsleistung aus der Perspektive der Datensatzerweiterung zu verbessern.

B Xu et al. [9] führten ein Framework für monokulares Sehen ein, das ein Bildübermalungsmodell verwendet, um verdeckte Fahrzeuge durch Maskierung wiederherzustellen und so die Leistung in Szenarien mit hohem Verkehrsaufkommen zu verbessern. Zusätzlich setzten B Xu et al. [10] eine Methode ein, die Schlüsselpunkterkennung und mehrere Deep-Learning-Modelle kombiniert und die Erkennung und Verfolgung von Fahrzeugachslastpunkten durch Integration der Modelle YOLO-v5 und HourglassNet mit einem Dynamikmodell optimiert. H Xu et al. [11] schlugen ein domänenübergreifendes Fahrzeugerkennungsmodell mit einem integrierten Fahrzeugerkennungsmodell mit Convolutional Block Attention Mechanism (CDCDMA) vor, das ein domänenübergreifendes Framework zur Objekterkennung aufbaut und Scheinwerferfunktionen hervorhebt, um mit Änderungen in komplexen Szenen umzugehen. C Zhang et al. [12] führten ein Systemframework ein , das die Herausforderungen der Fahrzeugwiederidentifizierung aus verschiedenen Ansichten bewältigt, indem es höherstufige Beziehungen und topologische Informationen lernt. Dieses Framework verbessert effektiv die Erkennungsgenauigkeit über verschiedene Ansichten hinweg durch das Modul zur Merkmalsextraktion, das Modul zur Graphenfaltung und das Modul zum Graphenabgleich.

Im Bereich der Fahrzeugerkennung sind die meisten Methoden typischerweise für einzelne Szenarien optimiert . Um dieses Problem zu lösen, schlägt dieses Dokument einen verbesserten Algorithmus auf Basis von YOLOv8 [13] vor – CDS- YOLOv8. Dieser Algorithmus führt während des Downsampling -Prozesses das Context-Guided-Modul **[14]** ein , **um die Merkmalsextraktionsfähigkeiten zu verbessern und sich so an komplexe Szenenvariationen anzupassen. Gleichzeitig wird die Dilated Reparam Block-** Struktur **[15] rekonstruiert und mit dem C2f-Modul integriert, um mehrskalige und spärliche Merkmale vollständig zu extrahieren, wodurch die Leistung des Modells bei der mehrskaligen Objekterkennung verbessert wird. Darüber hinaus verwendet der Algorithmus in der Nachbearbeitungsphase Soft-NMS anstelle von traditionellem NMS, wodurch durch einen sanften Vertrauensabfall mehr gültige Kandidatenboxen in dicht besiedelten Fahrzeugszenen erhalten bleiben. In Bezug auf die Fahrzeugverfolgung konzentrieren sich die meisten Verfolgungsalgorithmen typischerweise nur auf Erkennungsboxen mit hoher Vertrauenswürdigkeit und ignorieren solche mit niedriger Vertrauenswürdigkeit, was dazu führt, dass einige echte Ziele nicht erkannt werden . Um dieses Problem zu lösen, wird in diesem Dokument ByteTrack [16] als Basistracker ausgewählt und es werden Anpassungen am Zustandsvektor und der Kovarianzmatrix im Kalman-Filter vorgenommen. Darüber hinaus wird die Gaussian Smoothed Interpolation [17] eingeführt, um Fahrzeugtrajektorien na**

Die wichtigsten Beiträge dieses Artikels lassen sich wie folgt zusammenfassen:

1. **Wir haben ein verbessertes Fahrzeugerkennungsmodell vorgeschlagen, CDS-YOLOv8: Um die Einschränkungen von YOLOv8 in städtischen Umgebungen für die Fahrzeugerkennung zu beheben, führt das CDS-YOLOv8-Modell während des Downsamplings ein CG-Modul ein, um die Funktionen zur Merkmalsextraktion als Reaktion auf komplexe Szenenänderungen zu verbessern. Es rekonstruiert den DRB, um Probleme mit mehreren Maßstäben, die durch unterschiedliche Kameraabstände und Fahrzeuggrößen ve Darüber hinaus ersetzt Soft-NMS herkömmliche NMS, um mehr gültige Kandidatenboxen beizubehalten und so die Leistung in dicht besiedelten Fahrzeugszenen zu verbessern.**
2. **Wir haben den ByteTrack-Tracker verbessert: ByteTrack wurde als Basis-Tracker ausgewählt. Um die Einschränkungen des ursprünglichen Kalman-Filters bei der Erfassung der Fahrzeugbewegung zu beheben, wurden Verbesserungen am Zustandsvektor und der Kovarianzmatrix vorgenommen, um die nichtlinearen Änderungen der Fahrzeugbewegung besser handhaben zu können. Um Spurunterbrechungen aufgrund von Erkennungsverlusten zu verringern, wurde GSI verwendet, um Verfolgungslücken zu füllen, die Spurkontinuität zu verbessern und Identitätswechsel und Spurverluste zu reduzie**

Machine Translated by Google

3. Die Integration des verbesserten Detektors und Trackers wurde erreicht: CDS-YOLOv8 und der verbesserte ByteTrack wurden kombiniert und an öffentlichen Datensätzen validiert. Experimentelle Ergebnisse zeigen, dass diese Methode in mehreren Metriken eine verbesserte Leistung und Überlegenheit aufweist.

2. Verwandte Arbeiten

2.1. Fahrzeugerkennung

In den letzten Jahren wurden mit der rasanten Entwicklung der Deep-Learning-Technologie erhebliche Fortschritte im Bereich der Objekterkennung erzielt. Auf Deep Learning basierende Objekterkennungsmethoden können in einstufige und zweistufige Ansätze unterteilt werden. Der Vorteil einstufiger Methoden liegt in ihrer schnellen Verarbeitungsgeschwindigkeit, da sie Objektkategorien und Begrenzungsrahmen direkt über ein einziges Netzwerk generieren. Im Gegensatz dazu generieren zweistufige Methoden zuerst Kandidatenbereiche und führen dann eine Feinerkennung durch, mit dem Vorteil einer höheren Erkennungsgenauigkeit . Repräsentative einstufige Methoden umfassen die YOLO-Reihe, SSD [18] und RetinaNet [19], während repräsentative zweistufige Methoden Faster R-CNN [20] und Cascade R-CNN [21] umfassen. Das Arbeitsprinzip der YOLO-Reihe besteht darin, das Eingabebild in ein Raster mit fester Größe aufzuteilen, wobei jedes Raster den Standort und die Kategoriewahrscheinlichkeit des Objekts direkt vorhersagt. SSD extrahiert die Begrenzungsrahmen und Kategorien von Objekten auf Merkmalskarten unterschiedlicher Maßstäbe und erreicht so eine Erkennung in mehreren Maßstäben. RetinaNet verwendet ein Feature Pyramid Network zur Merkmalsextraktion und führt Focal Loss ein, um das Problem des Klassenungleichgewichts zu lösen. Faster R-CNN generiert zunächst Kandidatenregionen über ein Region Proposal Network und führt dann eine Feinklassifizierung der Objekte und eine Bounding-Box-Regression für diese Regionen durch. Cascade R-CNN, das auf Faster R-CNN aufbaut, führt eine mehrstufige Regressions- und Klassifizierungsstruktur ein, die Kandidatenregionen schrittweise verfeinert un

**Um den praktischen Anforderungen der Fahrzeugerkennung gerecht zu werden, wurden in mehreren Forschungsbemühungen allgemeine Detektoren weiter optimiert. Z. Chen et al. [22] schlugen eine verbesserte Methode auf SSD-Basis vor, indem sie Kanalaufmerksamkeitsmechanismen und Dekonvolutionsmodule zur Merkmalsfusion einführten und so die Leistung der Fahrzeugerkennung verbesserten. L. Kang et al. [23] stellten eine auf YOLOv5 basierende Typ-1-Fuzzy-Attention-Methode zur Verbesserung der Fahrzeugerkennung vor. Dieser Ansatz optimiert die Erkennungsleistung, indem er den Detektor auf die Zielmitte fokussiert.
M. Bie et al. [24] schlugen ein bidirektionales Feature-Pyramiden-Netzwerk auf Basis von YOLOv5 vor, um die Fähigkeiten des Modells zur Merkmalsextraktion zu verbessern. B. Wang et al. [25] führten eine verbesserte Methode auf Basis von YOLOv8 ein, indem sie den BiFormer-Aufmerksamkeitsmechanismus und das BiFPN-Netzwerk integrierten, um die Fähigkeiten zur Merkmalsfusion zu verbessern.**

Bei der Auswahl von Erkennungsalgorithmen erreicht YOLOv8 aufgrund seiner schnellen Erkennungsgeschwindigkeit, hohen Genauigkeit und kompakten Modellgröße ein gutes Gleichgewicht zwischen Geschwindigkeit, Genauigkeit und **Modellgröße. Daher wird in diesem Artikel ein auf YOLOv8 basierender CDS-YOLOv8-Algorithmus vorgeschlagen, um Probleme bei der Fahrzeugerkennung unter komplexen Wetterbedingungen, Maßstabsvielfalt und dichten Fahrzeugumgebungen zu lösen. Ohne die Rechenlast zu erhöhen, verbessert dieser Algorithmus effektiv die Fahrzeugerkennungsleistung.**

2.2. Fahrzeugverfolgung für mehrere Ziele

Erkennungsbasierte Multi-Target-Tracking-Algorithmen können in einstufige und zweistufige Methoden unterteilt werden. Einstufige Methoden verwenden im Allgemeinen eine End-to-End-Struktur, bei der Objekterkennung und -zuordnung in einem einzigen Schritt zusammengeführt werden, um direkt Erkennungs- und Tracking- Ergebnisse zu erzielen, mit dem Vorteil einer starken Echtzeitleistung. Zweistufige Methoden führen zuerst die Objekterkennung und dann die Objektzuordnung du Typische einstufige Methoden sind JDE [26] und FairMOT [27]. JDE führt die Objekterkennung und -einbettung im selben Netzwerk durch, wobei die Erkennungs- und Einbettungsteile die meisten Netzwerkparameter gemeinsam nutzen. FairMOT integriert Objekterkennungs- und ReID-Netzwerke in ein Netzwerk, verarbeitet Erkennungs- und Reidentifizierungsaufgaben jedoch über unabhängige Zweige.

Machine Translated by Google

Typische zweistufige Tracking-Algorithmen umfassen die SORT-Reihe. SORT [28], vorgeschlagen von A. Bewley et al., verwendet Faster R-CNN als Detektor in der Erkennungsphase, um Objektstandort- und Kategorieinformationen auszugeben. Die Tracking-Phase vervollständigt die Trajektorienassoziation durch die Kombination des Kalman-Filters und des Un Wojke et al. schlugen den DeepSORT- Algorithmus [29] vor , **der im Vergleich zu SORT die ReID-Methode einführt, indem er ein Merkmalsextraktionsnetzwerk verwendet, um die Erscheinungsmerkmale von Objekten zu erhalten, und während der Datenassoziation einen kaskadierten Abgleich der Erscheinungsmerkmale sowie Intersection Over Union (IoU) durchführt. StrongSORT verbessert basierend auf DeepSORT außerdem die Extraktion von Erscheinungsmerkmalen, die Verarbeitung von Bewegungsrauschen und die Kamerabewegungskompensation und ergänzt und optimiert Trajektorien mit fehlenden Erkennungen und Assoziationen . Y. Zhang et al. [16] schlugen den ByteTrack-Algorithmus vor, der die Tracking- Leistung durch Assoziation fast aller Erkennungsboxen verbessert. Anschließend schlugen N. Aharon et al.** den BoT-SORT- **Algorithmus [30] vor ,** der den Kalman-Filter auf der Grundlage von **ByteTrack verbessert und eine Kamerabewegungskompensation und die ReID-Methode hinzuf**

In Fahrzeugverfolgungsszenarien entwickelten Z. Zhao et al. [31] eine neue gemeinsame Bewertungsstrategie auf Basis des JDE-Frameworks und führten CBAM- und Transformer-Encoder ein , um die Erkennungsfunktionen zu verbessern und so die Verfolgungsleistung zu steigern. K. Zhang et al. [32] integrierten den CAENet-Aufmerksamkeitsmechanismus in YOLOX, um die Fahrzeugerkennung zu verbessern, und verwendeten DeepSORT zur Fahrzeugverfolgung. G. Han et al. [33] verfolgten Fahrzeuge auf Basis von FairMOT, optimierten das DLA-34-Netzwerk, um die Erkennungsleistung zu verbessern, und führten eine gemeinsame Verlustfunktion ein, um Probleme beim Identit

Im Vergleich zu einstufigen Tracking-Algorithmen sind zweistufige Tracking-Algorithmen flexibler und erreichen in komplexen Szenarien eine höhere Genauigkeit. Daher wird in diesem Dokument ByteTrack basierend auf CDS-YOLOv8 für die Fahrzeugerkennung und -verfolgung verbessert. Im Gegensatz zu den meisten Tracking-Methoden, die Erkennungsboxen mit niedriger Punktzahl ignorieren, berücksichtigt ByteTrack Erkennungsboxen mit niedriger Punktzahl bei der Trajektorienzuordnung vollständig und behebt so effektiv das Problem fehlender Trajektorienfragmente. Wenn beispielsweise eine teilweise Verdeckung auftritt, können die Erkennungsbox-Punktzahlen von hoch bis niedrig variieren. Durch die Berücksichtigung von Boxen mit niedriger Konfidenzpunktzahl kann ByteTrack teilweise verdeckte Ziele i

3. Methoden

Beim Erkennen und Verfolgen von Fahrzeugen gibt der CDS-YOLOv8-Detektor zunächst die Erkennungsboxen für das aktuelle Bild aus. Der verbesserte Kalman-Filter sagt dann die Begrenzungsboxen für das aktuelle Bild durch Tracklet-Vorhersage voraus. Anschließend werden die Erkennungsboxen und die vorhergesagten Boxen mithilfe des Ungarischen Algorithmus abgeglichen, wobei IoU als Grundlage für die Zuordnung dient. Der spezifische Prozess des sekundären Abgleichs ist in Abbildung 1 dargestellt .

Die mit CDS-YOLOv8 ausgegebenen Erkennungsboxen werden basierend auf ihren Vertrauenswerten in Boxen mit hoher und niedriger Vertrauenswürdigkeit unterteilt, gefolgt von zwei Runden der Assoziationsverarbeitung. In der ersten Assoziationsrunde sagt der verbesserte Kalman-Filter die verfolgten Tracklets und die verlorenen Tracklets voraus und gleicht dann die Boxen mit hoher Vertrauenswürdigkeit mit den vorhergesagten Boxen ab. Wenn der Abgleich erfolgreich ist, wird der Tracklet-Status aktualisiert; wenn nicht, durchlaufen die verfolgten Tracklets, die nicht mit Boxen mit hoher Vertrauenswürdigkeit abgeglichen wurden, eine zweite Assoziationsrunde mit Boxen mit niedriger Vertrauenswürdigkeit. In der zweiten Assoziationsrunde wird der Tracklet-Status aktualisiert, wenn der Abgleich erfolgreich ist; wenn nicht, wechselt das Tracklet in den verlorenen Status. Bei der Tracklet-Verwaltung werden Tracklets, die länger als 30 Frames nicht abgeglichen bleiben, gelöscht. Neue Tracklets werden nicht abgeglichenen Erkennungsboxen mit hoher Vertrauenswürdigkeit zugewiesen. Schließlich wird nach der Fahrzeugerkennung und -verfolgung GSI verwendet, um Lücken in den Trajektorien zu füllen, die durc

Das CDS-YOLOv8-Modell besteht aus drei Teilen: Rückgrat, Hals und Kopf, wie in Abbildung 2 dargestellt:

Machine Translated by Google

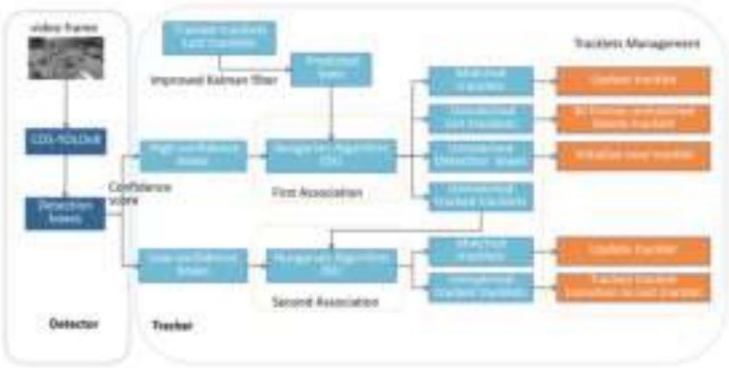


Abbildung 1. Flussdiagramm zur Fahrzeugerkennung und -verfolgung.

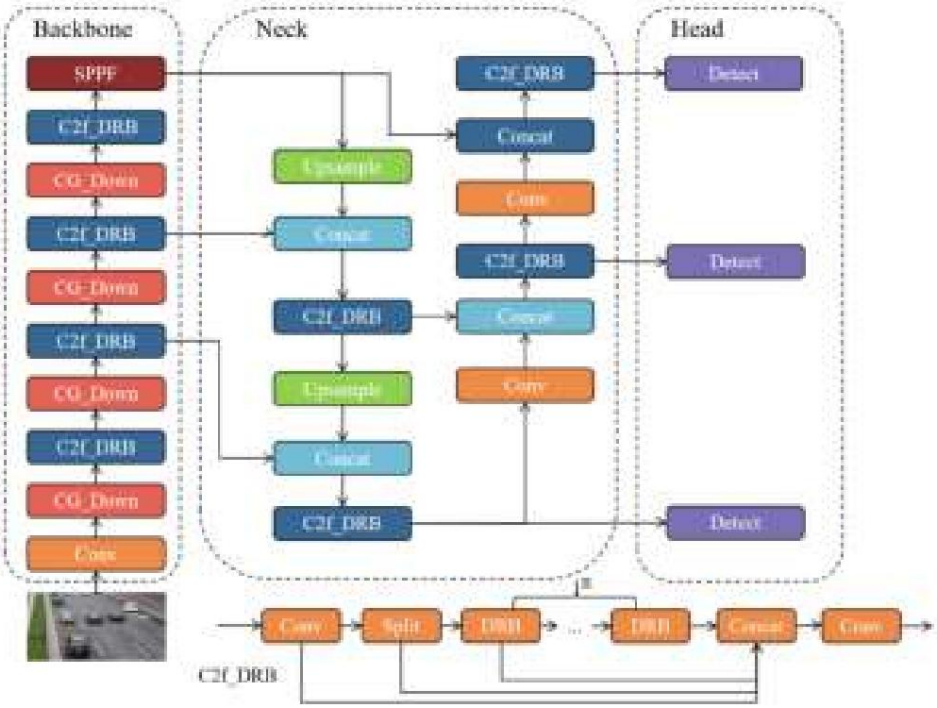


Abbildung 2. Struktur von CDS-YOLOv8.

In Abbildung 2 wird das Backbone zur Merkmalsextraktion verwendet, wobei mehrstufige Merkmale aus dem Eingabebild extrahiert werden. Der Neck ist für die Merkmalsfusion zuständig und kombiniert tiefe und flache Merkmale mithilfe von FPN und PAN. Der Head enthält drei Erkennungsköpfe, die Merkmalskarten unterschiedlicher Maßstäbe verarbeiten und die Positions- und Kategorieinformationen der Erkennungsboxen ausgeben. CDS-YOLOv8 behebt die Mängel von YOLOv8 bei der Erkennung von Stadtfahrzeugen mit den folgenden Verbesserungen: Das CG-Modul wird während des Downsamplings eingeführt und verbessert die Merkmalsextraktion in komplexen Umgebungen, indem lokale und umgebende Kontextinformationen kombiniert und globale Informationen zur Gewichtung verwendet werden. Im Backbone und Neck wird die DRB-Struktur neu strukturiert und in das C2f-Modul integriert, um Probleme mit der Maßstabsvielfalt besser zu lösen. In der Nachbearbeitungsphase ersetzt Soft-NMS das herkömmliche NMS, wodurch die aggressive Eliminierung von Kandidatenboxen vermiede

3.1. Kontextgesteuertes Downsampling

Die Umgebungen, in denen Fahrzeuge eingesetzt werden, umfassen komplexe Szenarien wie Nächte, bewölktes Wetter, sonnige Tage und regnerische Tage. Um die Funktionen zur Merkmalsextraktion in diesen komplexen Umgebungen zu verbessern, wird das kontextgesteuerte Modu

Machine Translated by Google

Downsampling-Prozess. Das CG-Modul umfasst mehrere Schlüsselkomponenten: Der Local Feature Extractor verwendet Standard-Faltung, um lokale Merkmale des Bildes zu extrahieren und liefert feinkörnige Merkmale und Randinformationen. Der Surrounding Context Extractor verwendet Dilated Convolution, um das rezeptive Feld lokaler Merkmale zu erweitern und so ein breiteres Spektrum an Kontextinformationen zu erfassen. Der Joint Feature Extractor verknüpft lokale Merkmale und umgebende Kontextmerkmale und verarbeitet sie mit Batch Normalization (BN) und SiLU-Aktivierung, wodurch gemeinsame Merkmale gebildet werden, die nichtlineare Ausdrucksmöglichkeiten verbessern. Der Global Context Extractor aggregiert globale Kontextinformationen des gesamten Bildes durch Global Avera
Dadurch werden den gemeinsamen Merkmalen Gewichte zugewiesen, wodurch nützliche Informationen verstärkt und irrelevante Informationen unterdrückt werden. Die Struktur des CG-Moduls ist in Abbildung 3 dargestellt:

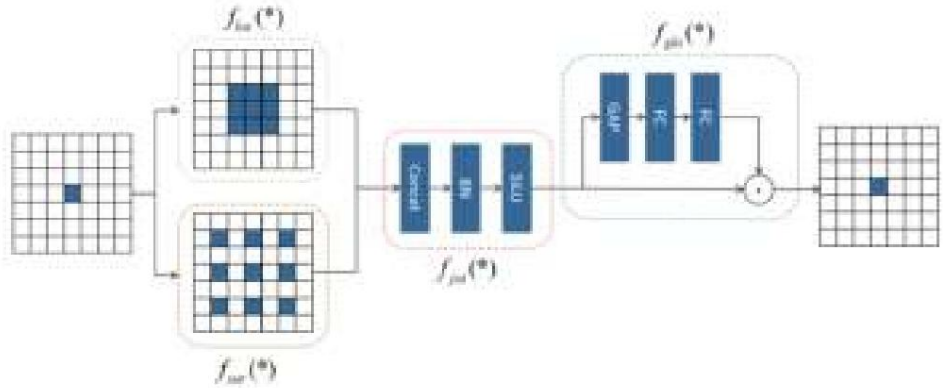


Abbildung 3. Struktur des kontextgesteuerten Moduls.

Im Einzelnen wird die Eingabe-Feature-Map zunächst einer 3 × 3-Faltung unterzogen, um lokale Features zu extrahieren, und einer 3 × 3-Faltung mit einer Dilatationsrate von 2, um umgebende Kontext-Features zu extrahieren. Als Nächstes verschmilzt der Joint Feature Extractor die lokalen Features und die umgebenden Kontext-Features entlang der Kanaldimension, gefolgt von der BN-Schicht und der SiLU-Aktivierung . Schließlich nutzt der Global Context Extractor die globalen Informationen des Bildes und verwendet Global Average Pooling, um einen globalen Feature-Vektor für jeden Kanal zu generieren. Dieser Vektor wird dann einer nichtlinearen Transformation durch zwei vollständig verbundene Schichten unterzogen, wodurch Kanalgewichte erzeugt werden. Diese Gewichte werden zur gewichteten Optimierung auf die gemeinsamen Features angewendet, wodurch wichtige Features hervorgehoben und unwichtige unterdrückt werden, wodurch die Feature-Extraktionsfähigkeit des Modells verbessert wird.

3.2. Rekonstruierter erweiterter Reparam-Block

Bei der Fahrzeugerkennung kommt es aufgrund unterschiedlicher Kameraabstände und Fahrzeuggrößen zu erheblichen Maßstabsabweichungen. Um dieses Problem zu lösen, haben wir den Dilated Reparam Block rekonstruiert und in die C2f-Struktur integriert. Die DRB-Struktur besteht aus einer Faltungsschicht mit großem Kernel und mehreren dilatierten Faltungsschichten mit kleinem Kernel. Die Faltungsschicht mit großem Kernel erfasst breitere räumliche Merkmale, während die dilatierten Faltungsschichten mit kleinem Kernel durch Dilatationsoperationen spärliche räumliche Merkmale erfassen. Die Ausgaben dieser Schichten werden durch die BN-Schicht verarbeitet und dann summiert. Mithilfe von Reparametrisierungstechniken werden sie äquivalent als eine einzelne Faltungsschicht mit großem Kernel dargestellt. Dieses Design hat zwei Hauptvorteile: Die Faltungsschichten mit kleinem Kernel erfassen durch Einbeziehung der Dilatation Merkmale über einen größeren Bereich. In Kombination mit der Faltungsschicht mit großem Kernel wird das rezeptive Feld effektiv erweitert, wodurch die Notwendigkeit entfällt, die Modelltiefe zu erhöhen oder größere Faltungskerne zu entwerfen. Während der Inferenz werden durch Neuparametrisierungstechniken mehrere Faltungsschichten zu einem einzigen gleichwertigen Faltungskernel zusammengeführt. Dabei bleiben die Kernelgröße und die Rechenkomplexität erhalten, während die Merkmalsextraktionsfunktion des Modells deutlich verbessert wird. Abbildung 4 veranschaulicht die DRB-Struktur.

Machine Translated by Google

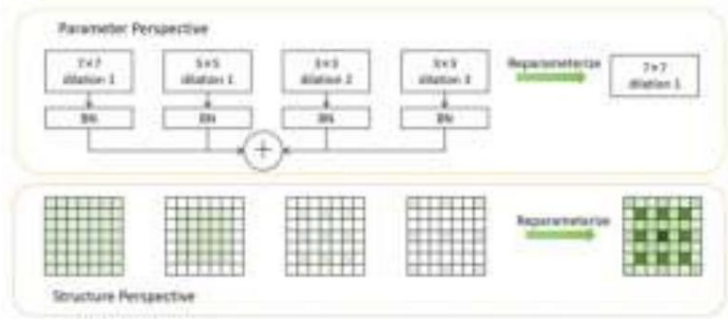


Abbildung 4. Struktur des erweiterten Reparamblocks.

Aus Parametersicht besteht der DRB aus einer 7 × 7-Faltungsschicht mit großem Kernel und drei Faltungsschichten mit kleinem Kernel und unterschiedlichen Dilatationsraten. Die Ausgaben aller Faltungsschichten werden durch BN-Schichten verarbeitet und während der Inferenz äquivalent **als einzelne 7 × 7-Faltungsschicht mit großem Kernel dargestellt ,** ohne **die Rechenlast wesentlich zu erhöhen.**

Strukturell zeigt Abbildung 4 , wie jede Faltungsschicht Bereichen basierend auf Kernelgröße und Dilatationsrate entspricht. Bei großen Objekten erweitert der DRB das rezeptive Feld effektiv, indem er Faltungsschichten mit großem Kernel mit dilatierten Faltungsschichten kombiniert. Bei kleinen Objekten erfassen die dilatierten Faltungsschichten im DRB mithilfe kleiner Kernel und unterschiedlicher Dilatationsraten effizient feinkörnige Merkmale und verbessern so die Leistung der Fahrzeugerkennung in mehreren Maßstäben vor komplexen Hintergründen.

3.3. Soft-NMS

Beim Fahrzeugerkennungs- und -verfolgungsprozess wird Soft-NMS anstelle von herkömmlichem NMS verwendet, um die Leistung in dicht besiedelten Fahrzeugszenarien zu verbessern. Die herkömmlichen NMS-Schritte sind wie folgt: 1.

Sortieren Sie die Kandidatenboxen in absteigender Reihenfolge basierend auf ihren Vertrauenswerten. 2.

Wählen Sie die Kandidatenbox mit dem höchsten Konfidenzwert aus und fügen Sie sie dem endgültigen Ergebnissatz hinzu.

3. Berechnen Sie den IoU zwischen dieser Kandidatenbox und den anderen Boxen. Wenn der IoU einer Kandidatenbox einen festgelegten Schwellenwert überschreitet, wird sie gelöscht.

4. Wählen Sie aus den verbleibenden Kandidatenfeldern das Feld mit dem höchsten Konfidenzwert aus und wiederholen Sie die Schritte 2 und 3, bis alle Kandidatenfelder verarbeitet wurden.

Soft-NMS optimiert den dritten Schritt des traditionellen NMS: In Soft-NMS wird nach der Berechnung des IoU zwischen der Kandidatenbox mit dem höchsten Vertrauenswert und den anderen Boxen statt der Boxen mit einem IoU über dem Schwellenwert eine sanfte Handhabungsmethode verwendet, indem ihre Vertrauenswerte verringert werden. Konkret verwenden wir eine Gaußsche Zerfallsmethode, wie in Gleichung (1) gezeigt:

$$si = si \times \exp\left(\ddot{y} \quad \frac{IoU2}{\ddot{y}}\right) \tag{1}$$

wobei si den Konfidenzwert der i-ten Kandidatenbox darstellt und ÿ der Abklingfaktor ist.

Wenn der gesunkene Vertrauenswert unter dem festgelegten Schwellenwert liegt, wird die Kandidatenbox entfernt. In Szenarien mit dichtem Fahrzeugverkehr kommt es zu einer starken Überlappung der Kandidatenboxen zwischen den Zielen . Im Vergleich zu herkömmlichen NMS kann Soft-NMS mehr gültige Kandidatenboxen beibehalten, benachbarte Ziele effektiver handhaben und die Anzahl ver

3.4. Verbesserter Kalman-Filter

Bei zweistufigen Tracking-Algorithmen wird die Ausgabe des Begrenzungsrahmens normalerweise mithilfe des Kalman-Filters (KF) geschätzt, um den Bewegungszustand des Ziels zu bestimmen. In ByteTrack wird der Zustandsvektor des Kalman-Filters in Gleichung (2) dargestellt:

Machine Translated by Google

$$x = xc, yc , a, h, \quad \dot{xc}, \dot{yc} , \dot{ein}, \dot{h}^{T} \tag{2}$$

wobei xc, yc die Mittelpunktskoordinaten, a das Seitenverhältnis und h die Höhe sind.

Bei teilweiser Okklusion und Perspektivwechseln ist das Seitenverhältnis dann schwer anzupassen. Daher ist die direkte Schätzung der Breite und Höhe die besser

Der verbesserte Zustandsvektor ist in Gleichung (3) dargestellt:

$$xk = xc(k), yc\,(k), w(k), h(k), \dot{xc}(k), \dot{yc}\,(k), \dot{w}(k), \dot{h}(k) \tag{3}$$

wobei (xc (k), **yc (k) die Mittelpunktskoordinaten, w(k) die Breite und h(k) die Höhe sind.**

Im Vergleich zur Verwendung des Seitenverhältnisses bietet die direkte Schätzung der Breite und Höhe folgende Vorteile: Einerseits kann sie die tatsächliche Größe des Objekts genauer wiedergeben und sich besser an Änderungen der Fahrzeugabmessungen anpassen; andererseits sind Änderungen der Form und Größe des Fahrzeugs während der Bewegung komplex, und die Verwendung des Seitenverhältnisses erhöht nichtlineare Fehler. Durch die direkte Schätzung der Breite und Höhe kann die **durch die Ansammlung von Fehlern verursachte Verzerrung des Begrenzungsrahmens verringert werden ,** wodurch **in komplexen Szenen eine bessere Stabilität aufrechterhalten wird.**

Im klassischen Kalman-Filter werden die Prozessrausch-Kovarianzmatrix Q und die Messrausch-Kovarianzmatrix R als feste Werte festgelegt, die während des gesamten Tracking-Prozesses unverändert bleiben. Um die Leistung des Kalman-Filters in dynamischen Umgebungen zu verbessern, werden Q und R dynamisch entsprechend den Echtzeitbedingungen angepasst, Q wie in Gleichung (4) und R wie in Gleichung (5) dargestellt.

$$Qk = \mathrm{diag}\left(\ddot{y}pw\hat{}^{2}\,k\ddot{y}1|k\ddot{y}1,\ \ddot{y}ph\hat{}^{2}\,k\ddot{y}1|k\ddot{y}1,\ \ddot{y}pw\hat{}^{2}\,k\ddot{y}1|k\ddot{y}1,\ \ddot{y}ph\hat{}^{2}\,k\ddot{y}1|k\ddot{y}1,\ \ddot{y}vw\hat{}^{2}\,k\ddot{y}1|k\ddot{y}1,\ \ddot{y}vh\hat{}^{2}\,k\ddot{y}1|k\ddot{y}1,\ \ddot{y}vw\hat{}^{2}\,k\ddot{y}1|k\ddot{y}1,\ \ddot{y}vh\hat{}^{2}\,k\ddot{y}1|k\ddot{y}1 \right) \tag{4}$$

wobei ÿp die Standardabweichung des Positionsrauschens und ÿp die Standardabweichung des Geschwindigkeitsrauschens ist.

$$Rk = \mathrm{diag}\left(\ddot{y}mw\hat{}^{2}_{k|k\ddot{y}1},\ \ddot{y}mh\hat{}^{2}_{k|k\ddot{y}1},\ \ddot{y}mw\hat{}^{2}_{k|k\ddot{y}1},\ \ddot{y}mh\hat{}^{2}_{k|k\ddot{y}1} \right) \tag{5}$$

wobei ÿm die Standardabweichung des Messrauschens ist.

Zeitabhängige Q- und R-Werte können flexibler auf nichtlineare Änderungen in der Umgebung reagieren , wie etwa die Verzögerung oder Beschleunigung eines Fahrzeugs. Im Vergleich zu festen Kovarianzmatrizen können zeitabhängige Q- und R-Werte basierend auf den aktuellen Vorhersage- und Messzuständen angepasst werden, wodurch Probleme wie Rauschen und Fehlerakkumulation effektiver angegangen werden. Dies ermöglicht eine bessere Anpassung an dynamische Umgebungen und verbessert die Gesamtleistung des Systems.

3.5. Gaußsche geglättete Interpolation

Um das Problem des Trajektorienverlusts zu lösen, der durch Erkennungsauslassungen bei der Fahrzeugverfolgung verursacht wird, wird die Methode der gaußschen geglätteten Interpolation als Nachbearbeitungsschritt zur Optimierung eingesetzt. Herkömmliche lineare Interpolationsmethoden basieren normalerweise nur auf den Positionen zweier Endpunkte und ignorieren die Komplexität der Fahrzeugbewegung, was zu weniger genauen Ergebnissen führt. GSI erfasst die nichtlineare Bewegungstrajektorie des Fahrzeugs durch die Gaußsche Prozessregression (GPR) und liefert dadurch präzis

Für die i-te Fahrzeugtrajektorie wird das GSI-Modell durch Gleichung (6) dargestellt:

$$pt = f\,(i)\,(t) + \ddot{y} \tag{6}$$

Machine Translated by Google

wobei t die Frame-ID, pt die Koordinate der Fahrzeugposition bei Frame t, f (i) (t) die von GPR angepasste Funktion zur Beschreibung der Fahrzeugtrajektorie und ÿ ÿ **N 0,** ÿ **das Gaußsche Rauschen darstellt .** 2

Der Gaußsche Prozess ist eine leistungsstarke nichtparametrische Bayes-Methode, die die nichtlinearen Beziehungen zwischen Datenpunkten erfassen kann, vorausgesetzt, die Funktion f (i) folgt eine Gaußscher Prozess, wie in Gleichung (7) dargestellt:

$$\mathbf{f}\ \mathbf{(i)}\ \ddot{y}\ \mathbf{GP\ 0,\ kx,}\ x\ddot{y} \tag{7}$$

wobei die Kovarianzfunktion k(x, xÿ) **der Radial Basis Function (RBF)-Kernel ist, der besonders empfindlich auf lokale Änderungen reagiert. Je näher die Datenpunkte beieinander liegen, desto höher ist ihre Korrelation, und die Reaktion des RBF-Kernels nimmt schnell ab, wenn der Abstand zwischen den Datenpunkten zunimmt. Diese Eigenschaft macht den RBF-Kernel besser geeignet, um nichtlineare Änderungen der Fahrzeugbewegung zu beschreiben, wie z. B. Anfahren, Wenden und Beschleunigen oder Abbremsen. Genauer gesagt wird dies wie in Gleichung (8) dargestellt:**

$$\mathbf{kx,}\ x\ddot{y} = \mathbf{exp}\ \ddot{y} \quad \frac{\ddot{y}\ \mathbf{x}\ \ddot{y}\ \mathbf{x} \quad |^{2}}{\mathbf{2}\ 2\ddot{y}} \tag{8}$$

Der Prozess der Vorhersage und Interpolation der Trajektorienposition läuft wie folgt ab: Für Frames mit fehlenden Detektionen wird die geglättete Position Pÿ **wie in Gleichung (9) gezeigt vorhergesagt:**

$$\mathbf{P}^{\ddot{y}} = K(F^{\ddot{y}}, F)\ \mathbf{K(F,}\ F) + \ddot{y} \quad {}^{2}{}_{\text{ICH}}{}^{\ddot{y}1}\ \mathbf{P} \tag{9}$$

wobei K(·, ·) **die Kovarianzmatrix basierend auf der Kovarianzfunktion** k(·, ·) ist, **I die Identitätsmatrix ist und P die bekannten Fahrzeugpositionen der Trajektorie darstellt.**

Darüber hinaus **bestimmt der Skalierungsparameter** ÿ für die RBF-Kernelfunktion ihre Empfindlichkeit **gegenüber Datenänderungen. Während der Fahrzeugverfolgung können unterschiedliche Fahrzeugtrajektorienlängen unterschiedlichen Bewegungsmustern entsprechen. Daher können durch Anpassen von** ÿ **zur Anpassung an die Länge l der Fahrzeugtrajektorie die Trajektorien geglättet werden, wodurch Rauschen reduziert und eine gleichmäßigere und genauere Verfolgung erreicht wird. Die Einstellung von** ÿ **is**

$$\ddot{y} = \ddot{y} \cdot \mathbf{log} \quad \frac{\mathbf{3}\ \ddot{y}}{\mathbf{m}} \tag{10}$$

wobei *ÿ* **eine Konstante zur Steuerung des Glättungsgrads ist, der auf Grundlage experimenteller Ergebnisse aus der Rastersuche auf 18 festgelegt wurde.**

Durch die oben genannten Schritte kann GSI fehlende Trajektorien im Fahrzeug effektiv ergänzen Tracking, wodurch die Genauigkeit und Stabilität von Flugbahnvorhersagen verbessert wird.

4. Experimente und Ergebnisse

4.1. Datensatz und experimentelle Konfiguration

Der UA-DETRAC-Datensatz [34], der gemeinsam von der University of Alberta und der University of Science and Technology of China veröffentlicht wurde, dient als Benchmark-Datensatz, der speziell für die Erkennung und Verfolgung von Fahrzeugen mit mehreren Objekten entwickelt wurde. Dieser Datensatz umfasst 10 Stunden Videomaterial, das mit einer Canon EOS 550D-Kamera (Canon USA, Inc., Lake Success, NY, USA) aufgenommen wurde. Die Videos werden mit einer Bildrate von 25 Bildern pro Sekunde aufgezeichnet und umfassen über 140.000 Bilder und 1,21 Millionen annotierte Objekterkennungsboxen. Der Trainingssatz besteht aus 60 Videosequenzen mit insgesamt 83.791 Bildern, während der Testsatz 40 Videosequenzen mit insgesamt 56.340 Bildern enthält. Der UA-DETRAC-Datensatz stellt verschiedene Herausforderungen dar. Repräsentative Szenarien sind in Abbildung 5 dargestellt. Der Datensatz deckt vier Fahrzeugkategorien ab: Autos, Busse, Lieferwagen und andere. Darüber hinaus umfasst es vier Wetterbedingungen: sonnig, Nacht, bewölkt

Machine Translated by Google



Abbildung 5. Repräsentative Szenarien im UA-DETRAC-Datensatz.

Bei den Fahrzeugerkennungsexperimenten wurden zunächst Bilder ohne Beschriftung ausgeschlossen, sodass ein Trainingsset von 82.085 Bildern und ein Testset von 56.167 Bildern entstand. Um ein Overfitting zu vermeiden, wurde jedes 10. Bild abgetastet, sodass letztlich 8209 Bilder für das Training und 5617 Bilder für die Auswertung entstanden. Für die Fahrzeugverfolgungsexperimente wurden 40 Videosequenzen aus dem Testset zur Auswertung verwendet.

Das in diesen Experimenten verwendete GPU-Gerät ist eine NVIDIA GeForce RTX 4090 (NVIDIA Corporation, Santa Clara, CA, USA), und das CPU-Gerät ist eine Core i9-9900K @ 3,60 GHz CPU (Intel Corporation, Chengdu, China). Die Experimente wurden auf einem Windows 10-System mit einer Deep-Learning-Umgebung basierend auf PyTorch 1.31.1, CUDA 11.6 und Python 3.8 durchgeführt. Die Trainingsparameter wurden wie folgt festgelegt: Die Anzahl der Trainingsepochen betrug 200, die Batchgröße betrug 32 und während des Trainings wurden keine vorab trainierten Gewichte verwendet .

4.2. Bewertungsmaßstäbe

Bei der Fahrzeugerkennung werden Präzision (P), Rückruf (R), mAP@0,5 und mAP@0,5:0,95 verwendet, um die Versuchsergebnisse auszuwerten.

P: der Anteil der echt-positiven Proben an allen als positiv erkannten Proben, wie in Gleichung (11) dargestellt:

$$P = \frac{TP}{TP + FP} \tag{11}$$

R: der Anteil der korrekt erkannten positiven Proben an allen tatsächlich positiven Proben, wie in Gleichung (12) dargestellt:

$$R = \frac{TP}{TP + FN} \tag{12}$$

Dabei ist TP die Anzahl der echt-positiven Proben, FP die Anzahl der falsch-positiven Proben und FN die Anzahl der tatsächlich positiven Proben, die fälschlicherweise als negative Proben erkannt wurden.

Durchschnittliche Präzision (AP): die Fläche unter der PR-Kurve für eine bestimmte Klasse von Proben, wie in Gleichung (13) dargestellt:

$$AP = \int_0^1 P dR \tag{13}$$

mAP@0.5 ist die mittlere durchschnittliche Präzision für alle Klassen mit einem IoU-Schwellenwert von 0,5.

mAP@0.5:0.95 ist die mittlere durchschnittliche Präzision für alle Klassen mit IoU-Schwellenwerten im Bereich von 0,5 bis 0,95 mit einer Schrittweite von 0,05.

Bei der Fahrzeugverfolgung werden MOTA, IDF1 und IDS zur Auswertung der experimentellen Ergebnisse verwendet.

MOTA misst die Genauigkeit des Tracking-Algorithmus und berücksichtigt dabei verschiedene Fehler , die während des Tracking-Prozesses auftreten, wie etwa falsche Positivwerte, verpasste Erkennungen und Identitätswechsel. Dies wird in Gleichung (14) dargestellt:

Machine Translated by Google

$$\text{MOTA} = 1 - \frac{\sum_t (\text{FNt} + \text{FPt} + \text{IDSWt})}{\sum_t \text{GTt}} \quad (14)$$

wobei FNt die Anzahl der verpassten Erkennungen zum Zeitpunkt t darstellt, FPt die Anzahl von falschen Positivmeldungen zum Zeitpunkt t, IDSWt stellt die Anzahl der Identitätswechsel zum Zeitpunkt t dar und GTt stellt die Anzahl der Ground-Truth-Ziele zum Zeitpunkt t dar.

IDF1: IDF1 misst die Identitätskonsistenz und Kontinuität des Tracking-Algorithmus, wie in Gleichung (15) gezeigt:

$$\text{IDF1} = 2 \cdot \frac{\text{IDP} \cdot \text{IDR}}{\text{IDP} + \text{IDR}} \quad (15)$$

wobei IDP den Anteil der korrekt identifizierten Ziele unter den erkannten Ziele. IDR stellt das Verhältnis der korrekt erkannten Zielidentitäten zur tatsächlichen Zielidentitäten.

IDS: Die Anzahl der Identitätswechsel, die während des Tracking-Prozesses auftreten, d.ÿh. dem gleichen Ziel werden unterschiedliche Identitäten zugewiesen. Ein niedrigerer Wert bedeutet eine bessere Leistung des Trackers bei der Aufrechterhaltung der Identitätskonsistenz.

4.3. Detektorablationsexperimente

Um die Wirksamkeit der vorgeschlagenen Verbesserungsmodule bei der Fahrzeugerkennung zu überprüfen, Ablationsexperimente wurden mit YOLOv8-S als Basismodell durchgeführt. Die Ergebnisse sind in Tabelle **1 aufgeführt :**

Tabelle 1. Ergebnisse des Ablationsexperiments im UA-DETRAC-Datensatz.

Modell	P	R	Karte@ 0,5	Karte@0,5:0,95	Flops
YOLOv8-S	71,8 %	58,6 %	64,2 %	46,6 %	28,8 G
YOLOv8-S+CG_Down	78,4 %	63,3 %	69,9 %	51,0 %	30,7 G
YOLOv8-S+C2f_DRB	81,5 %	62,6 %	71,4 %	51,6 %	24,7 G
YOLOv8-S+CG_Down+C2f_DRB	80,7 %	63,9 %	72,9 %	52,2 %	26,6 G
YOLOv8-S+CG_Down+C2f_DRB+Soft-NMS	82,4 %	62,3 %	73,2 %	55,4 %	26,6 G

Wie aus den Daten in Tabelle 1 ersichtlich, erhöhte die Einführung des CG-Moduls im Downsampling-Prozess P um 6,6 %, R um 4,7 %, mAP@0,5 um 5,7 % und mAP@0,5:0,95 um 4,4 %, was die Wirksamkeit des CG-Moduls bei der Extraktion von Fahrzeugmerkmalen bestätigt. Einführung des Moduls C2f_DRB erhöhte P um 9,7 %, R um 4,0 %, mAP@0,5 um 7,2 %, und mAP@0.5:0.95 um 5%, was den signifikanten Vorteil von C2f_DRB bei der Handhabung zeigt Multiskalenprobleme. Als die Module CG und C2f_DRB gleichzeitig eingeführt wurden, stieg P um 8,9 %, R um 5,3 %, mAP@0,5 um 8,7 % und mAP@0,5:0,95 um 5,6 %, Dies beweist, dass die Kombination dieser beiden Module das Fahrzeug besser extrahieren und integrieren kann Funktionen. Schließlich ersetzt das Ersetzen von traditionellem NMS durch Soft-NMS für die Nachbearbeitung leicht erhöhte mAP@0.5 um 0,3 Prozentpunkte, verbesserte mAP@0.5:0.95 jedoch deutlich um 3,2 %, was die Grundlage für eine präzisere Fahrzeugverfolgung in nachfolgenden Phasen bildet. Im Vergleich zum Basismodell, mit einer Reduzierung von 2,2 G in Flops, stieg P um 10,6 %, R um 3,7 %, mAP@0.5 um 9 % und mAP@0.5:0.95 um 8,8 %, was die Überlegenheit von CDS-YOLOv8 im Umgang mit Fahrzeugerkennungsszenarien.

Um die Fortschritte des verbesserten Modells bei der Fahrzeugerkennung weiter zu demonstrieren, Eine visuelle Analyse der Versuchsergebnisse wird in Abbildung 6 dargestellt.
Bilder von links nach rechts stellen das Originalbild dar, die Vorhersageergebnisse der Basislinie Modell YOLOv8-S, die Vorhersageergebnisse des verbesserten Modells CDS-YOLOv8 und die Anpassungen von CDS-YOLOv8 gegenüber YOLOv8-S, darunter die Reduzierung in verpassten Fahrzeugen. Die optimierten Teile sind mit rosa Kreisen markiert. Zusätzlich ist die Schwarze Bereiche in den Bildern sind ignorierte Bereiche, da die nachfolgende Fahrzeugverfolgung nur konzentriert sich auf fahrende Fahrzeuge auf der Hauptstraße.

Machine Translated by Google



Abbildung 6. Visualisierung der Erkennungsergebnisse.

In Abbildung 6 haben wir drei repräsentative Szenarien ausgewählt: sonnig, nachts und regnerisch. Im sonnigen Szenario, wie im rosa Kreis dargestellt, erkannte CDS-YOLOv8 einen Bus, der YOLOv8 verfehlte, sowie Autos, die weiter von der Kamera entfernt waren. Im Nachtszenario erkannte CDS-YOLOv8 fast alle Ziele, die YOLOv8 verfehlte, und identifizierte genau teilweise verdecktes weißes Auto, das YOLOv8 nicht als gültiges Ziel erkannte. Im regnerischen Szenario erkannte CDS-YOLOv8 fast alle kleinen Ziele, die weit von der Kamera entfernt waren und hat ein Taxi identifiziert, das teilweise verdeckt war und von YOLOv8 nicht als gültiges Ziel erkannt wurde.

In diesen Szenarien zeigte CDS-YOLOv8 signifikante Verbesserungen gegenüber YOLOv8-S in den Funktionen zur Merkmalsextraktion und zur Verarbeitung mehrerer Maßstäbe. Darüber hinaus wurden mehr valide Ziele in dichten Fahrzeugszenen durch die Handhabung stark überlappender Kandidatenboxen mithilfe von Soft-NMS mit Vertrauensverlust, anstatt diesen einfach durch einen IoU-Schwellenwert zu eliminieren . Zusammenfassend lässt sich sagen, dass CDS-YOLOv8 unter verschiedenen Licht- und Wetterbedingungen, effektive Handhabung von Problemen auf mehreren Ebenen und Szenarien mit dichtem Fahrzeugaufkommen.

4.4. Vergleichsexperiment der Detektoren

Um die Weiterentwicklung von CDS-YOLOv8 bei der Fahrzeugerkennung zu demonstrieren, verglichen wir mit gängigen Erkennungsalgorithmen. Die Ergebnisse sind in Tabelle 2 dargestellt:

Tabelle 2. Vergleichende experimentelle Ergebnisse des UA-DETRAC-Datensatzes. (Die besten Ergebnisse sind fett hervorgehoben.)

Modell	Karte@0,5	Karte@0,5:0,95	Flops
Faster-RCNN [20]	66,3 %	38,8 %	60,5 G
Kaskaden-RCNN [21]	65,2 %	41,1 %	88,1 G
RetinaNet [19]	56,6 %	32,1 %	45,2 G
ATSS [35]	58,8 %	32,1 %	49,2 G
YOLOv5-S [36]	62,0 %	44,9 %	16 G
YOLOv6-S [37]	59,0 %	43,0 %	44,9 G
YOLOv7 [38]	71,1 %	53,9 %	105,2 G
YOLOv8-S [13]	64,2 %	46,6 %	28,8 G
YOLOv8-M [13]	68,2 %	50,6 %	79,3 G
YOLOv8-L [13]	71,6 %	51,4 %	165,7 G
YOLOv9-S [39]	65,1 %	48,0 %	39,6 G
L. Kang et al. [23]	70,0 %	50,3 %	26,7 G
CDS-YOLOv8 (unser)	73,2 %	55,4 %	26,6 G

Aus den Vergleichsergebnissen geht hervor, dass CDS-YOLOv8 die fortschrittlichsten Leistung im Vergleich zu diesen herkömmlichen zweistufigen und einstufigen Erkennungsmethoden.

Machine Translated by Google

Die Leistung übertrifft nicht nur die neueste Version von YOLOv9-S, sondern übertrifft auch die großen Version von YOLOv8, mit der Verbesserungen von 1,6 % und 4 % bei mAP@0.5 und mAP@0.5:0.95 erreicht wurden, CDS-YOLOv8 bietet eine bessere Balance zwischen Genauigkeit und Rechenleistung Komplexität, was seine Überlegenheit in Szenarien zur Fahrzeugerkennung weiter beweist.

4.5. Tracker-Ablationsexperiment

Zur Überprüfung der Wirksamkeit des verbesserten Detektors und der vorgeschlagenen Verbesserung Module in der Fahrzeugverfolgung, wurde ein Ablationsexperiment wie in Tabelle 3 gezeigt durchgeführt. Die experimentellen Ergebnisse sind Durchschnittswerte basierend auf 40 Testsequenzen:

Tabelle 3. Ergebnisse des Ablationsexperiments im UA-DETRAC-Datensatz.

Detektor + Tracker	mIDF1	IDS	mMOTA
YOLOv8s+ BYTETrack	76,9 %	855	67,2 %
CDS-YOLOv8+ BYTETrack	78,2 %	785	70,6 %
CDS-YOLOv8+ BYTETrack+ Verbessertes KF	79,7 %	717	73,1 %
CDS-YOLOv8+ BYTETrack+ Verbessertes KF +GSI	80,3 %	530	73,9 %

Im Vergleich zum ursprünglichen Detektor erhöht die Einführung des verbesserten Detektors CDS-YOLOv8 mIDF1 um 1,3 %, reduziert IDS um 70 und verbessert mMOTA erheblich um 3,4 %. Dies spiegelt die entscheidende Rolle des Detektors im Trackingsystem wider. Nach der Verbesserung des Trackers durch Optimierung des Kalman-Filters erhöht sich mIDF1 um 1,5 %, IDS sinkt um 68 und mMOTA verbessert sich um 2,5 %. Dies deutet darauf hin, dass die verbesserte Kalman Filter können die nichtlineare Bewegung von Fahrzeugen flexibler handhaben. Auf dieser Grundlage wird die Einführung GSI zur Verarbeitung der Tracking-Trajektorien erhöht mIDF1 weiter um 0,6 %, deutlich reduziert IDS um 187 und verbessert mMOTA um 0,8 %, was die Wirksamkeit von diese Methode beim Füllen von Flugbahnlücken. Mit dem verbesserten Detektor und Tracker im Vergleich Im Vergleich zu vor den Verbesserungen steigt mIDF1 um 3,4 %, IDS sinkt um 325 und mMOTA verbessert sich um 6,7 %, was eine deutliche Steigerung der Leistung der Fahrzeugverfolgung bedeutet.

Um die Wirksamkeit des verbesserten Detektors und Trackers intuitiver zu demonstrieren Bei der Fahrzeugverfolgung wurde eine Visualisierungsanalyse der Verfolgungsergebnisse durchgeführt, wie gezeigt in Abbildung 7. Die vier Bildspalten von links nach rechts sind das Originalbild, das Tracking Ergebnisse mit dem ursprünglichen Detektor und Tracker, die Tracking-Ergebnisse mit dem verbesserten Detektor und Tracker sowie die Optimierungseffekte des verbesserten Modells im Vergleich zum Originaldetektor und -verfolger.



Abbildung 7. Visualisierung der Tracking-Ergebnisse.

In Abbildung 7 haben wir auch drei repräsentative Szenarien ausgewählt: sonnig, nachts und regnerisch. Im sonnigen Szenario, wie im rosa Kreis dargestellt, CDS-YOLOv8 – im Vergleich zu YOLOv8 – erkannte nicht nur mehr verpasste Ziele, sondern lieferte auch genauere Begrenzungen Box-Erkennung. CDS-YOLOv8 konnte beispielsweise eine weißer Lieferwagen und ein Taxi am Bildrand, während YOLOv8 nur einen Teil des

Machine Translated by Google

Region als Begrenzungsrahmen. Darüber hinaus war die Tracking-ID des weißen Lieferwagens 1281 mit der ursprüngliche Detektor und Tracker, aber nach Verbesserungen sank er deutlich auf 1099, was auf eine deutliche Verbesserung der Tracking-Konsistenz hinweist. Im Nachtszenario ist die Verbesserter Detektor identifizierte mehr kleine Ziele und die Tracking-ID des Busses auf der links sank deutlich von 3060 auf 2529. Im Regenszenario identifizierte YOLOv8 die Front eines Polizeiautos als separates Fahrzeug, während CDS-YOLOv8 korrekt erkannte die gesamte Begrenzungsbox des Polizeiautos. Darüber hinaus die Tracking-ID des Polizeiautos nach den Verbesserungen deutlich von 5898 auf 5026 gesunken.

Insgesamt reduzierte der verbesserte Detektor nicht nur die Anzahl der Fehlerkennungen sondern auch eine genauere Erkennung von Begrenzungsrahmen ermöglicht. Die deutliche Reduzierung der ID Zahlen für dasselbe Ziel zeigen, dass der verbesserte Tracker eine hohe Genauigkeit beibehielt und Konsistenz bei der Fahrzeugverfolgung. Der verbesserte Detektor und Tracker sind besser ausgestattet um komplexe Szenarien und nichtlineare Änderungen bei Fahrzeugen in städtischen Umgebungen zu bewältigen.

4.6. Vergleichsexperimente von Trackern

Um die Weiterentwicklung des verbesserten Trackers zu validieren, wurden Vergleichsexperimente durchgeführt mit herkömmlichen zweistufigen Objektverfolgungsalgorithmen durchgeführt. Um Fairness zu gewährleisten, Bei den Experimenten wurde CDS-YOLOv8 als Detektor verwendet. Die Tracking-Ergebnisse sind in Tabelle 4 dargestellt:

Tabelle 4. Vergleichsergebnisse des UA-DETRAC-Datensatzes. (Die besten Ergebnisse sind fett hervorgehoben.)

Tracker	mIDF1	IDS	mMOTA
DeepSORT [29]	73,5 %	2731	70,0 %
StrongSORT [17]	76,8 %	2994	71,3 %
OCSORT [40]	76,8 %	2191	71,9 %
DeepOCSORT [41]	77,1 %	1830	72,1 %
BYTETrack [16]	78,2 %	785	70,6 %
BOT-SORT [30]	78,6 %	727	72,9 %
Verbesserter BYTETrack (unserer)	80,3 %	530	73,9 %

Aus den Vergleichsergebnissen ist ersichtlich, dass der verbesserte BYTETrack Leistung auf dem neuesten Stand der Technik. Analyse der Ergebnisse von BYTETrack und BOT-SORT Tracking Algorithmen ist es offensichtlich, dass die Berücksichtigung von Scoreboxen mit geringer Konfidenz signifikant reduzieren IDS und beweisen damit ihre Wirksamkeit bei der Identitätskonsistenz. Darüber hinaus bietet DeepSORT, StrongSORT, DeepOCSORT und BOT-SORT verwenden alle die ReID-Methode, die reduziert die Betriebsgeschwindigkeit des Netzwerks. Im Gegensatz dazu wurde der verbesserte BYTETrack in Dieses Dokument verwendet nicht die ReID-Methode und erzielt dennoch die höchste Leistung über alle drei Tracking-Metriken, wodurch ein gutes Gleichgewicht zwischen Tracking-Geschwindigkeit und Genauigkeit erreicht wird.

5. Schlussfolgerungen

Dieses Dokument befasst sich mit den Herausforderungen bei der Fahrzeugerkennung und -verfolgung. wie komplexe Szenarien, Maßstabsvielfalt durch unterschiedliche Fahrzeuggrößen und Kamera Winkel und partielle Okklusionen in dichten Szenen. Eine Reihe von Verbesserungen wurden vorgeschlagen, um diese Probleme angehen.

In Bezug auf die Fahrzeugerkennung führen wir das CG-Modul während des Downsamplings ein Prozess, der lokale und umgebende Merkmale zu gemeinsamen Merkmalen kombiniert. Diese Gemeinsame Merkmale werden dann unter Verwendung globaler Informationen gewichtet, was die Merkmalsextraktion verbessert Leistungsfähigkeit des Detektors. Zusätzlich wurde das DRB-Modul rekonstruiert und integriert in das C2f-Modul. Durch die Verwendung von Faltungsschichten mit großem Kernel und mehreren Faltungsschichten mit kleinem Kernel Es wurden Faltungen mit Dilatationen, verschiedenen Skalen und spärlichen Merkmalen extrahiert. Die Reparametrisierungstechnologie wurde eingesetzt, um äquivalente Faltungskerne zu erzeugen, was für eine bessere umfassende Merkmalsextraktion während der Inferenzphase, wodurch Multiskalenprobleme effektiv angegangen werden. Wir haben auch das traditionelle NMS durch Soft-NMS ersetzt, das bewältigt partielle Okklusionsszenarien besser durch sanften Vertrauensabfall. In dem Aspekt der Fahrzeugverfolgung verbesserten wir den Zustandsvektor des Kalman-Filters und führten

Machine Translated by Google

eine zeitbezogene Kovarianzmatrix, um besser mit dynamischen Fahrzeugbewegungen umgehen zu können. Zusätzlich wurde die GSI-Methode integriert, um Trajektorienlücken zu füllen, die durch verpasste Erkennungen verursacht wurden. Wir führten Experimente mit dem UA-DETRAC-Datensatz durch, wobei wir Ablationsstudien verwendeten, um die Rolle jedes Moduls zu demonstrieren, und visuelle Analysen einsetzten, um die Optimierungseffekte des verbesserten Modells intuitiv darzustellen. Vergleichsexperimente zeigten die verbesserte Leistung des CDS-YOLOv8 und

In Zukunft werden wir uns stärker auf die Fahrzeugerkennung und -verfolgung aus verschiedenen Perspektiven konzentrieren und dabei vielfältigere Modalitäten wie Infrarotbilder, Radar und andere Methoden untersuchen, um die Wirksamkeit der Fahrzeugerkennung und -verfolgung zu v

Beiträge der Autoren: Konzeptualisierung, LY; Methodik, LY; Software, LY; Validierung, LY und YC; formale Analyse, CS; Untersuchung, CS und RL; Datenkuratierung, LY und CX; Schreiben – Überprüfung und Bearbeitung, LY; Visualisierung, LY; Überwachung, YC; Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Forschung wurde durch das Talent Fund-Projekt der China West Normal University (Nr. 463177) finanziert.

Datenverfügbarkeitserklärung: Die in der Studie vorgestellten Originalbeiträge sind im Artikel enthalten; weitere Anfragen können an den Erstautor gerichtet werden.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Verweise

1. **Boukerche, A.; Tao, Y.; Sun, P. Künstliche Intelligenz-basierte Verkehrsflussvorhersagemethoden zur Unterstützung intelligenter Transportsysteme.** Comput. Netw. **2020,** 182, **107484.** [CrossRef]
2. **Ju, Y.; Cao, Z.; Chen, Y.; Liu, L.; Pei, Q.; Mumtaz, S.; Dong, M.; Guizani, M. NOMA-unterstütztes sicheres Offloading für Fahrzeug-Edge- Computing-Netzwerke mit asynchronem Deep Reinforcement Learning.** IEEE Trans. Intell. Transp. Syst. **2024,** 25, **262** [Querverweis]
3. **Roy, D.; Li, Y.; Jian, T.; Tian, P.; Chowdhury, K.; Ioannidis, S. Multimodale Sensorik und Datenfusion zur Erkennung mehrerer Fahrzeuge.** IEEE Trans. Multimed. **2023,** 25, **2280–2295.** [CrossRef]
4. **Humayun, M.; Ashfaq, F.; Jhanjhi, NZ; Alsadun, MK Verkehrsmanagement: Mehrstufige Fahrzeugerkennung bei unterschiedlichem Wetter Bedingungen unter Verwendung von YOLOv4 und Spatial Pyramid Pooling Network.** Electronics **2022,** 11, **2748.** [CrossRef]
5. **Deshmukh, P.; Satyanarayana, GSR; Majhi, S.; Sahoo, Großbritannien; Das, SK Swin Transformer Based Vehicle Detection in Undisziplinierter Verkehrsumgebung.** Expert Syst. Appl. **2023,** 213, **118992.** [CrossRef]
6. **Wang, J.; Dong, Y.; Zhao, S.; Zhang, Z. Eine hochpräzise Methode zur Fahrzeugerkennung und -verfolgung basierend auf dem Aufmerksamkeitsmechanismus .** Sensors **2023,** 23, **724.** [CrossRef] [PubMed]
7. **Bui, T.; Wang, G.; Wei, G.; Zeng, Q. Fahrzeug-Mehrobjekterkennungs- und -verfolgungsalgorithmus basierend auf der verbesserten You Only Look Once 5s-Version und DeepSORT.** Appl. Sci. **2024,** 14, **2690.** [CrossRef]
8. **Wu, Y.; Wang, T.; Gu, R.; Liu, C.; Xu, B. Algorithmus zur nächtlichen Fahrzeugerkennung basierend auf Bildübersetzungstechnologie 1.** J. Intell. Fuzzy Syst. **2024,** 46, **5377–5389.** [CrossRef]
9. **Xu, B.; Liu, X.; Feng, G.; Liu, C. Ein monokulares Framework zur genauen Identifizierung der räumlich-zeitlichen Verteilung von Fahrzeugradlasten unter Okklusionsszenarien.** Eng. Appl. Artif. Intell. **2024,** 133, **107972.** [CrossRef]
10. **Xu, B.; Liu, C. Auf Schlüsselpunkterkennung basierende und Multi-Deep-Learning-Modell-integrierte Methode zur Ermittlung der Fahrzeugachslast Räumlich-zeitliche Verteilung.** Adv. Eng. Inform. **2024,** 62, **102688.** [CrossRef]
11. **Xu, H.; Lai, S.; Li, X.; Yang, Y. Domänenübergreifendes Autoerkennungsmodell mit integriertem Convolutional-Block-Aufmerksamkeitsmechanismus.** Image Vis. Comput. **2023,** 140, **104834.** [CrossRef]
12. **Zhang, C.; Yang, C.; Wu, D.; Dong, H.; Deng, B. Cross-View-Fahrzeugneuidentifizierung basierend auf Graph Matching.** Appl. Intell. **2022,** 52, **14799–14810.** [CrossRef]
13. **Ultralytics. YOLOv8: Objekterkennung und Bildsegmentierung in Echtzeit. GitHub-Repository. Online verfügbar: https://github.com/ultralytics/ultralytics (abgerufen am 28. Juli 2024).**
14. **Wu, T.; Tang, S.; Zhang, R.; Cao, J.; Zhang, Y. CGNet: Ein leichtes kontextgesteuertes Netzwerk für semantische Segmentierung.** IEEE Trans. Bildprozess. **2021,** 30, **1169–1179.** [CrossRef] [PubMed]
15. **Ding, X.; Zhang, Y.; Ge, Y.; Zhao, S.; Song, L.; Yue, X.; Shan, Y. UniRepLKNet: Ein universelles Perception Large-Kernel ConvNet für Audio-, Video-, Punktwolken-, Zeitreihen- und Bilderkennung.** arXiv **2024, arXiv:2311.15599.**
16. **Zhang, Y.; Sun, P.; Jiang, Y.; Yu, D.; Weng, F.; Yuan, Z.; Luo, P.; Liu, W.; Wang, X. ByteTrack: Multi-Object Tracking durch Zuordnung aller Erkennungsboxen. In** Computer Vision—ECCV 2022, Proceedings der 17. Europäischen Konferenz, Tel Aviv, Israel, 23.–27. Oktober 2022; **Avidan, S., Brostow, G., Cissé, M., Farinella, GM, Hassner, T., Hrsg.; Springer Nature Switzerland: Cham, Schweiz, 2022; S. 1–21.**

Machine Translated by Google

17. **Du, Y.; Zhao, Z.; Song, Y.; Zhao, Y.; Su, F.; Gong, T.; Meng, H. StrongSORT: Machen Sie DeepSORT wieder großartig.** IEEE Trans. Multimed. **2023,** 25, **8725–8737. [CrossRef]**
18. **Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, AC SSD: Single Shot MultiBox Detector. In** Computer Vision—ECCV 2016, Proceedings der 14. Europäischen Konferenz, Amsterdam, Niederlande, 11.–14. Oktober 2016; **Leibe, B., Matas, J., Sebe, N., Welling, M., Hrsg.; Springer International Publishing: Cham, Schweiz, 2016; S. 21–37.**
19. **Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollar, P. Brennpunktverlust bei der Erkennung dichter Objekte. In Proceedings of the IEEE International Konferenz über Computer Vision (ICCV), Venedig, Italien, 22.–29. Oktober 2017; S. 2980–2988.**
20. **Girshick, R. Fast R-CNN. In Proceedings der IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7.–13. Dezember 2015; S. 1440–1448.**
21. **Cai, Z.; Vasconcelos, N. Cascade R-CNN: Eintauchen in die Objekterkennung hoher Qualität. In Proceedings der IEEE-Konferenz zu Computer Vision und Mustererkennung (CVPR), Salt Lake City, UT, USA, 18.–23. Juni 2018; S. 6154–6162.**
22. **Chen, Z.; Guo, H.; Yang, J.; Jiao, H.; Feng, Z.; Chen, L.; Gao, T. Schneller Fahrzeugerkennungsalgorithmus im Straßenverkehr basierend auf verbesserter SSD.** Measurement **2022,** 201, **111655. [CrossRef]**
23. **Kang, L.; Lu, Z.; Meng, L.; Gao, Z. YOLO-FA: Typ-1-Fuzzy-Aufmerksamkeit basierender YOLO-Detektor zur Fahrzeugerkennung.** Expertensystem. Appl. **2024,** 237, **121209. [CrossRef]**
24. **Bie, M.; Liu, Y.; Li, G.; Hong, J.; Li, J. Echtzeit-Fahrzeugerkennungsalgorithmus basierend auf einem leichten „You-Only-Look-Once“-Algorithmus (YOLOv5n-L) Ansatz.** Expert Syst. Appl. **2023,** 213, **119108. [CrossRef]**
25. Wang, B.; Li, Y.-Y.; Xu , W .; **Wang, H.; Hu, L. Fahrzeug-Fußgänger-Erkennungsmethode basierend auf verbessertem YOLOv8.** 13, **2149. [CrossRef]**
26. **Wang, Z.; Zheng, L.; Liu, Y.; Li, Y.; Wang, S. Auf dem Weg zum Echtzeit-Multi-Objekt-Tracking. In** Computer Vision—ECCV 2020, Proceedings der 16. Europäischen Konferenz, Glasgow, Großbritannien, 23.–28. August 2020; **Vedaldi, A., Bischof, H., Brox, T., Frahm, J.-M., Hrsg.; Springer International Publishing: Cham, Schweiz, 2020; S. 107–122.**
27. **Zhang, Y.; Wang, C.; Wang, X.; Zeng, W.; Liu, W. FairMOT: Zur Fairness der Erkennung und Wiederidentifizierung bei der Verfolgung mehrerer Objekte.** Int. J. Comput. Vis. **2021,** 129, **3069–3087. [CrossRef]**
28. **Bewley, A.; Ge, Z.; Ott, L.; Ramos, F.; Upcroft, B. Einfaches Online- und Echtzeit-Tracking. In Proceedings der 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25.–28. September 2016; S. 3464–3468.**
29. **Wojke, N.; Bewley, A.; Paulus, D. Einfaches Online- und Echtzeit-Tracking mit einer Deep Association Metric. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Peking, China, 17.–20. September 2017; S. 3645–3649.**
30. **Aharon, N.; Orfaig, R.; Bobrovsky, B.-Z. BoT-SORT: Robuste Assoziationen bei der Verfolgung mehrerer Fußgänger.** arXiv **2022, arXiv:2206.14651.**
31. **Zhao, Z.; Ji, Z.; Yao, Y.; He, Z.; Du, C. Verbessertes Erkennungsmodell und gemeinsame Bewertungsstrategie für die Verfolgung mehrerer Fahrzeuge.** IEEE Zugang **2023,** 11, **30807–30818. [CrossRef]**
32. **Zhang, K.; Wu, F.; Sun, H.; Cai, M. Monokulare Fahrzeuggeschwindigkeitserkennung basierend auf verbessertem YOLOX und DeepSORT.** Neural Comput. Appl. **2024,** 36, **9643–9660. [CrossRef]**
33. **Han, G.; Jin, Q.; Rong, H.; Jin, L.; Zhang, L. Fahrzeugverfolgungsalgorithmus basierend auf Deep Learning in der Straßenrandperspektive.** Sustainability **2023,** 15, **1950. [CrossRef]**
34. **Wen, L.; Du, D.; Cai, Z.; Lei, Z.; Chang, MC; Qi, H.; Lim, J.; Yang, M.-H.; Lyu, S. UA-DETRAC: Ein neuer Benchmark und ein neues Protokoll zur Erkennung und Verfolgung mehrerer Objekte.** Comput. Vis. Image Underst. **2020,** 193, **102907. [CrossRef]**
35. **Zhang, S.; Chi, C.; Yao, Y.; Lei, Z.; Li, SZ Überbrückung der Lücke zwischen ankerbasierter und ankerfreier Erkennung durch adaptive Auswahl von Trainingsbeispielen. In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung (CVPR), Seattle, WA, USA, 13.–19. Juni 2020; S. 9759–9768.**
36. **Ultralytics. YOLOv5: Familie von Architekturen und Modellen zur Objekterkennung, vortrainiert auf dem COCO-Datensatz. GitHub-Repository. Online verfügbar: https://github.com/ultralytics/yolov5 (abgerufen am 28. Juli 2024).**
37. **Li, C.; Kleine.; Jiang, H.; Weng, K.; Geng, Y.; Kleine.; Ke, Z.; Li, Q.; Cheng, M.; Nie, W.; et al. YOLOv6: Eine einstufige Objekterkennung Framework für industrielle Anwendungen.** arXiv **2022, arXiv:2209.02976.**
38. **Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-YM YOLOv7: Trainable Bag-of-Freebies setzt neue Maßstäbe für Echtzeit-Objektdetektoren . In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung (CVPR), Vancouver, BC, Kanada, 17.–24. Juni 2023; S. 7464–7475.**
39. **Wang, C.-Y.; Yeh, I.-H.; Liao, H.-YM YOLOv9: Lernen, was Sie lernen möchten, mithilfe programmierbarer Gradienteninformationen.** arXiv **2024, arXiv:2402.13616.**
40. **Cao, J.; Pang, J.; Weng, X.; Khirodkar, R.; Kitani, K. Beobachtungszentriertes SORT: Neue Überlegungen zu SORT für robustes Multi-Object-Tracking. In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung (CVPR), Vancouver, BC, Kanada, 17.–24. Juni 2023; S. 9686–9696.**
41. **Maggiolino, G.; Ahmad, A.; Cao, J.; Kitani, K. Deep OC-Sort: Multi-Pedestrian Tracking durch adaptive Neuidentifizierung. In Proceedings der 2023 IEEE International Conference on Image Processing (ICIP), Kuala Lumpur, Malaysia, 8.–11. Oktober 2023; S. 3025–3029.**

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Entwicklung der Antennenstrahlungsparameter für Übergang von Luft zu Plasma

Tomasz Aleksander Mi´s

Abteilung für Plasmaphysik, Weltraumforschungszentrum der Polnischen Akademie der Wissenschaften, ul. Bartycka 18A, 00-716 Warszawa, Polen; tmis@cbk.waw.pl oder tomasz.a.mis@mailplus.pl

Zusammenfassung: Dieses Dokument enthält eine Beschreibung der Antennenparameter in Bezug auf ihre Strahlungs-/Empfangsfähigkeiten, die von den Plasmaparametern in der die Antenne umgebenden Umgebung beeinflusst werden, und ergänzt vorhandene Arbeiten zu den Antennenparametern (z. B. Impedanz oder Ströme). Die berücksichtigten Parameter sind die Radien der Strahlungszonen (induktiv, Fresnel, Fraunhofer), die Ausknickung und die Richtwirkung; außerdem wird eine Methode zur Umwandlung des in Luft/Vakuum gemessenen Strahlungs-/Empfangsmusters in das für gegebene Plasmabedingungen erwartete Muster berücksichtigt. Drei verschiedene vereinfachte Plasmabedingungen (unterschiedliche Elektronendichten: 1,4 × 1012 mÿ3 4 × 1011 mÿ3 und 108 mÿ3) werden berücksichtigt, mit unterschiedlichen Antennenlängen (1 m, 10 m, 100 m) und Signalausbreitungsmodi (klassisch- ionosphärisch, Whistler und Alfvén). **Die Erkenntnisse zeigen, dass die vorgestellten Antennenparameter und ihre Strahlungs-/ Empfangsmuster stark von den Plasmabedingungen abhängen. Diese Erkenntnisse können dazu genutzt werden, zusätzliche Anforderungen und Einschränkungen an die mechanische Konstruktion neuer Instrumente für Weltraumwettermessungen an Bord von Raumfahrzeugen zu formulieren (z. B. die Antennen vom Raumfahrzeug weg zu bewegen, um ihre Strahlungs-/Empfangsmuster nicht zu verändern oder das Plasma um das Raumfahrzeug herum nicht zu messen) oder eine genauere Datenverarbeitung von vorhandenen Weltraumwettersatelliten zu ermöglichen, die beispielsweise eine präzisere Triangulation der Signa**

Schlüsselwörter: Plasma; Plasmasphäre; Magnetosphäre; Weltraumantenne; aktive Antenne; Antennenparameter; Strahlungsmuster; Antennenprüfung; Richtwirkung; Scalloping

check for updates

Zitat: Mi´s, TA Evolution von Antennenstrahlungsparameter für Übergang von Luft zu Plasma. Elektronik **2024,** 13, **3040. https://doi.org/ 10.3390/Elektronik13153040**

Wissenschaftliche Herausgeber: Jun Wang, Kuikui Fan, Jianfeng Zhu und Fan Wu

Empfangen: 27. Juni 2024
Überarbeitet: 23. Juli 2024
Akzeptiert: 26. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 beim Autor. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung; Elektromagnetische Emissionen und Antennen in Plasmen

Seit Beginn des Weltraumzeitalters und der Erforschung der erdnahen Umwelt galt das Interesse der Untersuchung der Eigenschaften der den Planeten umgebenden Hülle aus ionisierten Teilchen – der Plasmasphäre –, ihrer Veränderlichkeit in Abhängigkeit von den Sonnenbedingungen, ihres Einflusses auf die sich ausbreitenden Radiosignale und ihrer Wirkung auf Raumfahrzeuge und bemannte Weltraummissionen [1]. Seit den 1960er Jahren wurden zahlreiche natürlich vorkommende Radiophänomene identifiziert, die durch die Wechselwirkungen zwischen dem Sonnenwind und der Magnetosphäre der Erde verursacht werden (z. B. die Aurorale Kilometrierstrahlung [2–4] und andere Arten von Emissionen [5]). Viele dieser Phänomene werden innerhalb des terrestrischen Wellenleiters (dem Raum zwischen der Erdoberfläche und den unteren Schichten der Ionosphäre) erzeugt und können sich entweder entlang der magnetischen Feldlinien oder unabhängig von ihnen weiter durch die Ionosphäre ausbreiten **[6–8]. Diese Emissionen variieren stark in ihren Ausbreitungswegen [9] und können als Fernerkundungsinstrument für die Ionosphäre [10] und die Plasmasphäre [8], einschließlich der Magnetopause und der plasmasphärischen Bugstoßwelle [11], eingesetzt werden. Die Ausbreitung dieser Signale hängt stark von den lokalen Plasmabedingungen sowohl in der Ionosphäre als auch in der Plasmasphäre ab, was beispielsweise während des Auftretens der Röntgen-Sonneneruptionen untersucht wurde, die beide ionisierten Umgebungen erheblich beeinflussten [12–15]. Es ist erwähnenswert, dass die komplexen Änderungen der Parameter des ionisierten Mediums (die nicht nur die Elektronendichte, sondern auch die Protonendichte umfassen) auch für die Analyse der planetennah**

Machine Translated by Google

Da viele dieser natürlichen Emissionen am Boden [6] und in Höhen näher an der Ionosphäre [4,17] relativ leicht empfangen werden können, sind **für ihre Ausbreitung durch die Magnetosphäre oft Messungen vor Ort (manchmal nur vor Ort [8]) erforderlich. Diese werden mit speziellen Raumfahrzeugen und Instrumenten durchgeführt, z. B. Interball-1 [18], POLRAD an Bord von Interball-2 [2], DEMETER [7], SWARM [3], PROGNOZ-8 [19], MAGION-4 [20], ISEE 1 und 2 [5], CESAR [21], CORONAS [21], CLUSTER [22] oder Arase [23]. Um die effektive Antennenöffnung für den Empfang und in manchen Fällen auch für aktive Experimente zu maximieren, werden Weltraumantennen üblicherweise als große (bis zu Dutzenden Metern Gesamtlänge) Dipole konstruiert, entweder fachwerk- oder drahtartig, mit optionaler Halterung für andere kleinere Sensoren, z. B. Magnetometer [23,24] oder Langmuir-Sonden [25]. Während sich Dipole in Luft-/Vakuumumgebung gut berechnen und testen lassen, weisen sie in einer ionisierten Umgebung deutlich andere Parameter und Leistungen auf. Es wurden aufwendige Arbeiten durchgeführt, um die Belastungsparameter solcher Antennen in Bezug auf den Winkel der lokalen Magnetfeldstärke [26–28] oder die Strommuster/Modi in den Antennendrähten in Abhängigkeit von den lokalen Plasmaparametern [29,30] zu bestimmen. Bei aktiven/sendenden Antennen ist die Bestimmung der Größe der Plasmahülle, die den Strahler umgibt und an der Bildung der Radiowelle beteiligt ist [31],** von entscheidender Bedeutung . **Außerdem sind die Formeln für die elektrische und magnetische Feldstärke im dreidimensionalen Raum (entscheidend für die Einhaltung der Genauigkeit [32]) um die Antenne herum zu bestimmen, beides approximiert für anisotropes P**

Diese Analysen zeigen deutlich die erheblichen Unterschiede zwischen den Luft-/Vakuum- und den Plasma-Fällen hinsichtlich des Antennendesigns und der grundlegenden Parameter der Antenne, wie Eingangsimpedanz und Reaktanz oder Strahlungsmuster. Allerdings werden nur wenige Informationen zur tatsächlichen räumlichen Strahlungsleistung der Antenne gegeben, z. B. wie sich das Strahlungsmuster unter verschiedenen Plasmabedingungen bildet und entwickelt (dies ist entscheidend, um z. B. die Ankunftsrichtung des registrierten Signals zu bestimmen). Diese Unterschiede sind sowohl in der Phase des Antennendesigns als auch im Antennenbetrieb wichtig, der darauf abzielt, wissenschaftlich nützliche Daten zu liefern, die die Phänomene, die innerhalb der Plasmasphäre auftreten, so genau wie möglich beschreiben. Während die ausgefeilteren und genaueren Formulierungen für verschiedene Plasmatypen während der Entwurfsphase der Antenne einbezogen werden können, kann die Überprüfung und Prüfung der Parameter (z. B. des Strahlungsmusters) praktisch nur unter Luftbedingungen durchgeführt werden, was eine einfache Methode zur Umwandlung der Ergebnisse in Luft in Plasmaergebnisse erfordert. Darüber hinaus verändert die Antenne während der Entstehung des eigentlichen Plasmas um die Weltraumantenne ständig ihr Strahlungsmuster und ihre Strahlungseigenschaften (Ausbuchtung, Richtwirkung, Strahlungszonen) – sowohl im Bereich der Plasmavariablen als auch im Betriebsfrequenzbereich. Die Bewertung dieser Änderungen, auch in einer vereinfachten, aber leicht umsetzbaren Form, ist entscheidend für die Analyse der empfangenen Spektrogramme (z. B. zur Bestimmung des Einfallswinkels der aufgezeichneten Signale oder der Empfindlichkeit des Antennensystems unter den gegebenen Plasmabedingungen im gegebenen Frequenzbereich) sowie für die Bewertung des möglichen Einflusses externer Objekte – Ausleger, Satellitenstrukturen, andere Antennen in der Nähe – auf das Plasma, dem die Antenn

Unter Berücksichtigung aller oben genannten Überlegungen werden in diesem Dokument vereinfachte Analysemethoden (für isotrope Umgebungen, Wellenzahl als eindimensionale Variable und weiter unten angegebene Magnetfeldbedingungen) zur Bestimmung der erwarteten Antennenparameter – Strahlungsmuster, Scalloping, Richtwirkung und spektrale Empfindlichkeit – für die **gegebenen plasmatischen und geometrischen Bedingungen zusammengefasst. Dadurch lässt sich der Betrieb der Antenne in einem Plasma besser beurteilen , die Antennenbewertung während des Tests wird erleichtert und die Genauigkeit der Analyse der registrierten Spektrogramme (auch Spektrogramme der oben genannten Weltraummissionen) wird verbessert. Dadurch wird die Lücke zwischen den vorhandenen Antennenanalysen (z. B. Impedanz- und Strommuster) und ihrer tatsächlichen Leistung im umgebenden Weltraum effektiv geschl**

Machine Translated by Google

2. Entwicklung der Strahlungszonen

Die Strahlungszonen um die Antenne herum können als die Bereiche beschrieben werden, in denen die von der Antenne abgestrahlte Energie zunächst (nahe der Antennenoberfläche) in reaktiver Form im freien Raum gespeichert wird und sich dann – mit zunehmender Entfernung zur Antenne – die Radiowellenfront mit ihren grundlegenden elektrischen und magnetischen Feldkomponenten bildet; die Energie nimmt die Strahlungsform an **[36]. Abbildung 1 zeigt das grundlegende Schema der Zonen um einen symmetrischen Dipol, wobei der Fresnel-Zonenradius** rFresnel angegeben ist, **hinter dem die reaktive Form oder emittierte Energie vorherrscht, und der Fraunhofer-Zonenradius** rFraunhofer, **nach dem die Radiowellenfront vollständig geformt ist. Zusätzlich ist für eine physikalische Antennenabmessung** h **kleiner als die Wellenlänge ÿ auch die durch die**

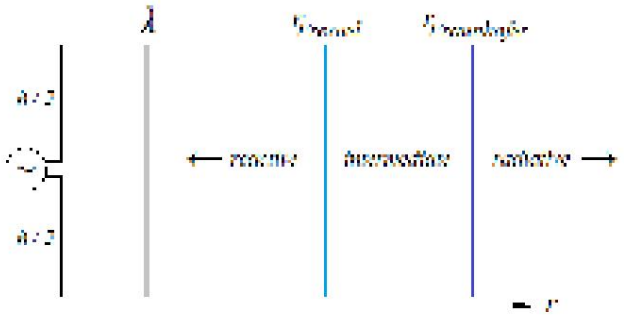


Abbildung 1. Schematische Darstellung der Strahlungszonen um einen symmetrischen Dipol.

Die nachfolgenden Zonenradien für das i-te ÿ **können wie folgt definiert werden [37]:**

$$rFresnel = \ddot{y}i \cdot \frac{h}{2} + \frac{\ddot{y}_3}{2} \cdot \frac{\ddot{y}\,h}{2\ddot{y}i} \quad (1)$$

$$rFraunhofer = \ddot{y}i \cdot \frac{2\,Std.^2}{\ddot{y}i} \quad (2)$$

Für die genaue mathematische Formulierung der elektrischen und magnetischen Feldkomponenten in Elevation ÿ, Orientierung ÿ **und radialer** Richtung r **von der Antenne aus, ist der Grenzwert** ÿ

Der Abstand zwischen reaktiver und strahlender Zone rFraunhofer_approx. kann **wie folgt definiert werden [38]:**

$$rFraunhofer_ca. = \frac{2\,4\ddot{y}}{\ddot{y}} \cdot \frac{\text{und } 2ne}{2\,Monate} \quad (3)$$

wobei ÿ0 [m] die Wellenlänge im freien Raum, e **die elektrische Ladung der Elektronen (1,602 × 10ÿ19 C),** ne **[1/m3] die Elektronenkonzentration/-dichte,** me **die Elektronenmasse (9,109 × 10ÿ31 kg), ÿ0 die elektrische Permittivität im freien Raum (8,854 × 10ÿ12 F/m) und** c **die Lichtgeschwindigkeit im freien Raum (300.000.000 m/s) ist. Für diese Formulierung liegt die nahe Strahlungszone (reaktiv) bei Abständen** r **<<** rFraunhofer_approx.; **für die Zwischenzone:** r **~** rFraunhofer_approx. **und für die Strahlungszone:** r **>>** rFraunhofer_approx..

Das i-te ÿ **kann wie folgt berechnet werden:**

$$\ddot{y}i = \frac{2\ddot{y}}{k_2} \quad (4)$$

wobei ki **die Wellenzahl für einen bestimmten Ausbreitungsmodus bezeichnet. In diesem Artikel werden drei grundlegende Ausbreitungsmodi für niedrigere Frequenzen betrachtet – der Whistler-Modus** (kW), **d. h. für die Frequenzen unterhalb der lokalen Elektronenplasmafrequenz, die als 1400,564 kHz angenähert wird [35], der Alfvén- Modus** (kA), **d. h. für die Frequenzen unterhalb der lokalen gyromagnetischen Ionenfrequenz, die als 40,215 Hz angenähert wird [35], und die Wellenzahl** kJ, **berechnet für die Ausbreitungsbedingungen, die durch die ionosphärische Grenzfrequenz (mit dem Index „klassisch“ gekennzeichnet) für kleine Stromelementantennen** bestimmt werden **[38].**

Machine Translated by Google

Für den Whistler-Modus, wie er von Bannister et al. [35] angegeben wurde, ist die Wellenzahl gleich

$$\mathbf{1}\ \text{kW} = \frac{}{\text{C}}\ \frac{2ÿ\ \text{f en}}{B0ÿ0} \qquad \textbf{(5)}$$

wobei f **[Hz] die Betriebsfrequenz und B0 [T] die magnetische Feldinduktion parallel zur Ausbreitungsrichtung der Welle (senkrecht zur Antennenachse) ist.**

Für den Alfvén- Modus wird nach Bannister et al. [35] die Wellenzahl **wie folgt formuliert :**

$$\text{kA} = \frac{2ÿ\ \text{f}}{\text{C}}\ \frac{\text{nimi}}{ÿ0} \qquad \textbf{(6)}$$

Dabei ist ni **[1/m3] die Ionenkonzentration, gleich** ne , **und** mi **[kg] ist die durchschnittliche Ionenmasse, angenähert als 3,17 × 10ÿ26 kg [35].**

Für den „klassischen" ionosphärischen Fall (in dem die Wellenfront in die Ionosphäre eintritt und Absorption, Brechung und Reflexion erfährt, die ausschließlich von der Frequenz und ne abhängen) wird die Wellenzahl von DJ Bem **[38] wie folgt angegeben :**

$$\text{kJ} = \frac{4ÿ2\ ÿ}{2\ 0}\ \frac{\text{e 2ne}}{meÿ0c^{2}} \qquad \textbf{(7)}$$

analog zur Herleitung der Formel (3).

Die vier Radien der nachfolgenden Strahlungszonen – induktiv (für Antennenlänge h **[m] << ÿ0), Fresnel, Fraunhofer und Fraunhofer klassisch – wurden für verschiedene exemplarische Elektronenkonzentrationen berechnet: ionosphärische F-Schicht unter Tageslichtbedingungen (1,4 × 1012 mÿ3 [38]), ionosphärische F-Schicht unter Nachtbedingungen (4 × 1011 mÿ3 [38]) und heliosphärische Bedingungen, verallgemeinert als 108 mÿ3 [39].**

Die drei betrachteten Wellenzahlen wurden in den Abbildungen 2–4 dargestellt. Jede Elektronenkonzentration wurde für Berechnungen für drei verschiedene Antennenlängen h verwendet: 1 m, 10 m und 100 m (Antenne als symmetrischer Dipol behandelt). Es fällt auf, dass die Größenordnungen der Wellenzahlen für die Alfvén- Fälle erheblich von denen der Whistler- und klassisch-ionosphärischen Fälle abweichen – dies wird später dazu führen, dass einige der Alfvén- Fälle aufgrund ihrer gigantischen Größe (im Vergleich zur tatsächlichen Antennenöffnung) vom direkten Vergleich mit anderen Fällen ausgeschlossen werden. Für den heliosphärische Fall fällt auch auf, dass die klassisch-ionosphärische Funktion von k **(Formel (7)) beginnt, sich mit dem Bereich für den Whistler-Modus der** k- **Funktion (Formel (5)) zu überschneiden – es ist daher wichtig, den Frequenzbereich zu berücksichtigen, auf den die gegebene Funktion anwendbar ist, und die endgültigen Größen der Radien der resultierenden Strahlungszonen, wobei dem kleineren Radius Vorrang eingeräumt wird.**

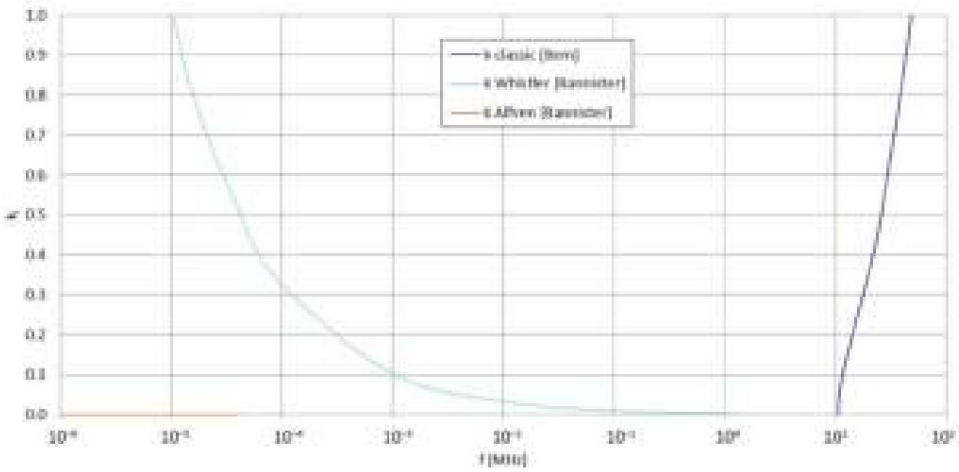


Abbildung 2. Berechnete Werte für die Wellenzahlen k **für den klassischen (ionosphärischen) Fall, den Whistler-Ausbreitungsmodus und den Alfvén- Ausbreitungsmodus für die ionosphärische F-Schicht-Tageslichtelektronendichte (1,4 ×**

Machine Translated by Google

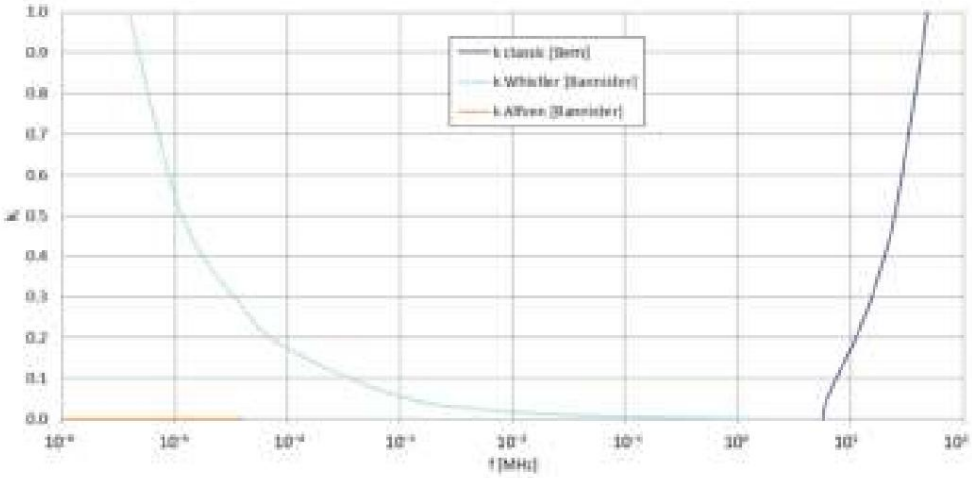


Abbildung 3. Berechnete Werte der Wellenzahlen k **für den klassischen (ionosphärischen) Fall, den Whistler-Ausbreitungsmodus und den Alfvén- Ausbreitungsmodus für die nächtliche Elektronendichte in der ionosphärischen F-Schicht**

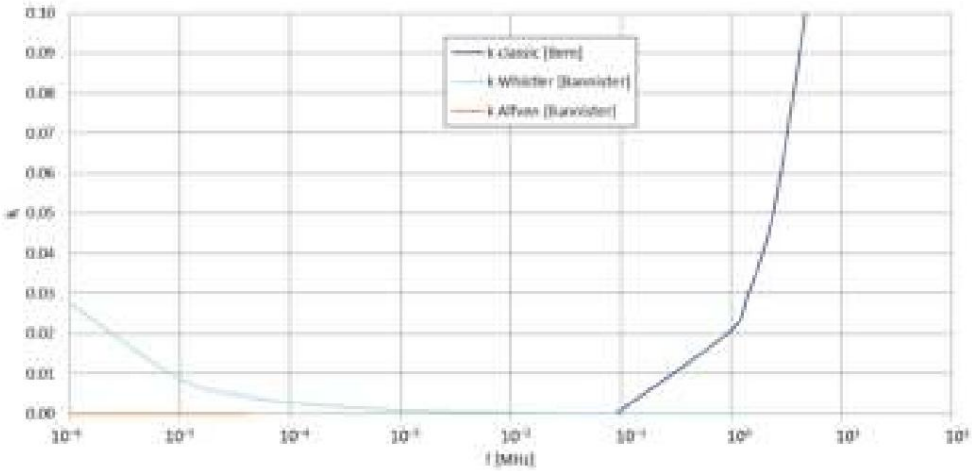


Abbildung 4. Berechnete Werte der Wellenzahlen k **für den klassischen (ionosphärischen) Fall, den Whistler-Ausbreitungsmodus und den Alfvén -Ausbreitungsmodus für die ungefähre heliosphärische Elektronendichte (108 mÿ3).**

Die Abbildungen 5–7 zeigen die Radien der Strahlungszonen – Fraunhofer-Näherung (Formel (3)), Fresnel-Näherung (Formel (1)), Fraunhofer-Näherung (Formel (2)) und Wellenlängen für nicht-plasmamatische Fälle (für ÿ **<< h) – für** kJ **und** h **= 1 m; Abbildungen 8–10: für** kJ **und** h **= 10 m und Abbildungen 11–13: für** kJ **und** h **= 100 m. Es ist zu erkennen, dass sich die Fraunhofer-Zone aus Formel (2) mit zunehmenden Abmessungen der Antenne deutlich ausdehnt, während die ungefähre Angabe der Fraunhofer-Zone (z. B. Angabe des Vorhandenseins – oder Nichtvorhandenseins – bestimmter Feldkomponenten) stärker zum induktiven Bereich (Wellenlänge) konvergiert, wobei die Fresnel-Zone in diesen Größenordnungen die größte ist. Diese Sätze von Funktionen und Diagrammen gelten im Allgemeine**

Abbildung 14 zeigt die ungefähren Radien (3) der Fraunhofer-Strahlungszonen für kW **und alle berücksichtigten** ne- **Bereiche (Formel h-unabhängig); Abbildungen 15–17:** für die **Radien (1) der Fresnel-Zonen** für kW **und alle berücksichtigten** ne- **Bereiche und** h- **Werte; Abbildungen 18–20: für die Radien (2) der Fraunhofer-Zonen für** kW **und alle berücksichtigten** ne- **Bereiche und** h- **Werte; Abbildungen 21–23: für die Radien (1) der Fresnel-Zonen für** kA **und alle berücksichtigten** ne- **Bereiche und** h- **Werte (die Radien der Fraunhofer-Zonen für** kA **wurden aufgrund großer Unterschiede in den Größenordnungen ihrer Werte im Vergleich zu den Werten der Antennenapertur h we**

Machine Translated by Google

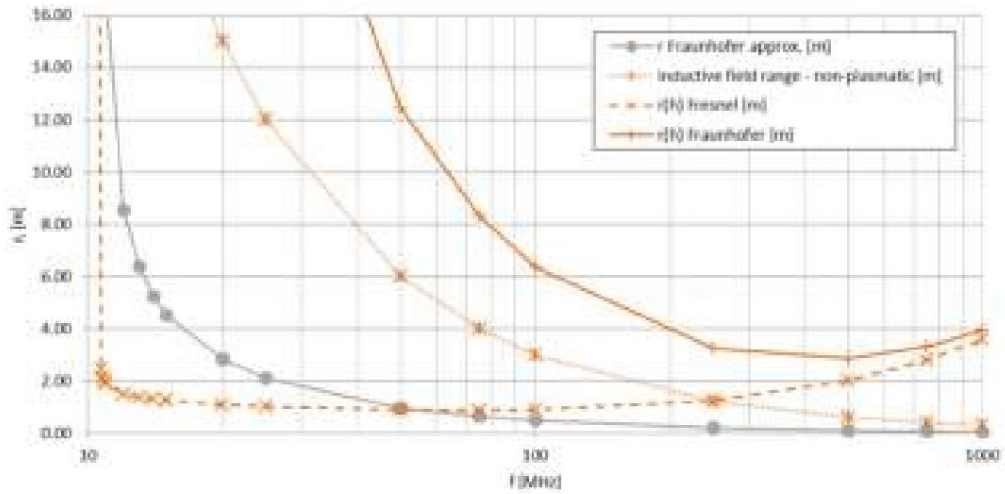


Abbildung 5. Berechnete Radien der Strahlungszonen für den klassisch-ionosphärischen Ansatz (kJ) **für die ionosphärische Tageslichtelektronendichte in der F-Schicht (1,4 × 1012 mÿ3) und** h **= 1 m.**

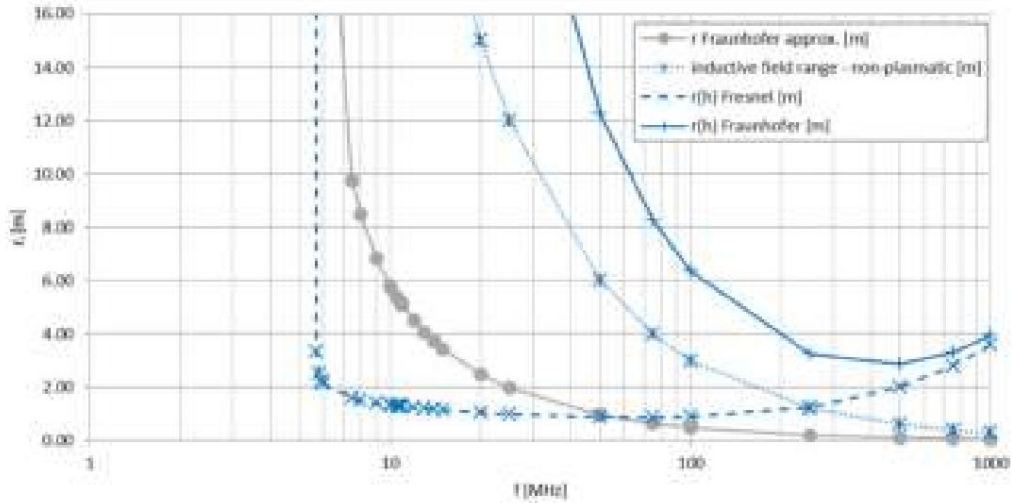


Abbildung 6. Berechnete Radien der Strahlungszonen für den klassisch-ionosphärischen Ansatz (kJ) **für die nächtliche Elektronendichte in der ionosphärischen F-Schicht (4 × 1011 mÿ3) und** h **= 1 m.**

tt
tt

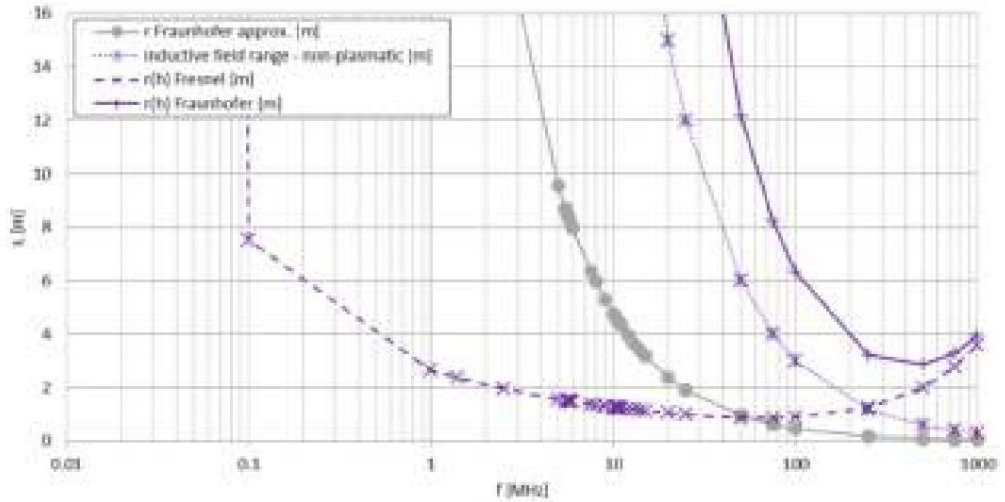


Abbildung 7. Berechnete Radien der Strahlungszonen für den klassisch-ionosphärischen Ansatz (kJ) **für die ungefähre heliosphärische Elektronendichte (108 mÿ3) und** h **= 1 m.**

Machine Translated by Google

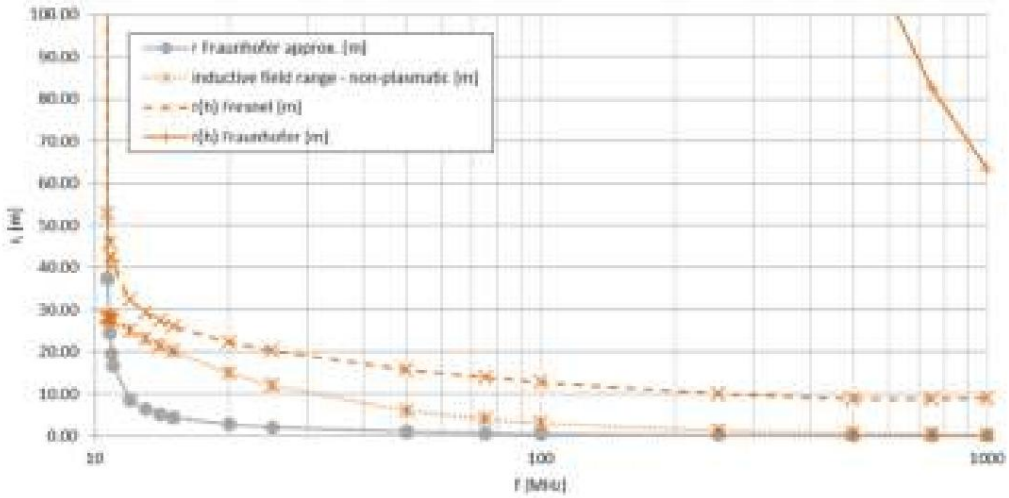


Abbildung 8. Berechnete Radien der Strahlungszonen für den klassisch-ionosphärischen Ansatz (kJ) **für die Tageslichtelektronendichte in der ionosphärischen F-Schicht (1,4 × 1012 mÿ3) und** h **= 10 m.**

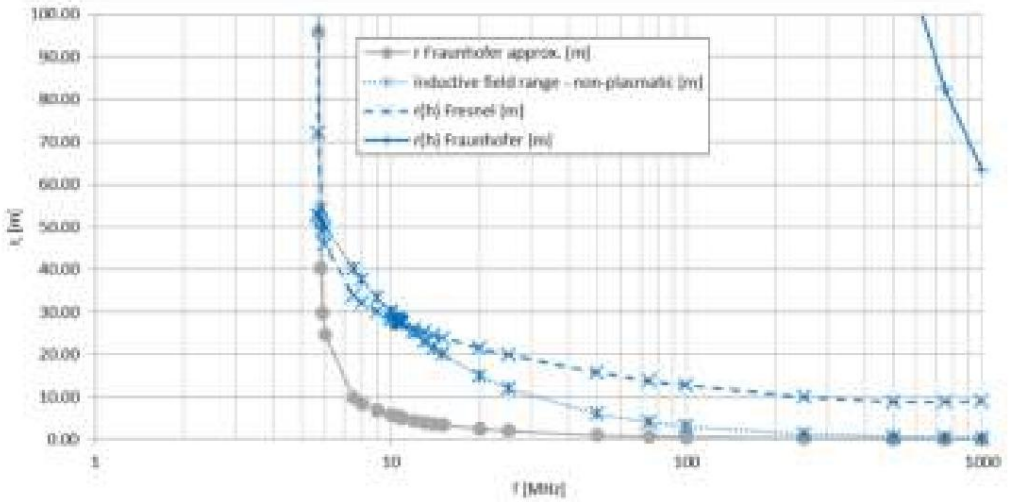


Abbildung 9. Berechnete Radien der Strahlungszonen für den klassisch-ionosphärischen Ansatz (kJ) **für die nächtliche Elektronendichte in der ionosphärischen F-Schicht (4 × 1011 mÿ3) und** h **= 10 m.**

tt tt

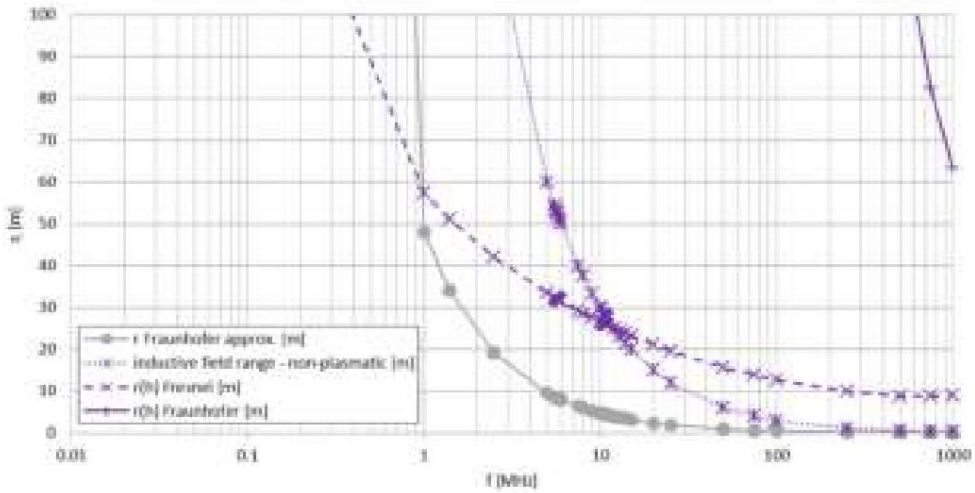


Abbildung 10. Berechnete Radien der Strahlungszonen für den klassisch-ionosphärischen Ansatz (kJ) **für die ungefähre heliosphärische Elektronendichte (108 mÿ3) und** h **= 10 m.**

Machine Translated by Google

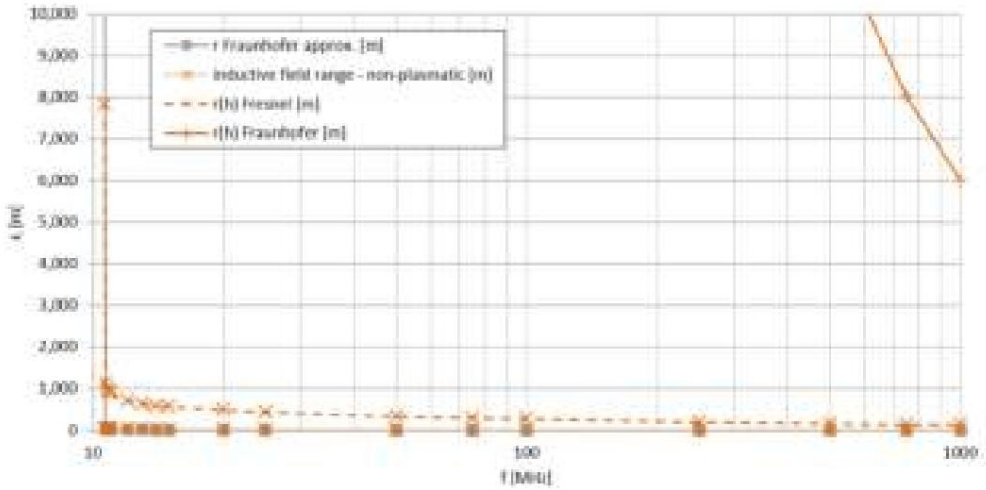


Abbildung 11. Berechnete Radien der Strahlungszonen für den klassisch-ionosphärischen Ansatz (kJ) **für die ionosphärische F-Schicht-Tageslichtelektronendichte (1,4 × 1012 mÿ3) und** h **= 100 m.**

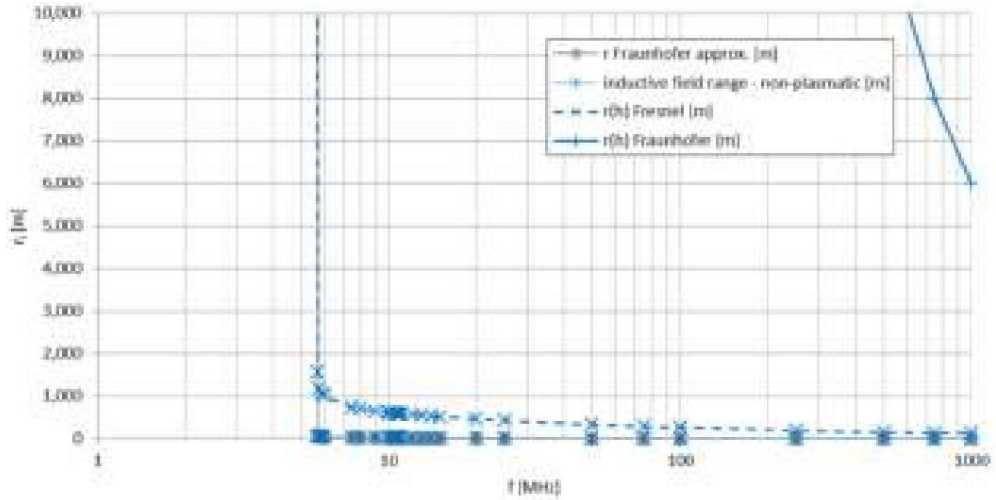


Abbildung 12. Berechnete Radien der Strahlungszonen für den klassisch-ionosphärischen Ansatz (kJ) **für die nächtliche Elektronendichte in der ionosphärischen F-Schicht (4 × 1011 mÿ3) und** h **= 100 m.**

tt
tt
tt

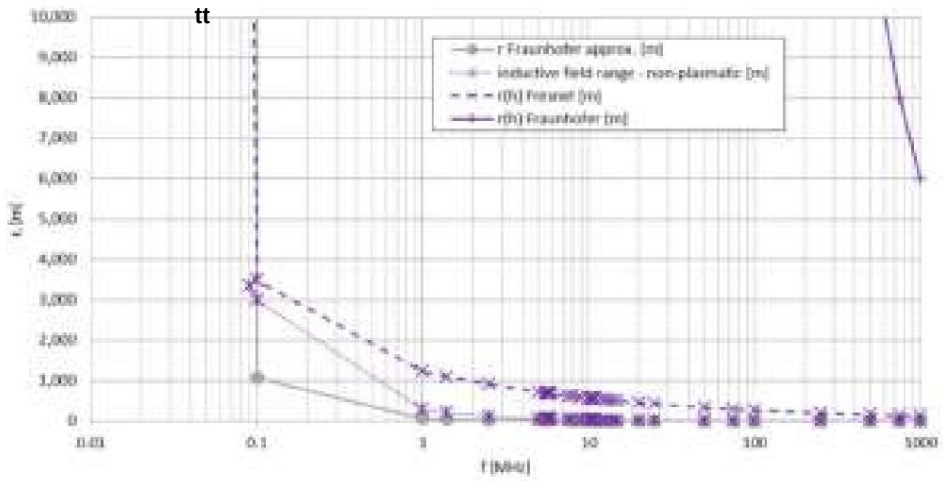


Abbildung 13. Berechnete Radien der Strahlungszonen für den klassisch-ionosphärischen Ansatz (kJ) **für die ungefähre heliosphärische Elektronendichte (108 mÿ3) und** h **= 100 m.**

Machine Translated by Google

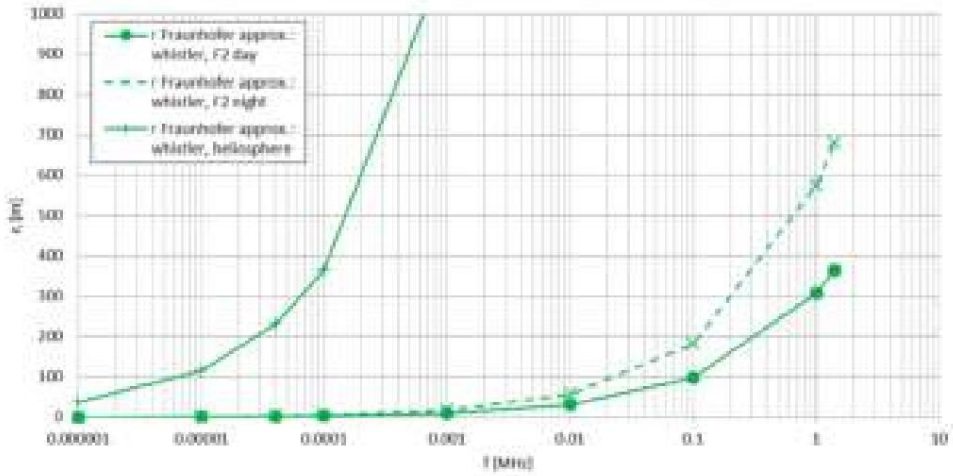


Abbildung 14. Berechnete ungefähre Radien der Fraunhofer-Strahlungszonen für die Whistler-Ausbreitung im ff ff ff-Modus (kW) **für** h **= 1 m und unterschiedliche Elektronenkonzentrationen.**

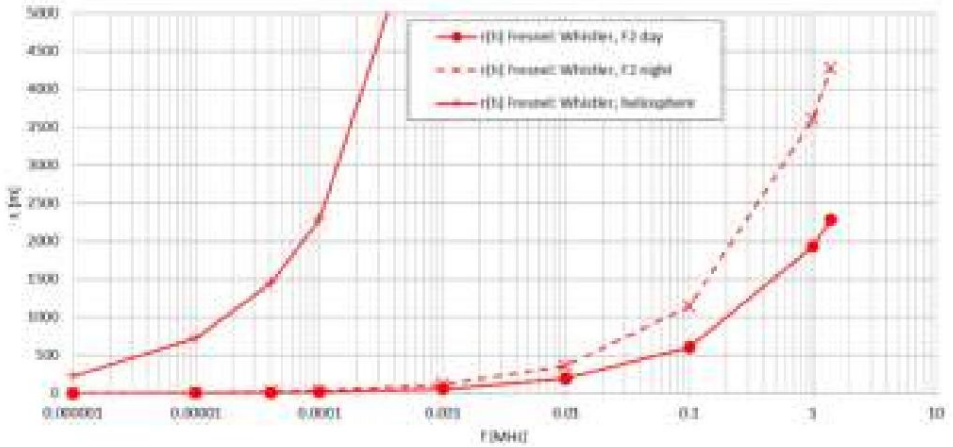


Abbildung 15. Berechnete Radien der Fresnel-Strahlungszonen für den Whistler-Ausbreitungsmodus (kW) **für ff ff** h **= 1 und verschiedene Elektronenkonzentrationen. ff**

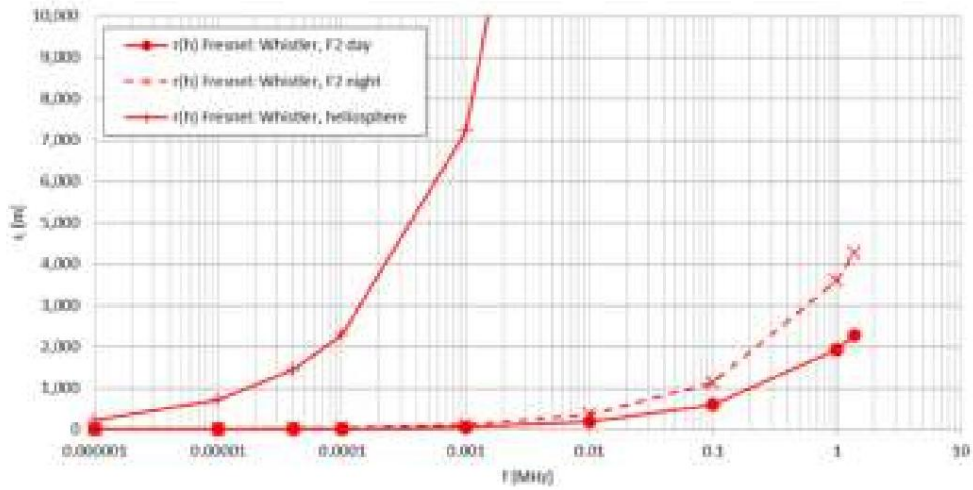
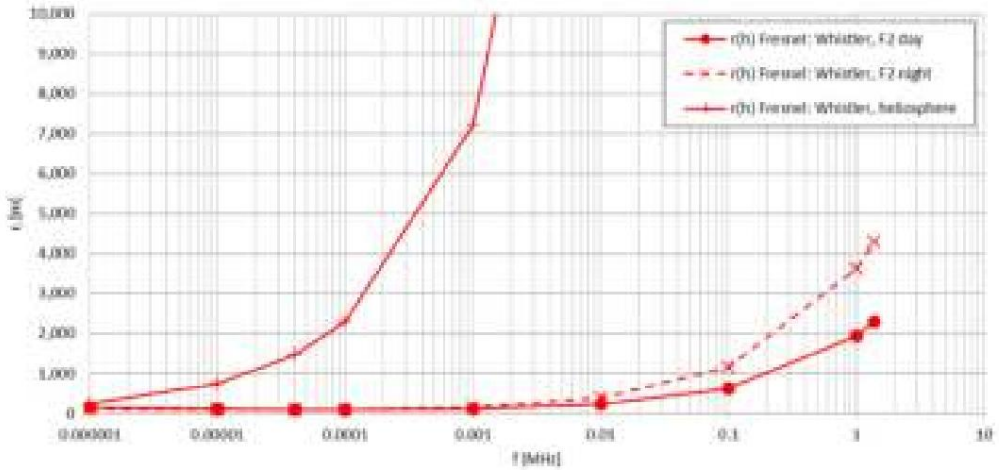


Abbildung 16. Berechnete Radien der Fresnel-Strahlungszonen für den Whistler-Ausbreitungsmodus (kW) **für** h **= 10 m und verschiedene Elektronenkonzentrationen. ff ff ff**

Machine Translated by Google



ff

Abbildung 17. Berechnete Radien der Fresnel-Strahlungszonen für den Whistler-Ausbreitungsmodus (kW) **für ff** h **= 100 m und verschiedene Elektronenkonzentrationen. ff**

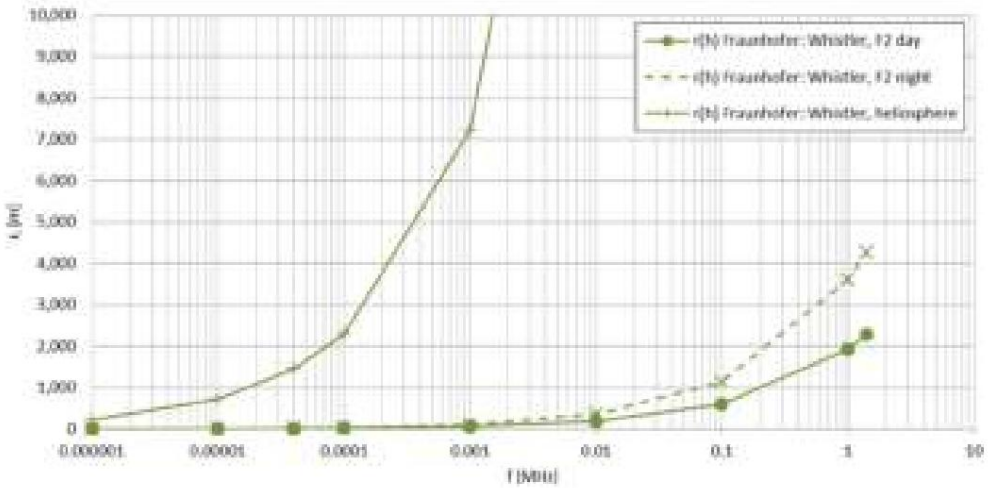


Abbildung 18. Berechnete Radien der Fraunhofer-Strahlungszonen für den Whistler-Ausbreitungsmodus (kW) **für** h **= 1 m und unterschiedliche Elektronenkonzentrationen.**

ff

ff ff

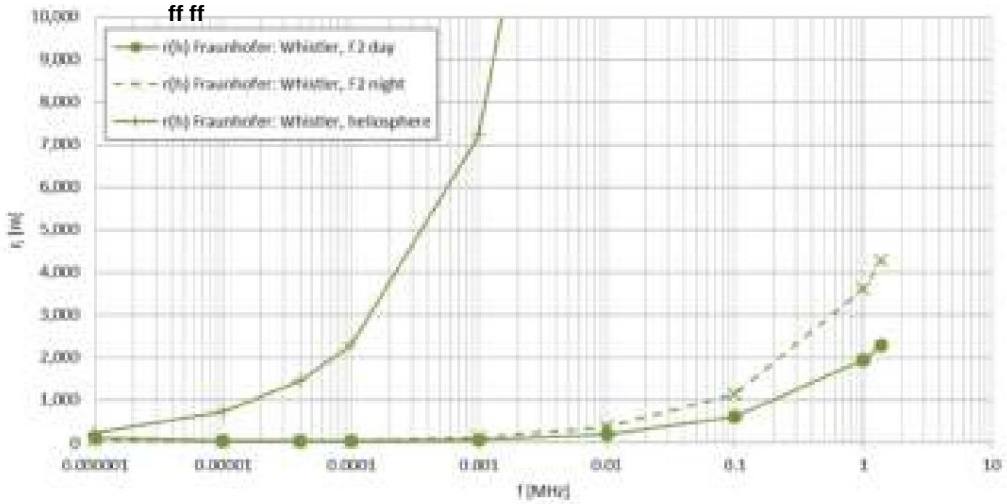


Abbildung 19. Berechnete Radien der Fraunhofer-Strahlungszonen für den Whistler-Ausbreitungsmodus (kW) **für** h **= 10 m und unterschiedliche Elektronenkonzentrationen.**

ff

ff ff

Machine Translated by Google

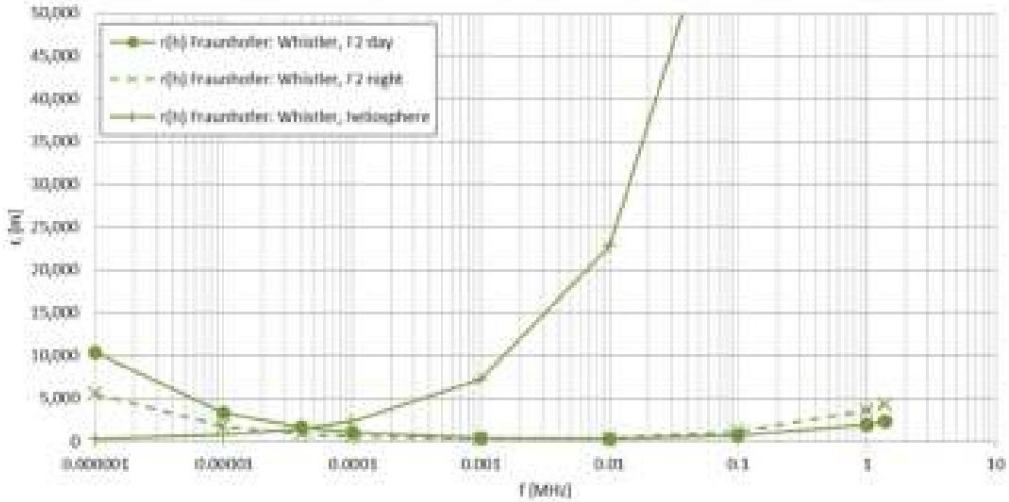


Abbildung 20. Berechnete Radien der Fraunhofer-Strahlungszonen für den Whistler-Ausbreitungsmodus (kW) **für** h **= 100 m und unterschiedliche Elektronenkonzentrationen.**

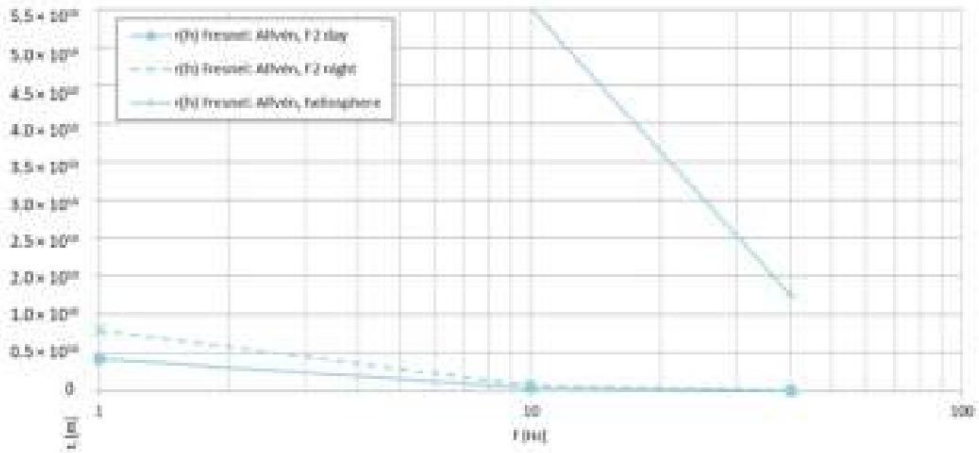


Abbildung 21. Berechnete Radien der Fraunhofer-Strahlungszonen für den Alfvén- Ausbreitungsmodus (kA) **für** h **= 1 m und unterschiedliche Elektronenkonzentrationen.**

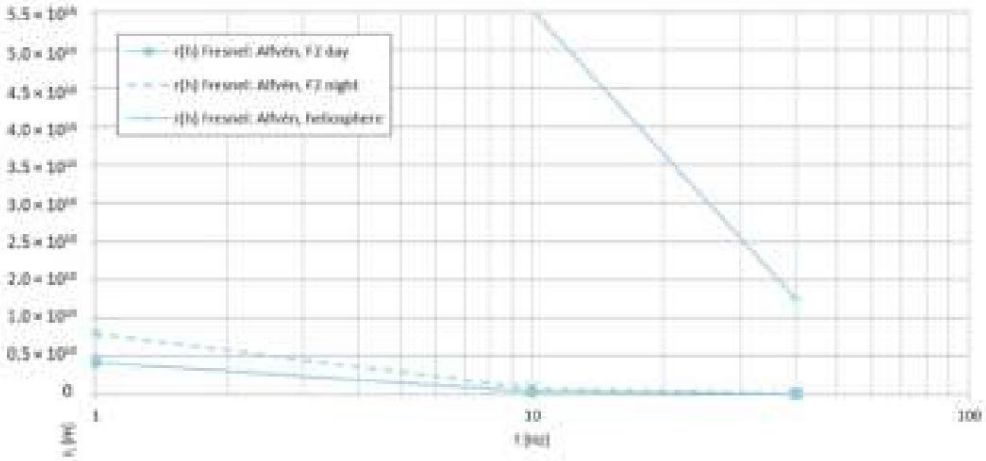


Abbildung 22. Berechnete Radien der Fraunhofer-Strahlungszonen für den Alfvén- Ausbreitungsmodus (kA) **für** h **= 10 m und unterschiedliche Elektronenkonzentrationen.**

Machine Translated by Google

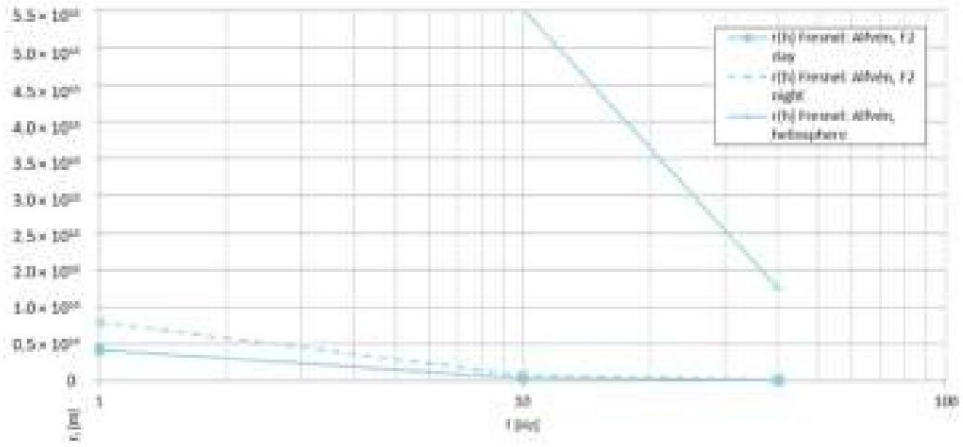


Abbildung 23. Berechnete Radien der Fraunhofer-Strahlungszonen für den Alfvén- Ausbreitungsmodus (kA) **für** h **= 100 m und unterschiedliche Elektronenkonzentrationen.**

Es fällt auf, dass die ungefähren Radien der Fraunhofer-Zone (3) für heliosphärische Bedingungen für eine bestimmte Frequenz um viele Größenordnungen höher bleiben als für die anderen elektronenreichen Umgebungen. Die Radien der Fraunhofer- und Fresnel-Zonen, die mit präzisen Formeln berechnet wurden, zeigen hauptsächlich eine Abhängigkeit von der Antennenlänge und ihrer Größenordnung, während die allgemeine (und dem vorherigen Fall ähnliche) Beziehung zwischen heliosphärischem und elektronenreichem Fall erhalten bleibt. Der Fall der Fraunhofer-Zone für die größte (100 m) Antennenlänge beginnt bei den niedrigeren Frequenzen (~kHz und niedriger) Änderungen/Radiuslängenzunahmen aufzuweisen , da die Antennenlä

3. Entwicklung der Scalloping-Funktion des Antennenstrahlungsmusters

Die Ausbuchtung des Strahlungsmusters (und nach dem Reziprozitätssatz [40] auch des Empfangsmusters) = ÿ ÿ der Antenne kann als längen- oder frequenzabhängiges Auftreten und Verlagerung von Minima und Maxima beschrieben werden, wodurch die von der Antenne gesendete Leistung in bestimmte Richtungen umgeleitet wird oder die Antenne für den Signalempfang in bestimmten Winkeln empfänglicher wird. Wenn die exakte Lösung [36], die die Anzahl der Maxima des Musters – also der Lappen – beschreibt, ähnlich wie (3) und (7) an elektronenreiche Bedingungen angepasst wird, kann die plasmatische Lösung für die Anzahl nPlasma **[-] der Ausbuchtungen des Musters wie folgt definiert werden:**

$$nPlasma = \ddot{y}_H \sqrt{\frac{4}{\ddot{y}2} \cdot \frac{e\ 2ne}{me\ddot{y}0\ddot{y}2c\ 2}} \ddot{y} \quad (8)$$

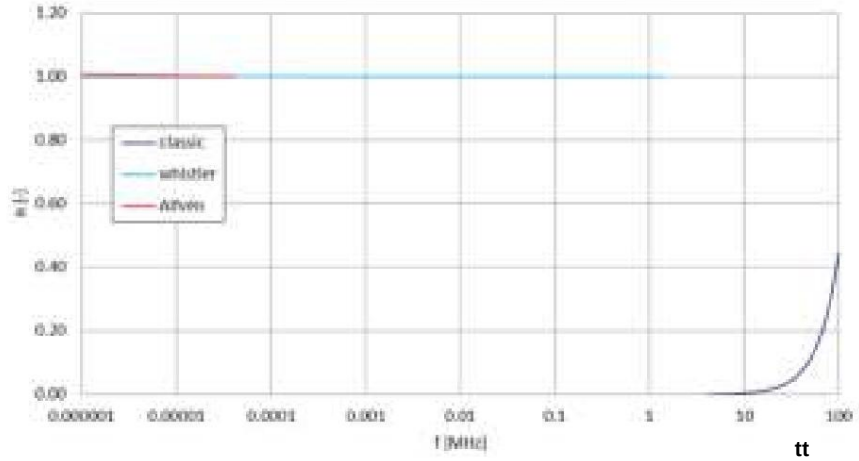
, wobei die Klammern die Obergrenze der Formel definieren (d. h. ein Wert über 1 ergibt nPlasma **= 2, ein Wert über 2 ergibt** nPlasma **= 3 usw.).**

Die Abbildungen 24 und 25 zeigen die Anzahl der Lappen (Werte der Formel in Klammern) für die heliosphärische Umgebung (ne **= 108 mÿ3) für alle berücksichtigten Ausbreitungsfälle (klassischer Ansatz, Whistler und Alfvén) und alle berücksichtigten Antennenlängen** h **(1, 10 und 100 m); Abbildungen 26–28: für die ionosphärische F-Schicht bei Tageslichtbedingungen** (ne **= 1,4 × 1012 mÿ3) und Abbildungen 29–32: ionosphärische F-Schicht bei Nachtbedingungen** (ne **= 4 × 1011 mÿ3).**

Es ist zu bemerken, dass bei heliosphärischen Bedingungen der MHz-Bereich der Signalfrequenzen einen exponentiellen Anstieg der Anzahl der Lappen des Musters erfährt, der mit der Verlängerung der Antenne deutlich zunimmt; bei anderen (niedrigeren) Frequenzen bleibt die Form des Musters nahezu konstant. Bei elektronendichteren Bedingungen (ionosphärisches F2 tagsüber) beginnen niedrigere Frequenzen einen starken exponentiellen Anstieg der Anzahl der Lappen des Musters zu zeigen, mit Ausnahme des Alfvén- Falls, der nahezu konstant bleibt. Ein ähnlicher Mechanismus, jedoch mit niedrigeren erreichten Werten, ist bei ionosphärischen F2- Elektronenkonzentrationsbedingungen nachts zu beobachten; er kann mit d

Machine Translated by Google

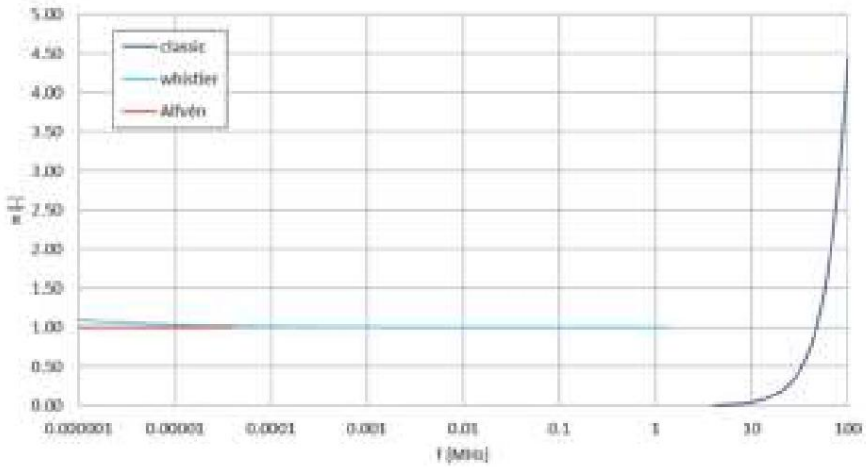
proportional zur Wellenlänge des Signals bei den niedrigeren Frequenzen und wird bei den höheren Frequenzen schnell überdimensioniert für die Wellenlänge des Signals – der erste Mechanismus kann als „verstärkt“ durch die erhöhte Elektronendichte wahrgenommen werden, und der letztere Mechanismus scheint global vorhanden zu sein.



tt

Abbildung 24. Berechnete Anzahl der Lappen des Strahlungs-/Empfangsmusters der Antenne für eine Antennenlänge von h **= 1 m und eine heliosphärische Umgebung** (ne **= 108 mÿ3).**

tt tt



tt

Abbildung 25. Berechnete Anzahl der Lappen des Strahlungs-/Empfangsmusters der Antenne bei einer Antennenlänge von h **= 10 m und einer heliosphärischen Umgebung** (ne **= 108 mÿ3).**

tt tt

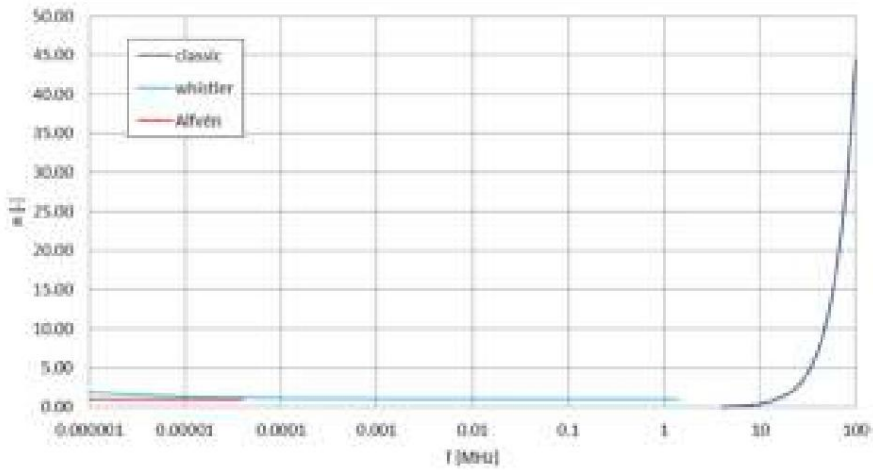


Abbildung 26. Berechnete Anzahl der Lappen des Strahlungs-/Empfangsmusters der Antenne für eine Antennenlänge von h **= 100 m und eine heliosphärische Umgebung** (ne **= 108 mÿ3).**

Machine Translated by Google

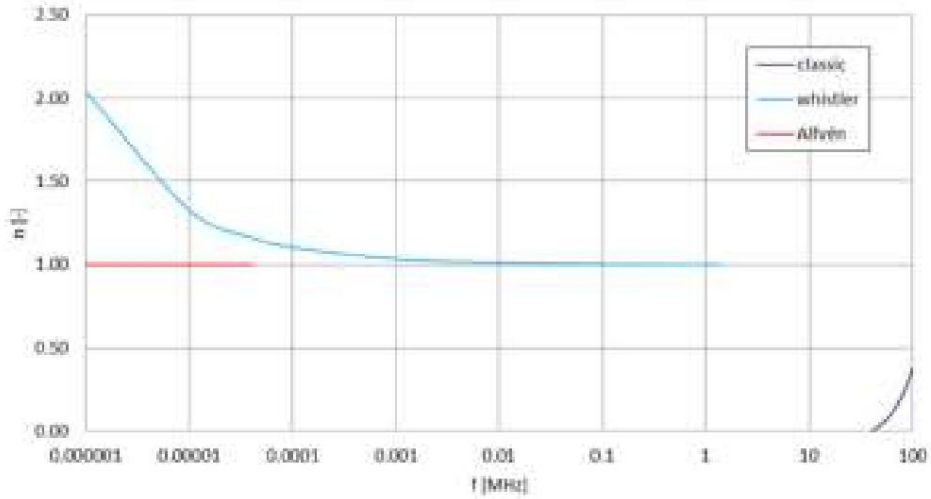


Abbildung 27. Berechnete Anzahl der Lappen des Strahlungs-/Empfangsmusters der Antenne bei einer Antennenlänge von h **= 1 m und einer ionosphärischen F2-Schicht-Tageslichtumgebung** (ne **= 1,4 × 1012 mÿ3).** **tt tt**

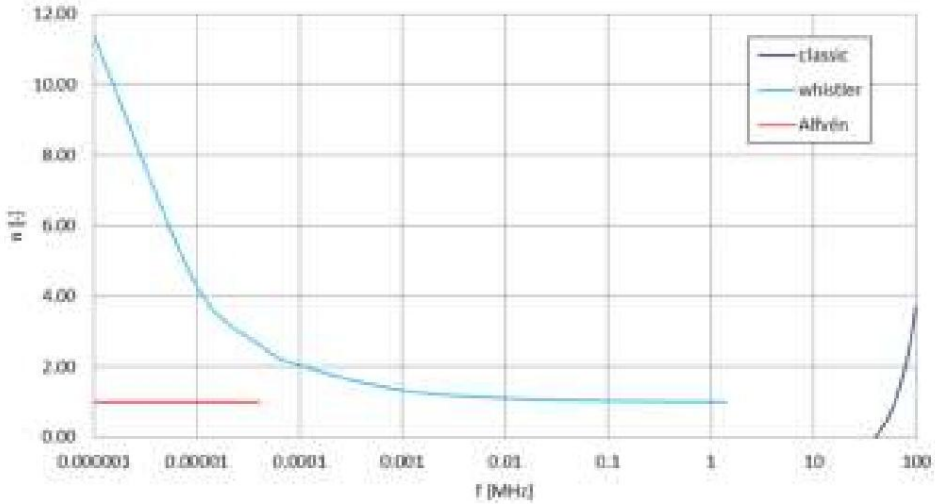


Abbildung 28. Berechnete Anzahl der Lappen des Strahlungs-/Empfangsmusters der Antenne bei einer Antennenlänge von h **= 10 m und einer ionosphärischen F2-Schicht-Tageslichtumgebung (**ne **= 1,4 × 1012 mÿ3).** **tt tt**

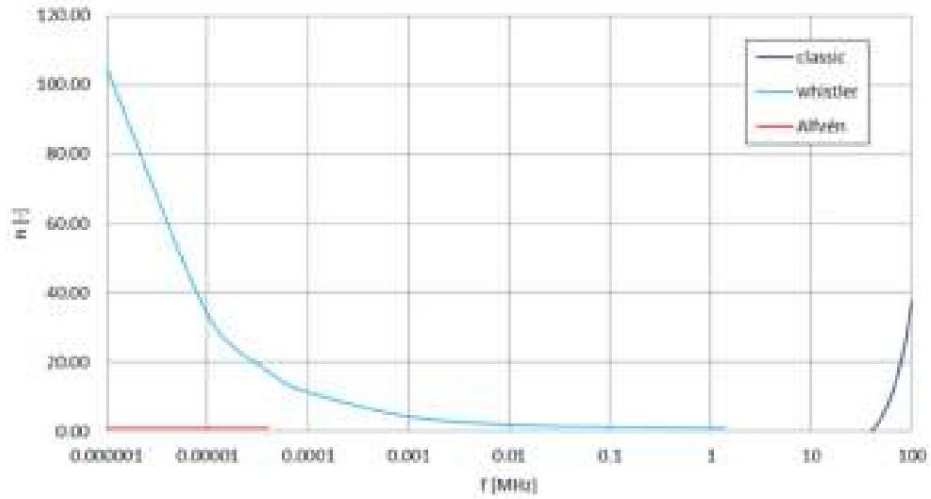


Abbildung 29. Berechnete Anzahl der Lappen des Strahlungs-/Empfangsmusters der Antenne bei einer Antennenlänge von h **= 100 m und einer ionosphärischen F2-Schicht-Tageslichtumgebung (**ne **= 1,4 × 1012 mÿ3).** **tt**

Machine Translated by Google

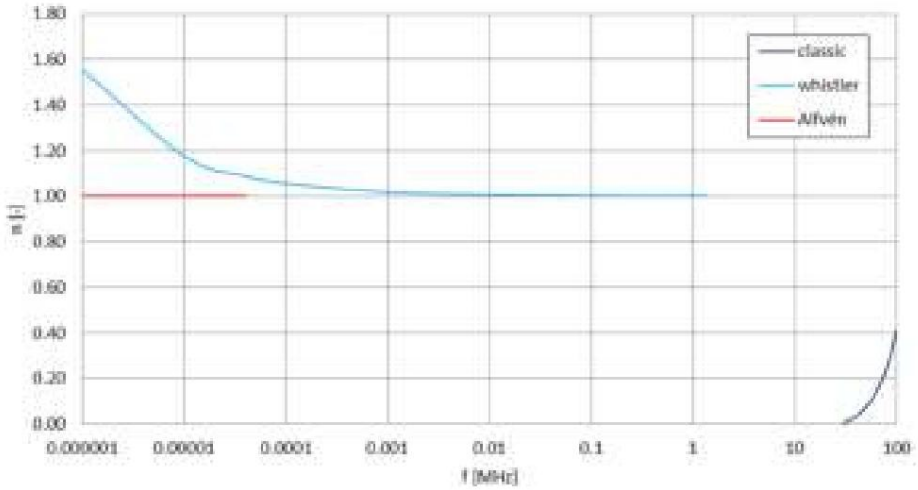


Abbildung 30. Berechnete Anzahl der Lappen des Strahlungs-/Empfangsmusters der Antenne bei einer Antennenlänge von h **= 1 m und einer ionosphärischen F2-Schicht in der Nachtumgebung** (ne **= 4 × 1011 mÿ3).**

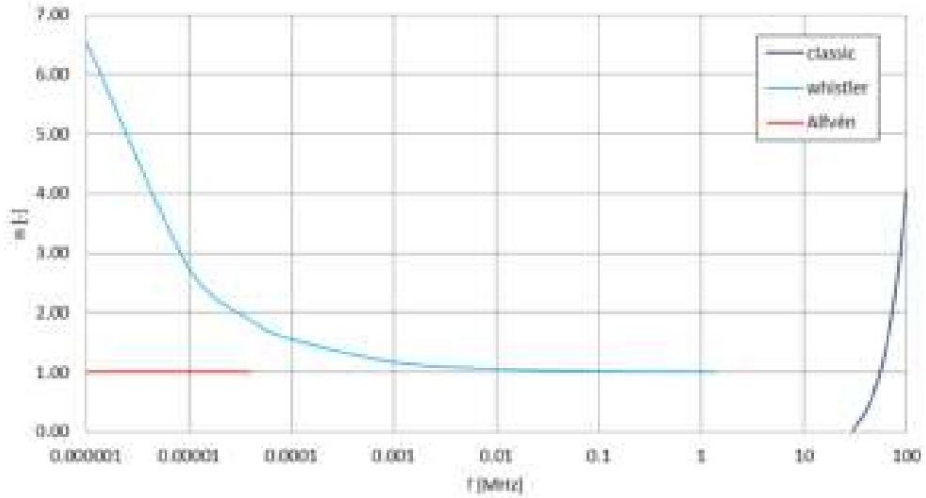


Abbildung 31. Berechnete Anzahl der Lappen des Strahlungs-/Empfangsmusters der Antenne bei einer Antennenlänge von h **= 10 m und einer ionosphärischen F2-Schicht in der Nachtumgebung** (ne **= 4 × 1011 mÿ3).**

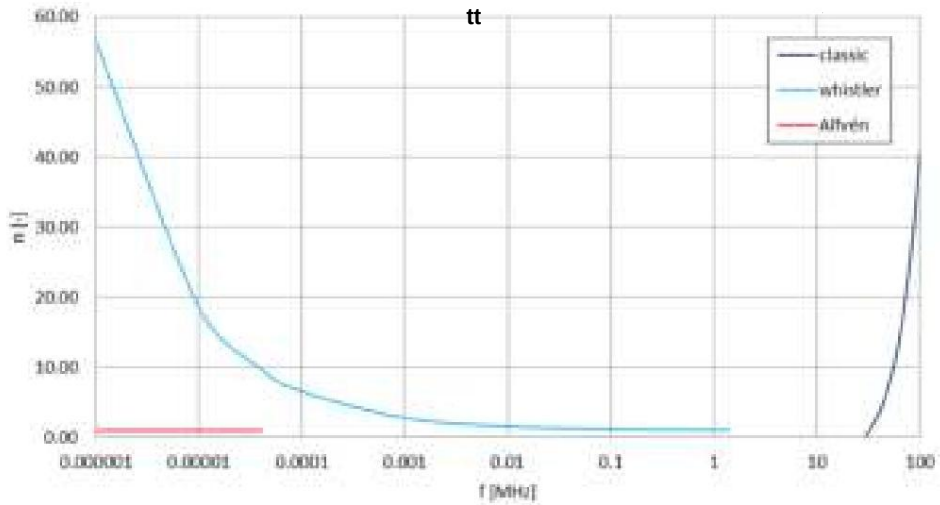


Abbildung 32. Berechnete Anzahl der Lappen des Strahlungs-/Empfangsmusters der Antenne bei einer Antennenlänge von h **= 100 m und einer ionosphärischen F2-Schicht in der Nachtumgebung** (ne **= 4 × 1011 mÿ3).**

Machine Translated by Google

4. Entwicklung der Richtwirkung des Antennenstrahlungsmusters

Der Parameter Richtwirkung der Antenne D **[-] kann als die Fähigkeit der Antenne erklärt werden, einen konzentrierten Strahl abgestrahlter Energie in einem bestimmten Winkel zu bilden – definiert als die Strahlungsintensität in einer gegebenen Richtung/einem gegebenen Winkel geteilt durch die durchschnittliche Strahlungsintensität über das gesamte Strahlungsmuster** *(2ÿ).* **Zusammen mit der Effizienz der Antenne** ÿA **[-] (die vom Strahlungswiderstand der Antenne und ihrem Verlustwiderstand abhängt) ergibt dies den Antennengewinn** G **[-]:** G = *DÿA* **[40]; der Antennengewinn ist daher direkt proportional zur Richtwirkung, wobei die Effizienz der Antenne stark von den Strahlungsverlusten bei der tatsächlichen Anwendung in einem Raumfahrzeug abhängt (der Strahlungswiderstand für einen Dipol als Funktion des Länge-zu-Wellenlänge-Verhältnisses i**

Die Richtwirkungsformel kann wie folgt ausgedrückt werden [36] :

$$Di = 2 \quad \frac{1}{3} \quad \frac{\cos(2hki)\sin(2hki) + }{(2hki)^2} \quad \frac{}{(2hki)^3} \quad ^{ÿ1} \tag{9}$$

wobei ki **für jeden Fall (Ionosphären-, Whistler- oder Alfvén-Modus) unter Verwendung der Formeln (5), (6) und (7) definiert ist. Die Abbildungen 33–35 zeigen die für den klassischen (7) Fall berechnete Richtwirkung für verschiedene zuvor betrachtete Elektronenkonzentrationen und Antennenlängen** h **= 1, 10 und 100 m; analog dazu wird in den Abbildungen 36–38** die Whistler-Wellenzahl (5) und **in den Abbildungen 39–41 die Alfvén- Wellenzahl (6) verwendet .**

In den klassisch-ionosphärischen Fällen ist die Richtwirkung umso konstanter, je länger die Antenne ist, und erreicht bei hohen MHz-Frequenzen einen Wert von 6. Im kHz- und darunterliegenden Frequenzbereich erreicht die Richtwirkung Werte unter 1, was auf das Fehlen gut ausgebildeter separater Lappen im Strahlungsmuster hinweist (wie in den Abbildungen im vorhergehenden Absatz). Beim Whistler- Ausbreitungsmodus verschiebt sich das allgemeine Verhalten der Richtwirkungsänderungen mit den Frequenzwerten nach oben, wenn die Antennenlänge zunimmt. Beide ionosphärischen Bedingungen zeigen zwei Maxima bei etwa 6,5, während der heliosphärische Fall bei der größten betrachteten Antennenlänge von 100 m größere Werte zu liefern beginnt. Beim Alfvén- Ausbreitungsmodus und -Frequenzbereich in Abbildung 40 führten die heliosphärische Bedingungen zu numerischen Fehlern in Formel (9) und wurden in die endgültige Darstellung nicht aufgenommen. je kleiner die Antenne, desto anfälliger scheint sie gegenüber höheren Elektronenkonzentrationen zu sein, wobei die größte betrachtete Antenne von 100 m eine hohe Variabilität gegenüber der niedrigsten (heliosphärischen) Elektronenkonzentration a

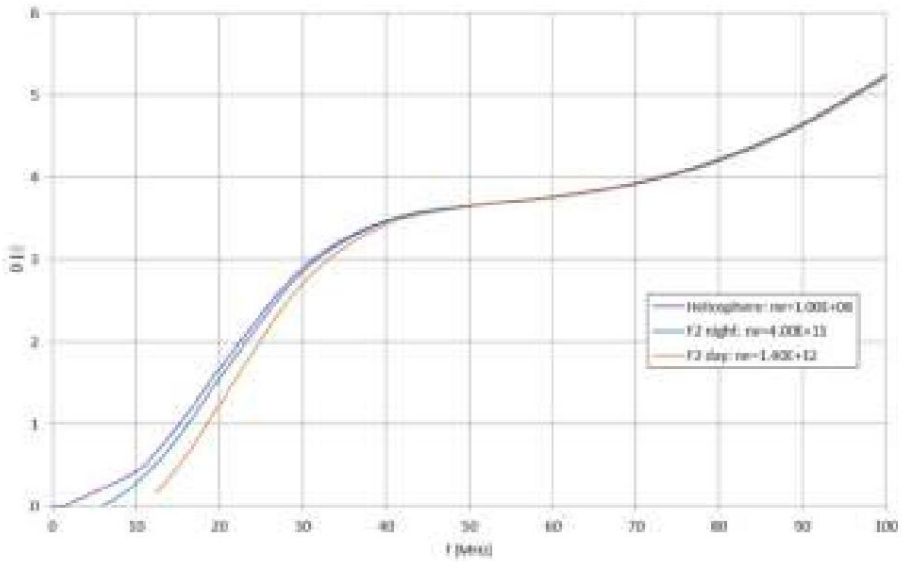


Abbildung 33. Berechnete Antennenrichtwirkung für Antennenlänge h **= 1 m und Wellenzahl** kJ **(7).**

Machine Translated by Google

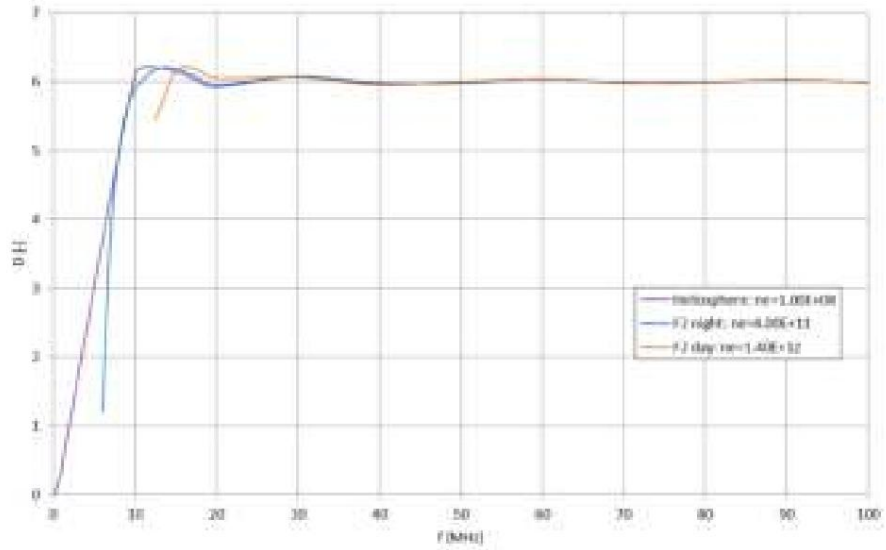


Abbildung 34. Berechnete Antennenrichtwirkung für Antennenlänge h **= 10 m und Wellenzahl** kJ **(7).**

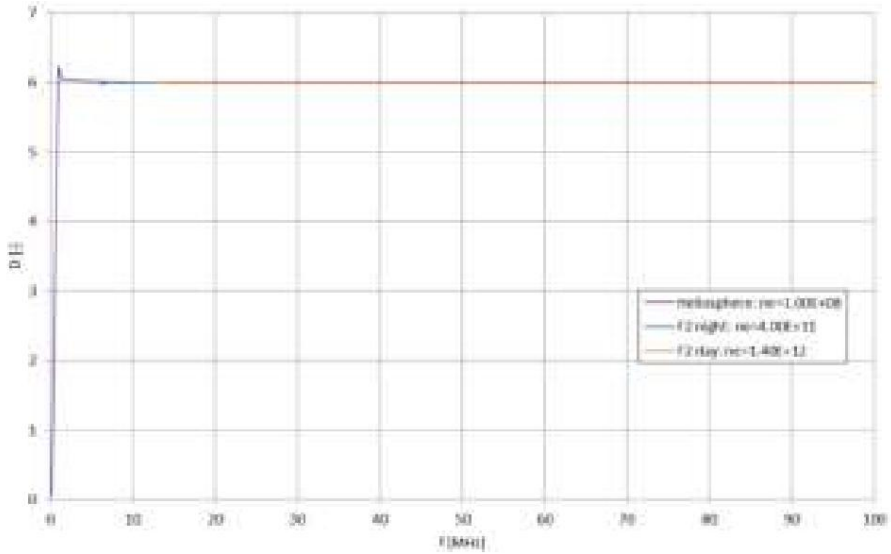


Abbildung 35. Berechnete Antennenrichtwirkung für Antennenlänge h **= 100 m und Wellenzahl** kJ **(7).**

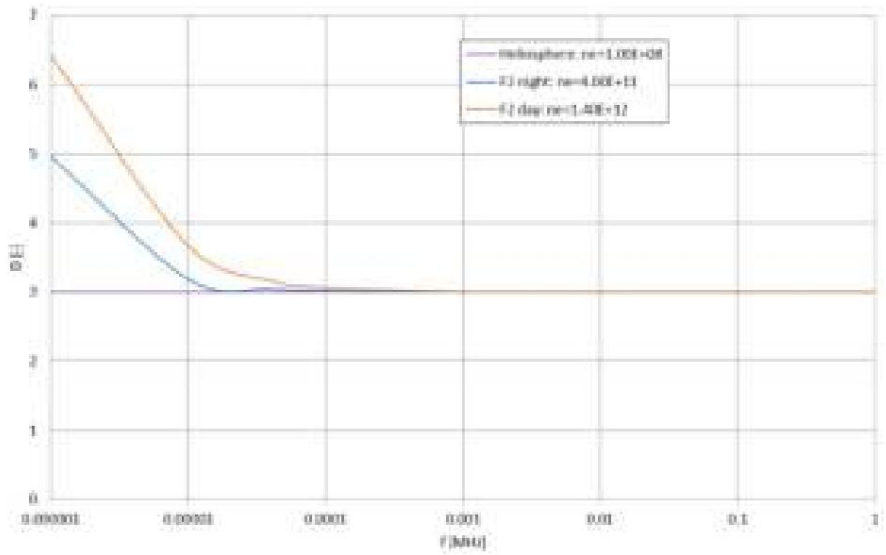


Abbildung 36. Berechnete Antennenrichtwirkung für Antennenlänge h **= 1 m und Wellenzahl** kW **(5).**

Machine Translated by Google

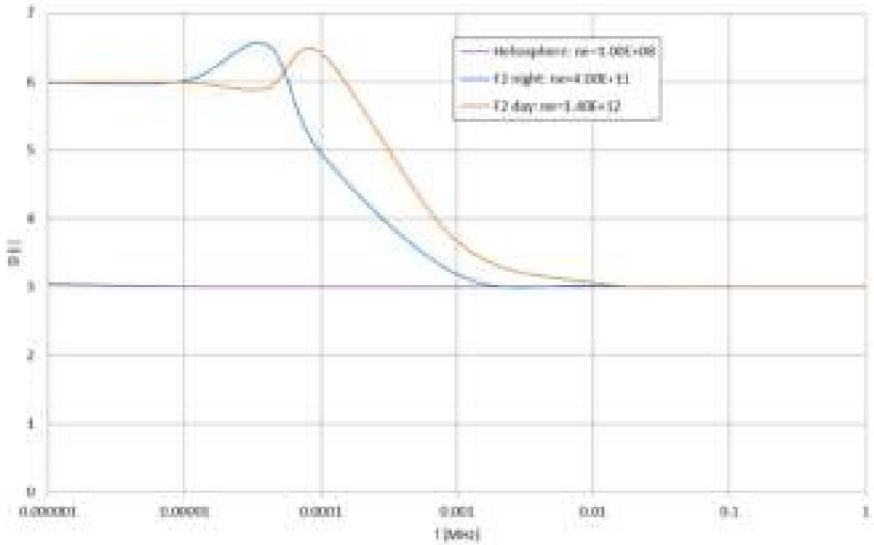


Abbildung 37. Berechnete Antennenrichtwirkung für Antennenlänge h **= 10 m und Wellenzahl** kW **(5).**

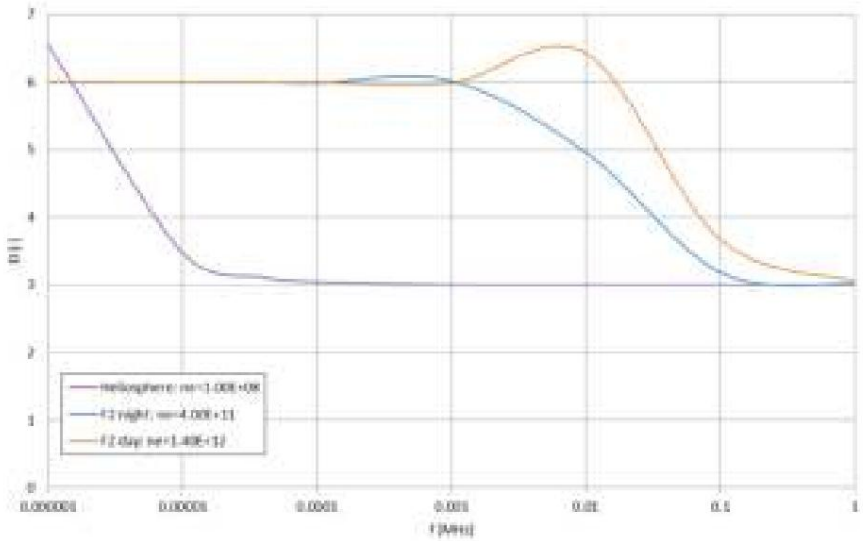


Abbildung 38. Berechnete Antennenrichtwirkung für Antennenlänge h **= 100 m und Wellenzahl** kW **(5).**

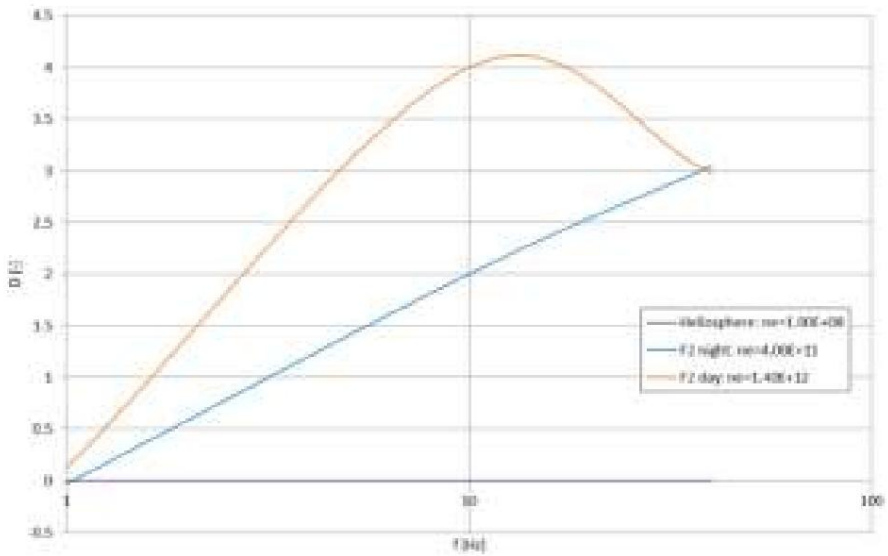


Abbildung 39. Berechnete Antennenrichtwirkung für Antennenlänge h **= 1 m und Wellenzahl** kA **(6).**

Machine Translated by Google

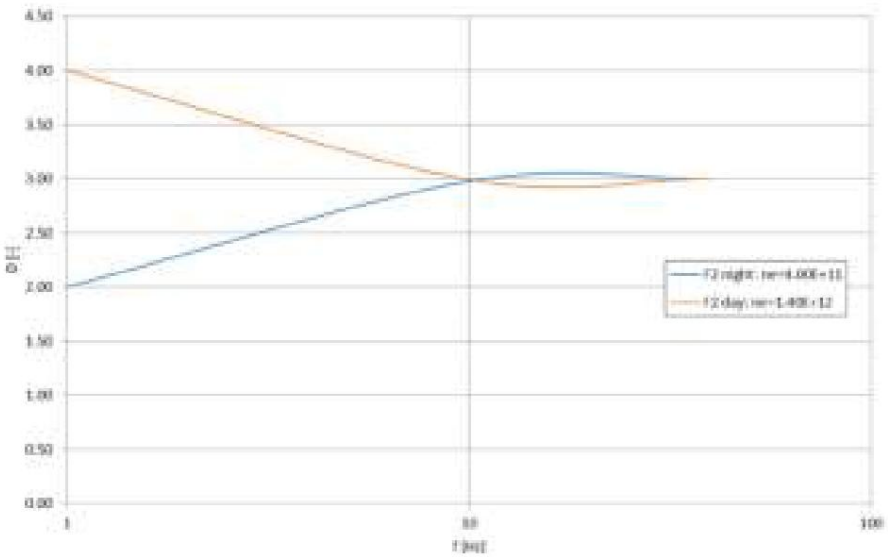


Abbildung 40. Berechnete Antennenrichtwirkung für Antennenlänge h **= 10 m und Wellenzahl** kA **(6).**

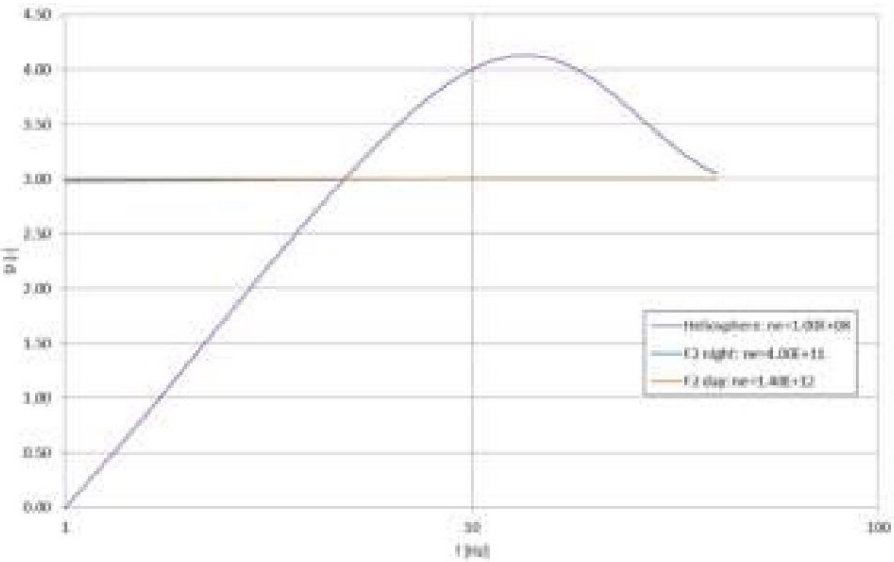


Abbildung 41. Berechnete Antennenrichtwirkung für Antennenlänge h **= 100 m und Wellenzahl** kA **(6).**

5. Strahlungsmuster-Transformationskoeffizient von Luft/Vakuum zu Plasmabedingungen

Die experimentelle Überprüfung der Form des Strahlungs-/Empfangsmusters der Antenne, die immer die tatsächliche mechanische Konfiguration der Antenne und der Antennenhalterung (z. B. auf einem Raumfahrzeug) einbezieht, ist hauptsächlich entweder für terrestrische atmosphärische Bedingungen oder nahezu Vakuumbedingungen möglich (wenn zusätzliche Wärmeübertragung, wie sie für die Weltraumumgebung typisch ist, für andere Komponenten nachgebildet werden muss). Es wurden analytische Lösungen für das Strahlungs-/Empfangsmuster einer gegebenen Antenne entwickelt, die in das Plasma mit gegebenen Parametern eingetaucht ist – z. B. das Strahlungsmuster von Bannister et al. für die i-te Wellenzahl k **[35]:**

$$\text{sin } \ddot{y}\ 1\ fP(\ddot{y}) = kih(1 \quad \ddot{y}\ \cos \ddot{y})\ \text{sin}\ (\ ) = \text{sin}\ \ddot{y}(1\ \ddot{y}\ )\ 1\ \ddot{y}\ \cos \ddot{y}\ 2 \tag{10}$$

wobei $\ddot{y}$ **[rad] der Winkel zwischen der Symmetrieebene der Antenne (betrachtet als symmetrischer Dipol der Gesamtlänge** h **[m]) und der Hauptachse der Antenne ist. Dieses Strahlungsmuster zeigt die (für große Plasmaantennen) übliche Eigenschaft, die Energie in eine Richtung nahe der Hauptachse der Antenne abzustrahlen, die sich mit zunehmender Länge der Antenne stärker manifestiert** [24,34,35] und völlig inkonsistent **mit der Form des Musters/Antennenverhaltens unter nicht-**

Machine Translated by Google

Diese Eigenschaft bleibt nicht nur für Antennentestzwecke und die eigentlichen Anwendungen von Sendeantennen im Plasma interessant, sondern auch für in Plasmen eingetauchte Empfangsantennen , die sich – während sich ihr Empfangsmuster im Verhältnis zum sie umgebenden Plasma weiterentwickelt – in einer Weise verändern können, die einige Richtungen zur Antenne vom Signalempfang ausschließt ; dies bleibt für die genaue Beschreibung beispielsweise der Eingangsrichtungen und der tatsächlichen Signalstärken der Aurora

Eine Methode, die auf einer einfachen Transformation der analytischen Definitionen der Form des Strahlungsmusters von den Luft-/Vakuumbedingungen zu den Plasmabedingungen basiert – unter der Annahme, dass dieser Ansatz für die experimentell definierten Musterformen gültig bleibt – kann für einen gegebenen *ÿ-* **Winkelwert wie folgt definiert werden:**

$$T\ddot{y}=\text{const.}(fA\ \text{real}(\ddot{y})\ \ddot{y}\ fP\ \text{real}(\ddot{y})) : fP\ \text{real} = fA\ \text{real}(\ddot{y})\ fA(\ddot{y}) \qquad \frac{fP(\ddot{y})}{} \qquad \textbf{(11)}$$

wobei fi *(ÿ)* **die Funktionen der Antennenstrahlungsmuster sind und die Indizes P,** A **und** real **anschließend Luft/Vakuum-, Plasma- und Realbedingungen beschreiben. Das Verhältnis der Plasma- zu Luft/Vakuum-Strahlungsmusterfunktionen erfordert eine zusätzliche Division jeder fi** (ÿ) **-Funktion durch den von ihr erreichten Maximalwert, um die nachfolgenden Fälle wertmäßig vergleichbar zu machen (dadurch würden jedoch die genauen Informationen zur Richtwirkung gelöscht). Natürlich darf für diese Transformation der Wert von fA(ÿ) nicht Null erreichen (kann einen Realwert von** *ÿ* **ÿ 0 erreichen) und bleibt für die folgende Bedingung gültig (sonst handelt es sich um eine Tautologie):**

$$\int_{\ddot{y}\ddot{y}0}^{\ddot{y}_{\ddot{y}y\,2}} fA(\ddot{y})d\ddot{y}\ \ddot{y} = \int_{\ddot{y}\ddot{y}0}^{\ddot{y}_{\ddot{y}y\,2}} fP(\ddot{y})d\ddot{y} \qquad \textbf{(12)}$$

Für die tatsächlichen Weltraumbedingungen und die gemessenen verschiedenen Radiosignale, die in Absatz 1 erwähnt wurden, ist die genaue Richtung des eingehenden Signals nicht immer bekannt, und die Entwicklung des Strahlungs-/Empfangsmusters der Antenne für verschiedene Plasmabedingungen wird noch mehr Variablen in seine Definition einführen. Wenn ein zweidimensionales Frequenzspektrum von einer solchen Antenne erhalten wird und grundlegende Plasmaparameter zugänglich sind (wie in Absatz 2 beschrieben), kann das Verhältnis aus der Formel (11) erweitert werden, um einen allgemeinen frequenz- und winkelabhängigen Koeffizienten w(ÿ,f) zu bilden, der das Plasmastrahlungsmuster (10) und das Luft/Vakuum-Strahlungsmuster für kurze lineare Antennen umfasst **[36]; die Unterschiede zwischen den Luft/Vakuum- und Plasma-Antennenstrahlungsmustern sind im Raum um den Antennendraht und über den gesamten möglichen Bereich von** *ÿ* definiert **(mit der integrationserleichternden Annahme, dass die Antenne ein symmetrischer Dipol der Gesamtlänge**

$$w(\ddot{y},\ f) = 4\ddot{y} \int_{\ddot{y}\ddot{y}0}^{\ddot{y}_{\ddot{y}y\,2}} \frac{fP(\ddot{y})}{\textbf{max[f }_{P}(\ddot{y})]} \frac{fA(\ddot{y})}{\textbf{max[fA(ÿ)]}} \ddot{y}1\ d\ddot{y} = 4\ddot{y} \int_{\ddot{y}\ddot{y}0}^{\ddot{y}_{\ddot{y}y\,2}} \frac{\text{Sünde } \frac{-}{1\,2}k\text{ih}(1\ \textbf{ÿ cos}\ \ddot{y})}{\textbf{sin}\ \ddot{y}(1\ \textbf{ÿ cos}\ \ddot{y})\ \textbf{cos2}\ \frac{2\ddot{y}\ \text{fh}}{c}\textbf{cos}\ \ddot{y}} \frac{\textbf{max[fA(ÿ)]}}{\textbf{max[f }_{P}(\ddot{y})]} d\ddot{y} \qquad \textbf{(13)}$$

Eine Näherungslösung kann wie folgt berechnet werden:

$$w(\ddot{y},\ f)\ \ddot{y}\ \ddot{y} \sum_{\textbf{max[f 2}}^{N\ \ i=1} \frac{\textbf{max[fA(ÿ)] 1}\ \ddot{y}i}{P(\ddot{y})]} \left[\frac{\text{Sünde } \frac{-}{1\,2}k\text{ih}(1\ \textbf{ÿ cos}\ \ddot{y}i)}{\textbf{sin}\ \ddot{y}i(1\ \textbf{ÿ cos}\ \ddot{y}i)\ \textbf{cos2}\ \frac{2\ddot{y}\ \text{fh}}{c}\textbf{cos}\ \ddot{y}i} + \frac{\text{Sünde } \frac{-}{1\,2}k\text{ih}(1\ \textbf{ÿ cos}\ \ddot{y}i\ddot{y}1)}{\textbf{sin}\ \ddot{y}i\ddot{y}1(1\ \textbf{ÿ cos}\ \ddot{y}i\ddot{y}1)\ \textbf{cos2}\ \frac{2\ddot{y}\ \text{fh}}{c}\textbf{cos}\ \ddot{y}i\ddot{y}1} \right] \textbf{ÿ ÿ} \qquad \textbf{(14)}$$

Da die Wellenzahlen ki die **B0- Magnetfeldkomponente enthalten können ,** ist es **notwendig, diese Komponente so zu approximieren, dass sie die Struktur und Richtung des Magnetfelds um die Antenne herum widerspiegelt. Für den Fall, dass die Magnetfeldinduktion** entlang der Z- Achse der Antenne **konstant bleibt , kann der** B0- **Parameter wie folgt gebildet werden:**

$$\textbf{B0} = \int_{0}^{2\ddot{y}} \textbf{B0}\ \text{max.}\ |\textbf{sin}\ \ddot{y}|\ d\ddot{y} = 4\text{B0 max.} \qquad \textbf{(15)}$$

Machine Translated by Google

wobei *ÿ* **[rad] der Winkel ist, der in der Ebene parallel zur Symmetrieebene des Dipols liegt. Wenn B0 entlang des Antennendrahts nicht konstant bleibt und eine bekannte Verteilungsfunktion mit dem Maximalwert von B0** max. **und dem Argument von** z hat, wird die Formel **wie folgt dargestellt :**

$$B0(z) = \int_0^{\frac{h}{2}} \int_0^{2\ddot{y}} \mathbf{B0}\ \text{max.}(z)|\sin \ddot{\mathbf{y}}|\, d\ddot{y}\, dz = \mathbf{4} \int_0^{\frac{h}{2}} \mathbf{B0\ max.(z)}\, dz \quad \mathbf{(16)}$$

Die Abbildungen 42–45 zeigen den w(f) **-Koeffizienten (nach der Integration über ÿ), berechnet für einen weiten Frequenzbereich für verschiedene Elektronenkonzentrationen (wie in den vorherigen Absätzen) und nachfolgende Antennenlängen: 1, 10 und 100 m. Da die Formeln (13) und (14) die Formel (10) einbeziehen , die im Wesentlichen für den Frequenzbereich des Whistler-Ausbreitungsmodus gültig ist , überschreitet die maximale Frequenz in diesen Abbildungen den Wert von 1,4 MHz nicht – die Berechnung über diese Grenze hinaus wäre nach der Klärung der Genauigkeit von Formel (10) oberhalb dieser Frequenz möglich.**

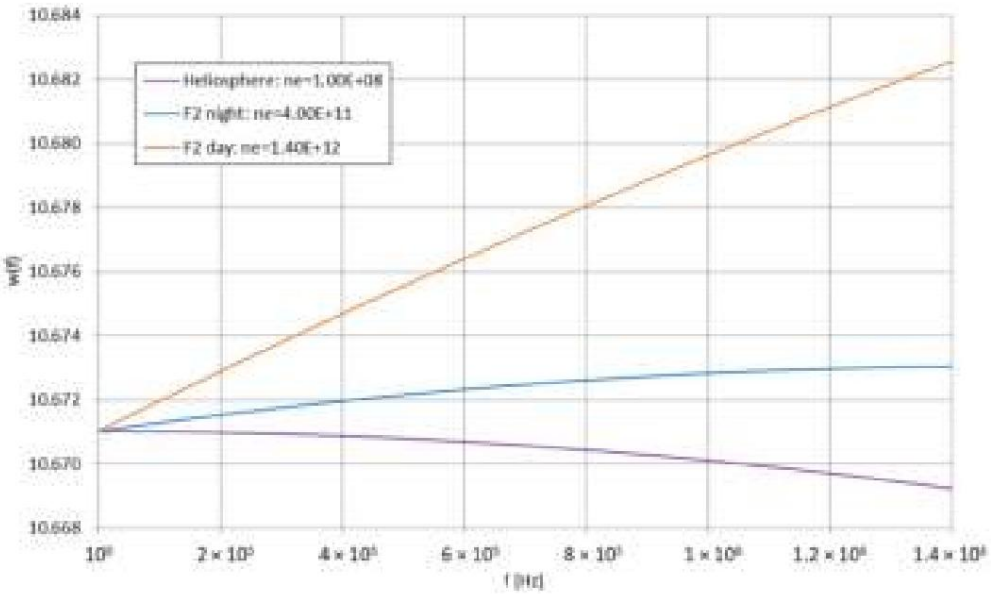


Abbildung 42. Der berechnete w(f) **-Transformationsparameter für eine Antennenlänge** h **= 1 m.**

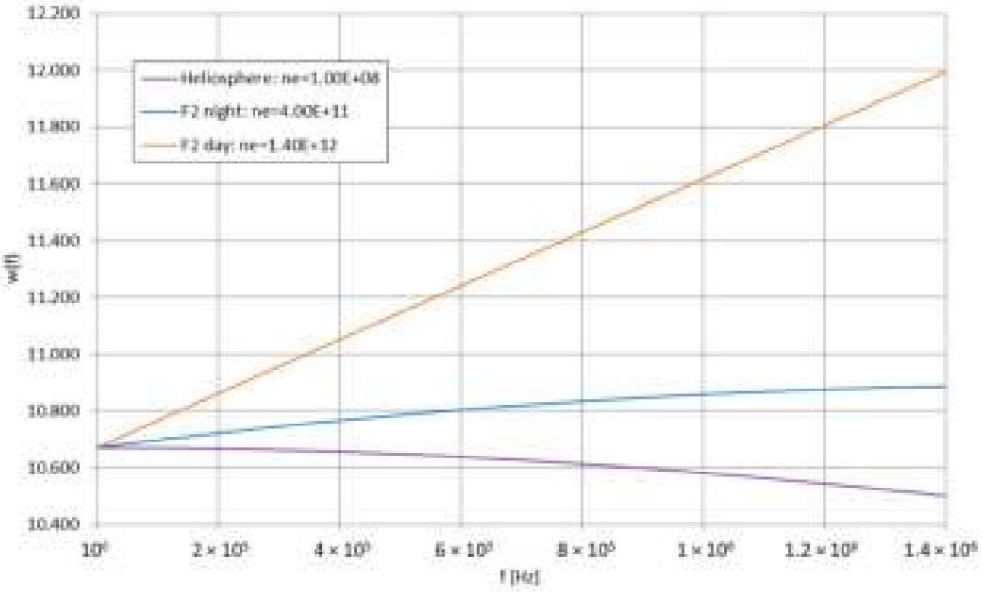


Abbildung 43. Der berechnete w(f) **-Transformationsparameter für eine Antennenlänge** h **= 10 m.**

Machine Translated by Google

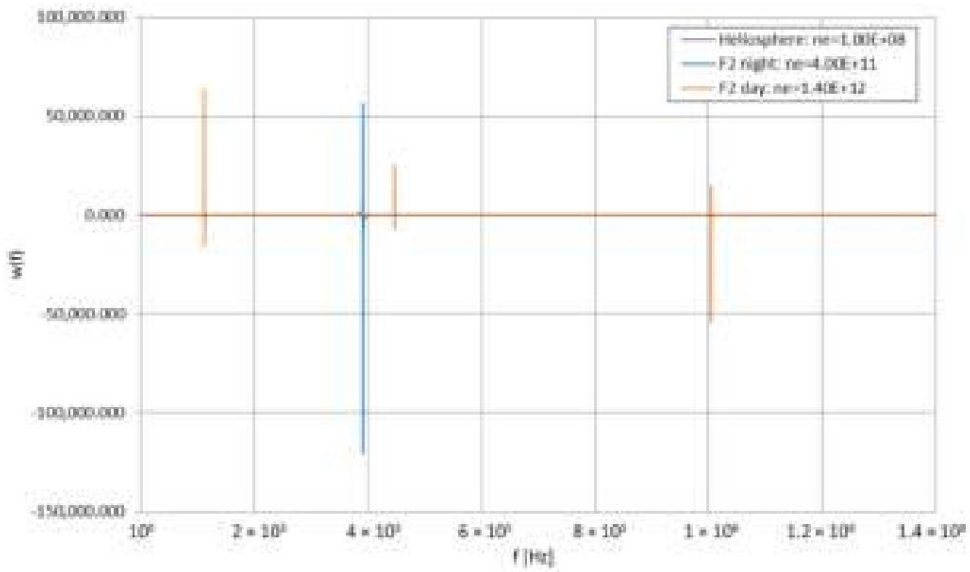


Abbildung 44. Der berechnete w(f) **-Transformationsparameter für eine Antennenlänge** h **= 100 m.**

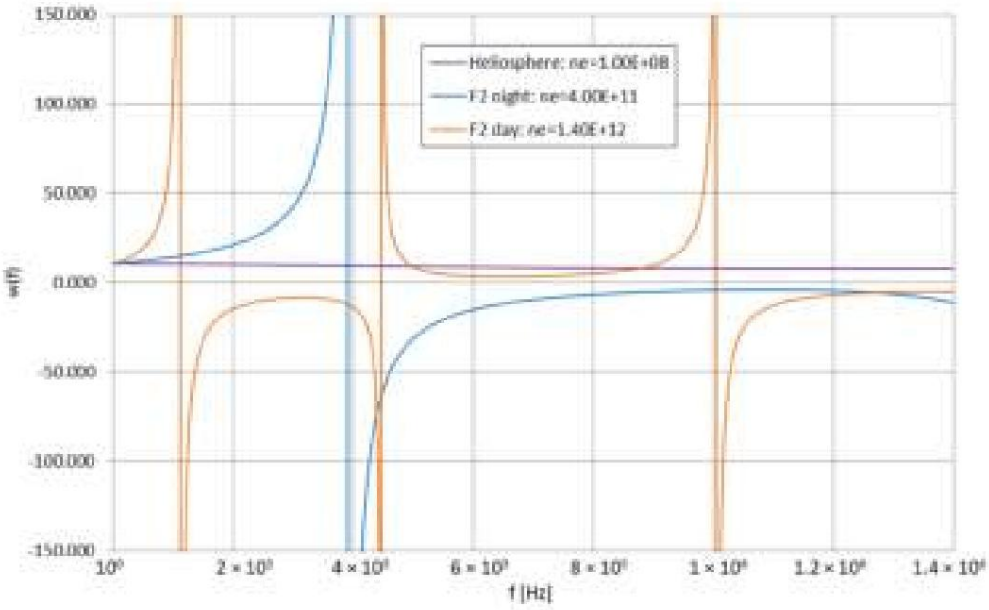


Abbildung 45. Der berechnete w(f) **-Transformationsparameter für eine Antennenlänge** h **= 100 m – vergrößert.**

Bei den kleineren Antennenlängen zeigen die berechneten Koeffizienten ein nahezu konstantes Verhalten, das bei dichteren Elektronenbedingungen (F2-Schicht bei Tageslicht) bei höheren Frequenzen leicht ansteigt und bei weniger dichten Elektronenbedingungen (heliosphärisch) leicht abnimmt. Bei der größten betrachteten Antennenlänge von 100 m zeigt der Koeffizient mehrere signifikante Wertänderungen bei den elektronendichteren Bedingungen, die die Unterschiede zwischen den Antennenstrahlungsm
Da der Koeffizient den Luft-/Vakuumfall wertmäßig in den Plasmafall bringen soll, reagiert er auf das Auftreten der Lappen im Muster, einschließlich aller Vorzeichenwechsel der Funktion, die – aus mathematischer Sicht – auftreten können.

6. Diskussion

Die dargestellten Entwicklungen der grundlegenden Antennenparameter, die ihre Fähigkeit beschreiben, elektromagnetische Energie abzustrahlen oder zu empfangen, zeigen deutlich die Abhängigkeit von den elektrischen Bedingungen oder dem umgebenden Raum – wie dargestellt, beeinflussen diese Bedingungen nicht nur die Impedanz der Antenne (wie in verschiedenen früheren Untersuchur

Machine Translated by Google

aber seine Funktion im Weltraum als Energiestrahler oder Energieempfänger. Dies ist aus zwei Hauptaspekten wichtig – dem Aspekt des Raumfahrzeug- und Antennendesigns und dem Aspekt der Interpretation und Verwendung von Daten des vorhandenen Raumfahrzeugs und seiner Instrumentierung.

Der Entwurfsaspekt konzentriert sich hauptsächlich auf die Anordnung der Antennen in der Nähe des Raumfahrzeugs, um zu verhindern, dass die Messungen des Plasmas direkt vom Raumfahrzeug beeinflusst werden (das um die Struktur des Raumfahrzeugs herum auftritt – es sei denn, dies ist eine spezielle Anforderung für die Mission). Dieser direkte Einfluss kann mit der **Anwesenheit der leitenden Strukturen (Ausleger/Träger, Solarpanele, Raumfahrzeugrahmen usw.) innerhalb der Fresnel- oder Fraunhofer-Zonen um die Antenne herum zusammenhängen .**

Die in Abschnitt 2 mit verschiedenen Formeln berechneten Zonenradien zeigen deutlich die Entwicklung mit den sich ändernden Plasmaparametern – sie bleiben hochempfindlich gegenüber unterschiedlichen Elektronenkonzentrationen ne . **Bei den kleineren Antennen (~100 m lang) sind alle Zonenradien relativ verdichtet; es ist daher nicht schwierig, in die Anforderungen an die mechanische Konstruktion der Instrumente die Anforderung aufzunehmen, größere Strukturen jenseits des Radius der Fraunhofer-Zone zu platzieren (bis zu einer maximalen Frequenz von Interesse).**
Mit zunehmender Antennenlänge (~101–102 m) wird die genaue (formelbasierte)
Die Fraunhofer-Zone wird zu groß, um sie einfach als Maßanforderung einbeziehen zu können; daher können die Radien der Fresnel-Zone oder die Näherungsformel für den Radius der Fraunhofer-Zone, die aus den Eigenschaften der elektromagnetischen Felder um die Antenne abgeleitet wird, in Betracht gezogen werden.

Wenn das Raumfahrzeug bereits entworfen, gebaut und in Betrieb ist, können diese Radien als Indikatoren verwendet werden, ob es sinnvoll wäre, das tatsächliche Strahlungs- /Empfangsmuster der Antenne auf eine Weise zu überprüfen, die die mechanischen Elemente des Raumfahrzeugs als passive Antennenreflektoren einbezieht. Dies ist insbesondere für den Empfang von Funksignalen wichtig, die sich im Whistler- und Alfvén-Wellenmodus ausbreiten, deren Radien in den Fraunhofer- und Fresnel-Zonen deutlich größer sind und bis zur Größe einer ganzen Satellitenkonstellation reichen.

Der Anstieg der Elektronenkonzentration ne beeinflusst auch die Ausformung des **Strahlungs-/Empfangsmusters der Antenne .** Bei den höchsten betrachteten Konzentrationen, **die denen in der ionosphärischen F2-Schicht bei Tageslichtbedingungen (Photoionisation usw.) proportional sind, ist das Auftreten der getrennten Musterkeulen (Maxima und Minima) sowohl für den Whistler-Ausbreitungsmodus bei niedrigeren Frequenzen (~kHz und niedriger) als auch für die hohen MHz- Frequenzen vorhanden. Die Zunahme der Ausformung ist auch direkt mit dem Verhältnis der Wellenlänge des eingehenden Signals zur Antennenlänge verbunden, was sich gleichzeitig mit dem Einfluss des wachsenden** ne manifestiert .

Ausreichend ausgeprägte Lappen erhöhen den Richtwirkungsparameter, der – zusammen mit dem Effizienzparameter der Antenne – verwendet werden kann, um den Antennengewinn zu definieren oder einfach die Wirksamkeit der Antennenstrahlung beim Empfang bei einer gegebenen Frequenz in einem Plasma mit gegebenen Parametern zu beschreiben. Das Verhalten der Richtwirkung ist für jeden Ausbreitungsmodus und jeden betrachteten ne- **Wert deutlich unterschiedlich – dies erfordert eine sorgfältige Herangehensweise für einen präzisen Satz von Messvoraussetzungen, z. B. Frequenzbereich und erwartete Plasmaparameter. Im Allgemeinen ist festzustellen, dass kleinere Antennen (Länge proportional zu 100 m) anfälliger für Änderungen der Richtwirkung sind, insbesondere im hohen MHz- Frequenzbereich – ähnlich dem bei diesen Antennen beobachteten Scalloping-Verhalten. Bei hohen** ne **-Werten weist der Whistler-Ausbreitungsmodus zwei globale Richtwirkungsmaxima auf, die in jedem betrachteten Antennenlängenfall auftreten, in der Umgebung mit geringem Elektronengehalt aber nicht vorhanden sind.**

Abbildung 45 zeigt deutlich die Veränderungen, die die Plasmaumgebung in den Strahlungs-/Empfangsmustern der Antenne bei großen Antennenöffnungen (~102 m) hervorruft – der dargestellte Koeffizient, der das Luft-/Vakuummuster in das Plasmamuster umwandeln soll, spiegelt das entwickelte Plasmamuster wider, das reich an Maxima und Minima ist (tatsächliche Diagramme des Plasmastrahlungsmusters , die die scheinbare Strahlung entlang der Hauptachse der Antenne zeigen, finden sich z. B. in [35]). Dies beweist, dass die Messungen auf der Erde kein vollständiges Verständnis **des Antennenbetriebs in der Weltraumu**

Machine Translated by Google

Die vorgeschlagene Methode der Mustertransformation ermöglicht jedoch eine grundlegende Beschreibung des angenäherten Antennenverhaltens im Plasma, die für die Analyse der spektralen elektromagnetischen Daten vorhandener Satelliten nützlich sein dürfte – die Definition beispielsweise der Ausbreitungsrichtung des empfangenen Signals oder seiner tatsächlichen Stärke (unter Berücksichtigung der unterschiedlichen Richtwirkung und Auskrümmung der Antenne).

7. Schlussfolgerung

In diesem Artikel werden die grundlegenden Antennenparameter, die mit ihrer Fähigkeit zusammenhängen, elektromagnetische Energie auszustrahlen oder zu empfangen, als Variablen dargestellt, die von den vereinfachten Bedingungen des Plasmas abhängen, in das die Antenne eingetaucht ist. Es werden drei Ausbreitungsmodi (klassisch-ionosphärisch, Whistler und Alfvén**) sowie drei exemplarische Elektronenkonzentrationen (verallgemeinerte heliosphärische und ionosphärische F2-Schicht: Tageslicht und Nacht) und eine konstante magnetische Feldinduktion berücksichtigt. Die berücksichtigten Antennenparameter sind die Strahlungszonen (induktiv, Fresnel, Fraunh**

Es zeigt sich, dass sich diese Parameter mit den sich ändernden Plasmabedingungen stark verändern . Die Strahlungszonen bleiben bei den kleineren Antennengrößen verdichtet und dehnen sich bei den größeren Antennenlängen deutlich aus, was trotz der Vergrößerung der Antennenöffnung als technischer Nachteil angesehen werden kann. Außerdem werden bei größeren Antennen die Fresnel-Zonen die größten physikalisch anwendbaren Zonen. Die Ausbuchtung (Anzahl der Lappen im Strahlungsmuster) wird sowohl durch die Erhöhung der Elektronendichte als auch durch das Verhältnis von Antennenlänge zu Wellenlänge verstärkt.

In diesem Artikel wird auch eine Methode zur Umwandlung des auf der Erde gemessenen Antennenstrahlungs-/Empfangsmusters in das in der Weltraumplasmaumgebung zu erwartende Muster vorgeschlagen. Diese Methode funktioniert bei größeren Antennenlängen (> 10 m) effektiv und ermöglicht eine nützliche Umwandlung des Antennenstrahlungsmusters von der Luft-/Vakuumumgebung (Testumgebung) in die erwarteten Plasmabedingungen.

Die präsentierten Daten – in einer praktischen Form von Diagrammen für verschiedene Parameter und Variablen oder in Form von Formeln, die sich leicht an einzelne Plasmafälle anpassen lassen – ergänzen die bestehenden Arbeiten zu Impedanzen, Strömen und Hüllen von Plasmaantennen wirksam **und können zur tatsächlichen Bewertung der Antennenleistung in bestehenden Weltraumwettersatelliten verwendet werden . Dies soll eine verbesserte Triangulation der registrierten eingehenden Signale sowie zusätzliche Anforderungen/Einschränkungen für das Instrumentendesign des neuen Satelliten** ermöglichen, was beispielsweise hilfreich sein soll, um die Messausrüstung vom **direkten Einfluss des Plasmas wegzubewegen, das** um das Raumfahrzeug selbst herum auftritt.

Finanzierung: Diese Forschung erhielt keine externe Finanzierung.

Datenverfügbarkeitserklärung: Daten sind auf Anfrage erhältlich.

Interessenkonflikte: Der Autor erklärt, dass keine Interessenkonflikte vorliegen.

Verweise

1. **Green, JL; Inan, US Blitzeffekte auf Weltraumplasmen und Anwendungen. In** Plasmaphysik Angewandte; **Forschungswegweiser; Grabbe, C., Hrsg.; American Physical Society: College Park, MD, USA, 2007.**
2. **Morozova, EA; Mogilevsky, MM; Hanasz, J.; Rosanov, AA Die Feinstruktur des elektromagnetischen Feldes des AKR, gemessen mit der Satellit Interball-2.** Cosm. Res. **2002,** 40, **404–410. [CrossRef]**
3. **Chuchra-Konrad, A.; Matyjasiak, B.; Schreiber, R.; Rothkaehl, H. Eine Korrelation zwischen AKR-ähnlichen Emissionen und feldorientierten Strömungen. In Proceedings of the URSI GASS 2023, Sapporo, Japan, 19.–26. August 2023.**
4. **Bresler, K.; Zawiÿa, R.; Ciarka, M.; Shmyk, A.; Mi´s, TA; Mystkowski, R.; Martyniak, J.; Wiater, K. TOTORO-Testbeobachtungen von transienten Objekten und RadiO an Bord von BEXUS 33. In Proceedings des 26. ESA PAC Symposiums zu europäischen Raketen- und Ballonprogrammen und damit verbundener Forschung, Luzern, Schweiz, 19.–23. Mai 2024.**
5. **Gurnett, DA; Anderson, RR; Scarf, FL; Fredricks, RW; Smith, EJ Erste Ergebnisse der ISEE-1- und -2-Plasmawellenuntersuchung .** Space Sci. Rev. **1979,** 23, **103–122. [CrossRef]**
6. **Gallet, RM; Helliwell, RA Ursprung von „sehr niederfrequenten Emissionen“.** J. Natl. Bur. Stand.—D Radio Propag. **1959,** 63D, **21–27. [Querverweis]**
7. **Ferencz, C.; Lichtenberger, J.; Hamar, D.; Ferencz, OE; Steinbach, P.; Székely, B.; Papagei, M.; Lefeuvre, F.; Berthelier, J.-J.; Clilverd, M. Eine ungewöhnliche VLF-Signatur, aufgezeichnet vom DEMETER-Satelliten.** J. Geophys. Res. **2010,** 115, **A02210. [CrossRef]**

Machine Translated by Google

8. **Siingh, D.; Singh, AK; Patel, RP; Singh, R.; Singh, RP; Veenadhari, B.; Mukherjee, M. Gewitter, Blitze, Sprites und magnetosphärische Whistler-Mode-Radiowellen.** Surv. Geophys. **2008,** 29, **499–551.** [CrossRef]
9. **Xiao, F.; Zhou, Q.; Su, Z.; He, Z.; Yang, Z.; Liu, S.; He, Y.; Gao, Z. Erklärung des Vorkommens von Polarlichtstrahlung in Van Allen-Strahlungsgürtel.** Geophys. Res. Lett. **2016,** 43, **971–978.** [CrossRef]
10. **Kishore, A.; Deo, A.; Kumar, S. Fernerkundung der oberen Atmosphäre mithilfe von ELF-VLF-Blitzen erzeugten Tweek- und Whistler-Sferics.** South Pac. J. Nat. Appl. Sci. **2016,** 34, **12–20.** [CrossRef]
11. **Bÿ ÿecki, J. ULF/ELF-Plasmaturbulenz in Verbindung mit kleinräumigen Strukturen in der Nähe der Erdmagnetopause und des Bugstoßes . In Proceedings des 2. Nationalen Symposiums PLASMA '95 „Forschung und Anwendungen von Plasmen“, Warschau, Polen, 26.–28. Juni 1995; Eingeladene Beiträge; Band 2.**
12. **Mansilla, GA Ionosphärische Effekte eines intensiven geomagnetischen Sturms.** Stud. Geophys. Geod. **2007,** 51, **563–574.**

[CrossRef] **13. Šuli´c, DM; Sre´ckovi´c, VA; Mihajlov, AA Eine Studie über VLF-Signalvariationen im Zusammenhang mit den Änderungen des Ionisations D-Region als Folge solarer Bedingungen.** Adv. Space Res. **2016,** 57, **1029–1043.** [CrossRef]

14. **Thomson, NR; Clilverd, MA; McRae, WM Nächtliche ionosphärische D-Region-Parameter aus VLF-Phase und Amplitude.** J. Geophys. Res. Atmos. **2007,** 112. [CrossRef]
15. **Žigman, V.; Dominique, M.; Grubor, D.; Rodger, CJ; Clilverd, MA Elektronendichte in der unteren Ionosphäre und effektive Rekombinationskoeffizienten aus Weltraumbeobachtungen mit mehreren Instrumenten und VLF-Messungen am Boden während Sonneneruptio** 2023 , 247, **106074.** [CrossRef]
16. **Jarvinen, R.; Kallio, E. Globale Hybridmodellierung des Weltraumwetters auf Merkur und BepiColombo-Beobachtungen. In Proceedings of the 4th URSI AT-RASC, Gran Canaria, Spanien, 19.–24. Mai 2024.**
17. **Mikhailov, YM; Mikhailova, GA; Kapustina, OV ELF- und VLF-elektromagnetischer Hintergrund in der äußeren Ionosphäre über seismoaktiven Regionen (Satellit „INTERCOSMOS-24“). In Proceedings des International Symposium PLASMA '97 „Research and Applications of Plasmas“, Jarnoÿtówek bei Opole, Polen, 10.–12. Juni 1997; Beiträge; Band 1.**
18. **Klimov, S.; Romanov, S.; Savin, S.; Nikolaeva, N.; Skalsky, A.; Petrukovich, A.; Nozdrachev, M.; Triska, P.; Amata, E.; Ciobanu, M.; et al. INTERBALL-Schwanzsonde. Untersuchung magnetosphärischer Plasmagrenzen. In Proceedings des International Symposium PLASMA '97 „Forschung und Anwendungen von Plasmen“, Jarnoÿtówek bei Opole, Polen, 10.–12. Juni 1997; Beiträge; Band 1.**
19. **Bÿ ÿecki, J.; Kosacki, K.; Wronowski, R.; 'Savin, S.; Simunek, J.; Smilauer, J.; Juchniewicz, J.; Klimov, S.; Triska, P. Elektrostatische und elektromagnetische Komponenten der Plasmawelle, die an Bord der Satelliten Prognoz-8 und Magion-4 in der polaren Spitze beobachtet wurden. In Proceedings des Internationalen Symposiums PLASMA '97 „Forschung und Anwendungen von Plasmen“, Jarnoÿtówek bei Opole, Polen, 10.–12. Juni 1997; Beiträge; Band 1.**
20. **Bÿ ÿecki, J.; Kosacki, K.; Wronowski, R.; Nemecek, Z.; Safrankova, J.; Simunek, J.; Smilauer, J.; Borodkova, N.; Romanov, SA; Savin, SP; et al. Von Magion-4-Subsatellit im Plasmablatt beobachtete Elektronen-Zyklotron-Harmoniken. In Proceedings des International Symposium PLASMA '97 „Forschung und Anwendungen von Plasmen“, Jarnoÿtówek bei Opole, Polen, 10.–12. Juni 1997; Beiträge; Band 1.**
21. **Rothkaehl, H.; Bÿ ÿecki, J.; Kÿos, Z. Die Signatur der ionosphärischen-magnetosphärischen Kopplung, untersucht mit dem Plasmawellenpaket PWP an Bord des CESAR-Satelliten. In Proceedings des Internationalen Symposiums PLASMA '97 „Forschung und Anwendungen von Plasmen“, Jarnoÿtówek bei Opole, Polen, 10.–12. Juni 1997; Beiträge; Band 1.**
22. **Wooliscroft, LJC Plasmawellenmessungen mit der ESA/NASA CLUSTER-Mission. In Proceedings des 2. Nationalen Symposiums PLASMA '95 „Forschung und Anwendungen von Plasmen“, Warschau, Polen, 26.–28. Juni 1995; Eingeladene Beiträge; Band 2.**
23. **McCollough, JP; Miyoshi, Y.; Ginet, GP; Johnston, WR; Su, Y.-J.; Starks, MJ; Kasahara, Y.; Kojima, H.; Matsuda, S.; Shinohara, I.; et al. Weltraum-zu-Weltraum-Radioübertragung mit sehr niedriger Frequenz in der Magnetosphäre mithilfe der Satelliten DSX und Arase.** Earth Planets Space **2022,** 74, **64.** [CrossRef]
24. **Lauben, D.; Linscott, I.; Inan, U.** Der Breitbandempfänger (BBR) für Demonstrationen und wissenschaftliche Experimente (DSX) Raumfahrzeug; **Abschlussbericht Nr. AFRL-RV-PS-TR-2022-0114; Hanson Experimental Physics Laboratory, Stanford University: Stanford, CA, USA, 15. Dezember 2022.**
25. **Neuffer, A.; Lunk, A. Diagnostik einer planaren Magnetron-Sputterquelle mit Langmuir-Sonden. In Proceedings des International Symposium PLASMA '97 „Research and Applications of Plasmas“, Jarnoÿtówek bei Opole, Polen, 10.–12. Juni 1997; Eingeladene Beiträge; Band 2.**
26. **Galejs, J. Impedanz einer endlichen isolierten zylindrischen Antenne in einem kalten Plasma mit einem longitudinalen Magnetfeld.** IEEE Trans. Antennas Propag. **1966,** 14, **727–736.** [CrossRef]
27. **Galejs, J. Impedanz einer endlichen isolierten Antenne in kaltem Plasma mit senkrechtem Magnetfeld.** IEEE Trans. Antennas Propag. **1966,** 14, **737–748.** [CrossRef]
28. **Wunsch, DA Strahlung von einem Streifen elektrischen Stroms in magnetoionischem Medium.** Can. J. Phys. **1967,** 45, **1675–1691.** [CrossRef]
29. **Bhat, B. Stromverteilung auf einer unendlichen röhrenförmigen Antenne, die in ein kaltes Kollisionsmagnetoplasma eingetaucht ist.** Radio Sci. **1973, 8 483–492.** [CrossRef]
30. **Seshadri, SR Eingangsadmittanz einer zylindrischen Antenne in magneto-ionischem Medium.** J. Appl. Phys. **1968,** 39, **2407–2412.** [CrossRef]
31. **Song, P.; Reinisch, BW; Paznukhov, V.; Sales, G.; Cooke, D.; Tu, J.-N.; Huang, X.; Bibl, K.; Galkin, I. Hochspannungsantenne-Plasma-Wechselwirkung bei der Übertragung von Whistler-Wellen: Plasmamanteleffekte.** J. Geophys. Res. **2007,** 112, **A03205.** [CrossRef]
32. **Soloviev, OV Niederfrequente Radiowellenausbreitung im Erd-Ionosphären-Wellenleiter gestört durch ein großräumiges drei-dimensionale Unregelmäßigkeit.** Radiophys. Quantum Electron. **1998,** 41, **392–402.** [CrossRef]

Machine Translated by Google

33. Wait, JR Theorie der Strahlung aus Quellen in anisotropen Medien. J. Res. Natl. Bur. Stand.—B. Math. Math. Phys. **1964,** 68, **119–136. [CrossRef]**

34. Ginet, GP VLF Fernfeldstrahlung von einer linearen Dipolantenne, die in ein Plasma eingebettet ist: Basismodell für DSX; **Bericht Nr. 95-0002; Lincoln Laboratory, Massachusetts Institute of Technology: Lexington, MA, USA, 15. Juni 2020.**

35. Bannister, PR; Harrison, JK; Rupp, CC; King, RWP; Cosmo, ML; Lorenzini, EC; Dyer, CJ; Grossi, MD Orbitalsender und Antenne für weltraumgestützte Kommunikation bei ELF/VLF zu untergetauchten U-Booten. In Proceedings der AGARD-Konferenz Proceedings 529: ELF/ VLF/LF-Funkausbreitung und Systemaspekte, Symposium des Electromagnetic Wave Propagation Panel, Brüssel, Belgien, 28. September–2. Oktober 1992.

36. Poisel, RA Antennensysteme und Anwendungen für elektronische Kriegsführung; **Artech House: Boston, MA, USA; London, UK, 2012.**

37. Modelski, J.; Jaszczyszyn, E.; Chaci ´nski, H.; Majchrzak, P. Messungen von Antennenparametern; **OFPW, Technische Universität Warschau : Warschau, Polen, 2004.**

38. Bem, DJ- Antennen und Ausbreitung von Radiowellen; **WNT: Warschau, Polen, 1973.**

39. Celi ´nski, Z. Plasma; **PWN: Warschau, Polen, 1980.**

40. Bogdanov, AF; Vasin, VV; Dulin, VN; Ilin, VA; Krivitskiy, BH; Kuznetsov, VA; Labutin, VK; Molochkov, YB; Piertsov, SV; Stiepanov, BM; et al. Kompendium zur Radioelektronik; **Kulikovskiy, AA, Hrsg.; WKiÿ: Warschau, Polen, 1971.**

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Eine ressourcenbasierte dynamische Preisgestaltung und erzwungene Weiterleitung Anreizalgorithmus im Socially Aware Networking

Xuemin Zhang 1,2, Yuan Li 1,2,*, Zenggang Xiong [1,2,*], Yanchao Liu [1], Shihui Wang [2] und Delin Hou 1,2

[1] Fakultät für Computer- und Informationswissenschaften, Hubei Engineering University, Xiaogan 432000, China; zhxm9701@163.com (XZ); yanchaoliu@cuc.edu.cn (YL); hdl00@stu.hubu.edu.cn (DH)

[2] Fakultät für Informatik und Informationstechnik, Hubei-Universität, Wuhan 430062, China; wsh@hubu.edu.cn

* Korrespondenz: 202221116012706@stu.hubu.edu.cn (YL); jkxxzg2003@163.com (ZX)

Zusammenfassung: In sozial bewussten Netzwerken verhalten sich Knoten aufgrund von Ressourcenbeschränkungen typischerweise egoistisch und soziale Korrelationen, was zu einer geringen Netzwerkleistung führt. Um egoistische Knoten dazu zu motivieren, aktiv an der Nachrichtenweiterleitung teilnehmen, schlägt dieses Papier eine ressourcenbasierte dynamische Preisgestaltung und erzwungene Forwarding Incentive Algorithm (DFIA). Zunächst führt der Algorithmus virtuelle Währung als Transaktionsmedium ein und entwirft dann eine Preisfunktion basierend auf Faktoren wie dem Ressourcenstatus des Knotens, Beteiligungsbeitrag, Standortrelevanz und soziale Konnektivität. Es stellt sicher, dass die Weiterleitung Dienstleistung wird durch Verhandlungsregeln zu einem angemessenen Preis abgewickelt. Zweitens eine erzwungene Weiterleitung Strategie wird implementiert, um egoistische Knoten zu zwingen, die nicht bereit sind, an den Aktivitäten anderer Knoten teilzunehmen. Nachrichtenweiterleitung, um eine bestimmte Anzahl nicht lokaler Nachrichten weiterzuleiten. In der Zwischenzeit, um Verhindern, dass Knoten Nachrichten verwerfen und eine erfolgreiche Weiterleitung an das Ziel gewährleisten, Es werden bestimmte Regeln verwendet, um Beitragswerte an Knoten zuzuweisen, die erfolgreich an Nachrichten teilnehmen. Weiterleitung. Um falsches Quotierungsverhalten zu vermeiden, wird schließlich die Blockchain-Technologie eingesetzt. Transaktionsinformationen werden in Blöcke verpackt und nach Konsensvalidierung zur Blockchain hinzugefügt durch andere Knoten im Netzwerk, wodurch die Transparenz und Unveränderlichkeit der Transaktionsdaten gewährleistet wird. Die Ergebnisse der Simulation zeigen, dass dieser Algorithmus im Vergleich zu den bestehenden Anreizalgorithmen nicht verbessert nicht nur die Wahrscheinlichkeit der Nachrichtenübermittlung, sondern reduziert auch effektiv die durchschnittliche Latenz.

Schlüsselwörter: sozial bewusste Vernetzung; egoistische Knoten; falsche Zitate; virtuelle Währung

check for updates

Zitat: Zhang, X.; Li, Y.; Xiong, Z.; Liu, Y.; Wang, S.; Hou, D. A Ressourcenbasierte dynamische Preisgestaltung und Anreiz zur erzwungenen Weiterleitung Algorithmus in Socially Aware **Vernetzung.** Electronics **2024,** 13, **3044. https://doi.org/10.3390/Elektronik13153044**

Wissenschaftlicher Redakteur: Domenico Rosaci

Erhalten: 25. Juni 2024
Überarbeitet: 23. Juli 2024
Akzeptiert: 30. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz. Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennung (CC BY)-Lizenz (https://creativecommons.org/licenses/by/4.0/).

1. Einleitung

Mit der rasanten Entwicklung der Netzwerk-Informationstechnologie wird die mobile Umgebung, die durch intelligente Geräte (Smartphones, Tablets, Geräte im Fahrzeug) gebildet wird, mit nachbarschaftsbasierte drahtlose Technologien (Bluetooth, WIFI) [1] haben mehr und mehr Aufmerksamkeit, und Socially Aware Networking (SAN) [2,3] ist dementsprechend entstanden. SAN Die Forschung ermittelt sinnvolle Protokolle zur Bereitstellung von Diensten für Anwendungen der oberen Ebene durch Erfassen und Analysieren der Kontextinformationen des mobilen Knotens [4] im Netzwerk. Dies wird durch die Analyse sozialer Netzwerke und Technologien des maschinellen Lernens erreicht, die Analysieren und lernen Sie aus dem Verhalten und den Umweltinformationen der Träger, die Sie erhalten haben von intelligenten Mobilgeräten. Aus diesen Daten werden soziale Attribute wie Gemeinschaft, Interaktionsintensität und Mobilitätsmuster abgeleitet. Durch die Integration dieser sozialen Attribute mit einem Carrier-Verhalten, effektiven Routing-Protokollen und Nachrichtenweiterleitungsstrategien sind darauf ausgelegt, qualitativ hochwertige Dienste für Anwendungen der oberen Schicht bereitzustellen. Beispielsweise Die Studien in [5,6] trafen Weiterleitungsentscheidungen auf der Grundlage sozialer Merkmale der Nutzer und Ähnlichkeiten [7].

SAN (siehe Abbildung 1) nutzt die sozialen Eigenschaften von Mobilfunkanbietern (d. h. Abonnenten), um effiziente Netzwerkprotokolle zu entwickeln. Es stellt einen neuen Ansatz in der Verzögerungstechnologie dar. Tolerant Network (DTN) [8] und Opportunistic Network (OppNet) [9]. In Bezug auf das Funktionsprinzip ähnelt SAN DTN und OppNet [10]. Einerseits aufgrund der

Machine Translated by Google

Aufgrund der dynamischen Topologie, die durch die Mobilität der Knoten innerhalb des Netzwerks verursacht wird, mangelt es an stabilen End-to-End-Kommunikationspfaden zwischen den Quell- und Zielknoten. Wenn die Verbindung zwischen den Knoten unterbrochen ist, können die herkömmlichen Routing-Methoden die Nachrichtendaten nicht weiterleiten, und häufig werden probabilistische Begegnungen verwendet, um die Weiterleitung im Multi-Hop- Verfahren zu realisieren. Andererseits verwenden Zwischenknoten während der Nachrichtenweiterleitung normalerweise die Methode „Store-Carry-Forward“, um Nachrichten vorübergehend zu speichern [11–13], wobei sie ihre Geräte mit geringer Reichweite und geringen Kosten verwenden, um miteinander zu kommunizieren und so Datenobjekte (wie Bilder, Werbung usw.) zwischen den Knoten auszutauschen **[7]. Im Gegensatz zu DTN und OppNet werden mobile Geräte in SAN jedoch im Allgemeinen von Objekten wie Menschen, Fahrzeugen und Tieren getragen, die sich an verschiedene Orte bewegen und verschiedene Aktivitäten ausführen, was zu einer ständig wechselnden Netzwerktopologie führt. Daher gibt es bestimmte soziale Eigenschaften zwischen den Knoten in SAN. Es ist äußerst wichtig,** effiziente Lösungen zu entwickeln, um

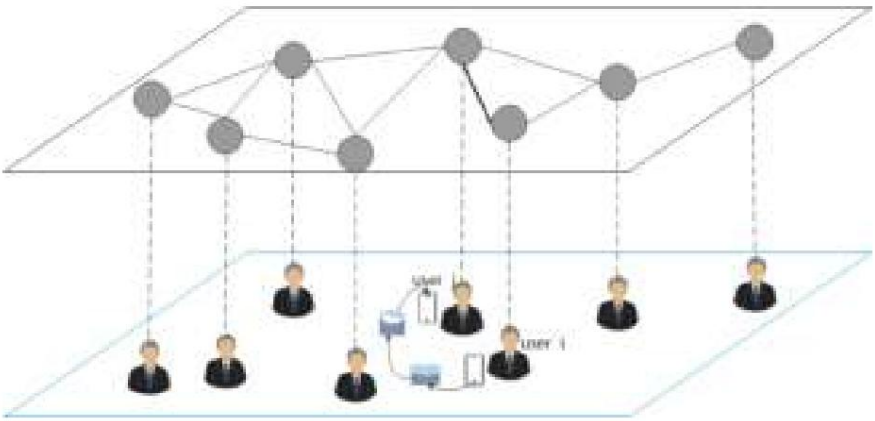


Abbildung 1. Sozial bewusstes Networking.

Als neues Netzwerkmodell ist SAN noch mit zahlreichen Herausforderungen hinsichtlich der Erforschung von Knotenegoismus, Sicherheitsproblemen bei der Nachrichtenweiterleitung und anderen damit verbundenen Technologien konfrontiert . In SAN können einige Knoten aufgrund ihrer begrenzten Ressourcen egoistisches Verhalten zeigen, was zu schlechter Zustellbarkeit oder schlechten sozialen Beziehungen zu anderen Knoten führt. Diese Verhaltensweisen wirken sich negativ auf die Gesamtleistung und -effizienz des Netzwerks aus. Das egoistische Verhalten der Knoten manifestiert sich als ind Knoten, die individuellen Egoismus zeigen, empfangen weder Informationen von anderen Knoten noch leiten sie Informationen an diese weiter, wenn ihre Ressourcen knapp sind. Umgekehrt bieten Knoten, die sozialen Egoismus zeigen, nur Datenweiterleitungsdienste für bekannte Knoten an und weigern sich, Datennachrichten von Knoten mit schlechten sozialen Beziehungen weiterzuleiten oder zu empfangen. Im SAN besteht neben den Herausforderungen, die der Knotenegoismus mit sich bringt, auch das Risiko böswilliger Angriffe wie Manipulation und Diebstahl weitergeleiteter Nachrichten und Transaktionsdatensätze [16]. Es ist daher klar, dass wirksame Anreizmechanismen und strenge Sicherheitsmaßnahmen erforderlich sind, um Knoten zur Teilnahme an der Zusammenarbeit zu motivieren und böswillige Angriffe zu verhindern, was wiederum einen effizienten Betrieb des SAN und eine sichere Datenübertragung gewährleistet.

Derzeit ist die Entwicklung von Anreizmechanismen für Knoten die wichtigste Methode, um das Problem des Knotenegoismus in SAN zu lösen. Vorhandene Anreizmechanismen für die Knotenkooperation können in drei Typen eingeteilt werden: auf Reputation basierende Anreizmechanismen, auf virtueller Währung basierende Anreizmechanismen und auf Spieltheorie basierende Ar

Wie in der Arbeit [17] beschrieben, berechneten die Autoren die Reputation genau, indem sie **Glaubwürdigkeit und Vertrauen der Empfehlungsgeber kombinierten ,** und bewerteten jede Empfehlung fair **Sie setzten einen Mechanismus fairer Leistungsdifferenzierung ein, um eine aktive Teilnahme zu fördern.**

In [18] definierten Roy et al. den Reputationswert von Knoten und die Kosten für die Nachrichtenweiterleitung. Diese Strategie verbessert die Effizienz der Nachrichtenübermittlung, indem Knoten mit einem hohen Reputationswert und geringen Kosten bei der Multicast-Datenübertragung als Next-Hop-Knoten für die Nachrichtenweiterl Bounouni et al. [19] kombinierten direkte und indirekte Reputation, um die Vertrauenswürdigkeit von Knoten zu bewerten.

Machine Translated by Google

ness, wodurch Knoten den Status benachbarter Knoten in ihrer Umgebung genau verstehen können. Anschließend werden die Knoten anhand ihrer Reputationswerte in verschiedene kooperative Zustände eingeteilt und entsprechende Anreize implementiert, um die Zusammenarbeit basierend auf diesen Zuständen zu fördern. Diese Anreizmechanismen stehen bei der Überwachung, Speicherung und Übertragung von Reputationwerten vor Herausforderungen aufgrund potenzieller Angriffe durch böswillige Knoten, die zu Inkonsistenzen bei den Reputationwerten führen können. Darüber hinaus können Knoten in der Nähe des festgelegten Schwellenwerts eine Zufallsmentalität aufweisen und gerade genug Reputationswert aufrechterhalten, um eine aktive Teilnahme an der Weiterleitung zu vermeiden.

Wang et al. [20] schlugen ein evolutionäres Anreizspielmodell vor, das sowohl normale als auch abnormale Knoten durch den evolutionären Prozess des Überlebens des Stärkeren zur Teilnahme an der Netzwerkkooperation ermutigte , wodurch ein hohes Maß an Kooperation sichergestellt und die kontinuierliche Stabilität des Netzwerks aufrechterhalten wurde. In der Literatur **[21] wurde ein repetitives Spielmodell (CGTFT) auf Grundlage der Kooperationsbewertungsfunktion benachbarter Knoten entwickelt. Für rationale Knoten im Netzwerk wird die Kooperationsgradbewertungsfunktion durch Kombination der Restenergie und der erfolgreichen Weiterleitungsrate der Knoten entwickelt. Anschließend wird das CGTFT-Modell für das Spiel verwendet, wobei Knoten belohnt werden, die aktiv an der Kooperation teilnehmen, und egoistische und unkooperative Knoten bestraft werden. Zusätzlich wird ein Vergessensfaktor eingeführt, um die Kooperationsrekonstruktion für Knoten zu erleichtern, die in die Bestrafungsphase eintreten. Die in [22,23]** berichtete Studie **schlug ein spieltheoretisches Preismodell auf Grundlage des Rubinstein-Bargain-Modells vor, das es Relaisknoten ermöglicht, ihre Dienste nach der Restenergie und dem Wert der übertragenen Informationen zu bepreisen und so die Verteilung des Nachrichteninhalts innerhalb des Netzwerks zu verbessern. Diese Mechanismen verwenden zur Analyse die Spieltheorie. Sie erfordern jedoch einen hohen Ressourcenaufwand und sind eher für symmetrische Netzwerke geeignet, wodurch sie für asymmetrische Netzwerke weniger geeignet sind.**

Xiao et al. [24] schlugen ein auf äquivalenten Austauschen basierendes Anreizsystem für die Datenweiterleitung (EEIS) vor. EEIS kombiniert den Ressourcenstatus und den Vermögensstatus von Knoten, um die weiterzuleitenden Nachrichten zu bepreisen und den Vermögens- und Ressourcenstatus beider Handelsparteien öffentlich bekannt zu geben. In [25] führten Li et al. einen auf Servicepreisen basierenden zweistufigen Anreizalgorithmus (STIA) ein. STIA bepreist den Weiterleitungsdienst unter Berücksichtigung von Servicequalität, Nachrichtenattributen und Ressourcenstatus, und die Autoren schlugen einen zweistufigen Anreizmechanismus vor. Es formuliert auch ein Währungsüberziehungssystem und einen zweistufigen Strafmechanismus, um zu verhindern, dass Knoten nach Überziehung der Währung in Zahlungsverzug geraten. Während der Servicetransaktion greifen beide Knoten auf ein Credit Clearing Center (CCC) zu, um die Situation des jeweils anderen zu verstehen und falsche Notierungen zu verhindern. Die Verwendung einer vertrauenswürdigen Drittinstitution zur Verwaltung und Verhinderung falscher Notierungen erhöht jedoch das Risiko von Datenlecks, wenn Sicherheitslücken oder böswillige Angriffe vorliegen. Rao et al. [26] schlug einen ressourcenbeschränkten und auf sozialem Egoismus basierenden Anreizalgorithmus (RSIA) vor. Knoten werden durch Feilschen und Verhandeln mit virtuellen Währungen belohnt, als Anreiz für egoistische Knoten, an der Kooperation teilzunehmen. Der Algorithmus führt den Gini-Koeffizienten ein, um die Einkommensbilanz der Knoten zu bewerten, und belohnt oder bestraft Knoten auf der Grundlage ihrer Einkommensbilanz . RSIA konfiguriert jeden Knoten im Netzwerk mit einem vertrauenswürdigen Sicherheitsmodul. Knoten fügen den entsprechenden Währungswert nur hinzu, wenn die Transaktionsinformationen als unverändert verifiziert werden, wodurch falsche Notierungen vermieden werden. Dieser Algorithmus berücksichtigt jedoch nicht die Positionsbeziehung zwischen Knoten. Bei der Konfiguration des Sicherheitsmoduls H **für Knoten zur Durchführung** einer ACK- **Validierung könnte die Abhängigkeit von Herstellern hinsichtlich des Knotenvertrauens die Sicherheit des ge Daher müssen die Echtheit und Genauigkeit der Transaktionen sichergestellt und falsche Preisangaben verhindert werden.**

Zusammenfassend lässt sich sagen, dass dieses Papier in Bezug auf die obige Analyse einen ressourcenbasierten dynamischen Preis- und Zwangsweiterleitungsanreizalgorithmus (DFIA) vorschlägt, um egoistische Knoten zu motivieren, aktiv an der Nachrichtenweiterleitung teilzunehmen und zu verhindern, dass Knoten falsche Angebote machen. Dieser Algorithmus analysiert das egoistische Verhalten von Knoten und berücksichtigt ihre Ressourcensituation , ihr Beitragsniveau, ihren Vermögensstatus und ihre soziale Relevanz, um Preise für die Nachrichtenweiterleitung auszuhandeln. Zusätzlich

Machine Translated by Google

egoistische Knoten zur Weiterleitung nicht-lokaler Nachrichten und schränken dadurch das Verhalten von Knoten ein, die sich ausschließlich auf ihre eigenen Interessen konzentrieren und nicht bereit sind, an der Netzwerkkooperation teilzu Um falsche Notierungen zu verhindern, werden die Transaktionsinformationen in der Blockchain aufgezeichnet. Dabei wird die dezentrale und transparente Natur der Blockchain-Technologie ausgenutzt und sichergestellt, dass die Preise nicht manipuliert werden können.

Der Rest dieses Papiers ist wie folgt strukturiert: Abschnitt 2 stellt das Netzwerkmodell und die Annahmen des Algorithmus vor. Abschnitt 3 beschreibt die relevanten Definitionen und den Implementierungsprozess des vorgeschlagenen Algorithmus. Abschnitt 4 stellt die Simulationsumgebung vor . Wir legen Spray And Wait [27] als zugrundeliegenden Routing-Algorithmus fest und kombinieren die vorgeschlagene DFIA, RSIA aus dem Papier [26], die STIA aus dem Papier [25] und die EEIS aus dem Papier [24] mit dem zugrundeliegenden Routing-Algorithmus für einen experimentellen Vergleich. Der letzte Teil fasst das gesamte Papier und die Zukunftsaussichten zusammen.

2. Netzwerkmodell und Annahmen

Dieses Dokument präsentiert eine abstrakte Darstellung der Verbindungen zwischen Knoten in Socially Aware Networking (SAN) als ungerichteten Graph G(D, E), wobei D die Menge aller Knoten und E die Menge der Kanten darstellt. Dabei bezeichnet N die Anzahl der Knoten im SAN.
Der Quellknoten ist der Träger der Nachricht vor der Übertragung, und der Relaisknoten ist der Träger, der die Nachricht empfängt. Beim Zusammentreffen von Quellknoten und Relaisknoten wird die Bildung einer Kante zwischen den Knoten eingeleitet. E(i, j) stellt die Kante zwischen den Knoten i und j dar , wobei i ÿ= j.

Darüber hinaus erstellt dieses Dokument in der Anfangsphase der Netzwerksimulation einen Genesis- Block. Wenn Transaktionen für Nachrichtenweiterleitungsdienste zwischen Knoten erfolgreich abgeschlossen werden, werden neue Transaktionsdatensätze in neue Blöcke verpackt und nach dem Genesis-Block sequenziell verknüpft. Dadurch wird sichergestellt, dass alle Transaktionsdatensätze sicher und geordnet in der aus dem Genesis-Block gebildeten Blockchain gespeichert werden. Knoten im Netzwerk nehmen gemeinsam an einer öffentlichen, transparenten und manipulationssicheren Blockchain teil und pflegen diese.

Entsprechend den Anforderungen werden in diesem Dokument die folgenden Annahmen und Definitionen getroffen für das Modell:

1. Jeder Knoten im Netzwerk hat genau eine ID und alle Knoten sind rational, das heißt, sie streben die Maximierung ihres wirtschaftlichen Nutzens oder anderer strategischer Ziele an und halten sich dabei an die Netzwerkprotokolle und -regeln.
2. Das Netzwerk besteht aus zwei unterschiedlichen Knotentypen: normalen Knoten und egoistischen Knoten.
   Normale Knoten nehmen wie erwartet an der Speicherung, Übermittlung und Weiterleitung von Nachrichten teil .
   Egoistische Knoten sind jene Knoten, die sich aufgrund geringer Vorteile oder schwacher sozialer Verbindungen weigern, Nachrichten von anderen Knoten anzunehmen und weiterzuleiten .
3. Gehen Sie davon aus, dass sich das potenziell böswillige Verhalten von Knoten im SAN auf falsche Zitate beschränkt und nicht auf Nachrichtenmanipulationen oder geheime Absprachen beschränkt.
4. Zu einem bestimmten Zeitpunkt ist nur ein Knoten im Netzwerk berechtigt, einen neuen Block zu bilden.

3. Ein ressourcenbasierter Algorithmus für dynamische Preisgestaltung und Anreize zur erzwungenen Weiterleitung

In diesem Dokument wird die Weiterleitung von Nachrichten zwischen Knoten als Transaktion behandelt. Der Käufer ist der Überbringer der Nachricht vor der Weiterleitung, und der Verkäufer ist der Überbringer, der die Nachricht empfängt. Der vom Käufer erbrachte Weiterleitungsdienst ist die Ware, und beide Parteien verwenden virtuelle Währung als Transaktionsmedium. Beide Parteien bieten einen Anfangspreis basierend auf einer Preisfunktion. Wenn das Gebot angemessen ist, ist die Transaktion erfolgreich. Andernfalls werden die Preise gemäß bestimmten Verhandlungsregeln angepasst und die Transaktion erneut versucht. Wenn die Transaktion immer noch fehlschlägt, wird die nächste Nachricht verarbeitet. Sobald eine Transaktion erfolgreich ist, werden die Transaktionsinformationen in einen Block geschrieben und mit der Blockchain verknüpft, wodurch Transparenz und Unveränderlichkeit der Informationen gewährleistet w

Machine Translated by Google

3.1. Relevante Definitionen

3.1.1. Zustellungsdringlichkeit von Knoten

Die Zustellungsdringlichkeit eines Knotens bezieht sich auf die Dringlichkeit, mit der ein Knoten Nachrichten empfangen und weiterleiten kann. Sie wird in erster Linie durch den verbleibenden Pufferspeicher des Knotens, die verbleibende Energie und die verbleibende Gültigkeitsdauer (Time To Live, TTL) der Nachricht bestimmt. Wenn alle drei ausreichend sind, weist dies darauf hin, dass die Zustellungsdringlichkeit des Knotens gering ist und der Knoten über ausreichend Ressourcen und Zeit verfügt, um Nachrichten weiterzuleiten und zu empfangen. Da kontinuierlich Nachrichten empfangen und weitergeleitet werden, verringern sich der Pufferspeicher des Knotens, seine Energie und die Gültigkeitsdauer der Nachricht unvermeidlich allmählich , wodurch die Zustellungsdringlichkeit des Knotens zunimmt. In diesem Dokument verwenden wir einen Statusindex, um diese Zustellung

Der Pufferzustandsindex des Knotens ist das Verhältnis des verbleibenden Pufferspeicherplatzes des Knotens zu einem bestimmten Zeitpunkt zum anfänglichen Gesamtpufferspeicherplatz des Knotens und spiegelt die Pufferkapazität des Knoten Der Energiezustandsindex des Knotens ist das Verhältnis der verbleibenden Energie des Knotens zu einem bestimmten Zeitpunkt zur anfänglichen Gesamtenergie des Knotens und spiegelt die Fähigkeit des Knotens wider, Nachrichten weiterzuleiten. Der TTL-Index der Nachricht ist das Verhältnis der verbleibenden TTL der Nachricht zu einem bestimmten Zeitpunkt zur anfänglichen TTL der Nachricht. Dies ist einer der kritischen Faktoren, die die Fähigkeit des Knotens beeinflussen, die Nachricht zuzustellen. Daher wird die Zustellungsdringlichkeit Qh **des Knotens** x **zum Zeitpunkt** h **wie folgt definiert :**

$$Q^{H}_{x,m} \text{ (1 ÿ).} \quad \frac{\text{Brx Erx Trm} = ÿ \text{ ÿ (1 ÿ)} + ÿ \text{ ÿ (1 ÿ)} + ÿ \text{ ÿ}}{\text{Binitx Einitx Tinitm}} \quad \textbf{(1)}$$

Die Gewichtswerte ÿ, ÿ und ÿ **werden eingeführt, um verschiedene Optimierungsziele auszugleichen. Da es sich bei dem betreffenden Problem um NP-schwer handelt, haben die Gewichtswerte zwischen verschiedenen Instanzen eine erhebliche Auswirkung auf die Leistung des Algorithmus. Daher haben wir in unserer Studie verschiedene Gewichtswerte durch eine große Anzahl von Experimenten angepasst und dann die optimale Gewichtskonfiguration für eine bestimmte Probleminstanz ermittelt.** Brx , Erx **und** Trm **stellen den verbleibenden Pufferplatz, den verbleibenden Energiewert und die verbleibende TTL der Nachricht** m **des Knotens** x **zum Zeitpunkt h dar.** Binitx , Einitx **und** Tinitm **stellen den anfänglichen Gesamtpufferplatz, den anfänglichen Gesamtenergiewert und die anfängliche TTL der Nachricht** m **für Knoten x dar.**

3.1.2. Positionsrelevanz von Knoten

Die Positionsrelevanz von Knoten bezieht sich auf die entfernungsbezogene Situation zwischen dem Knoten und dem Nachrichtenzielknoten. Sie wird mithilfe einer Gaußschen Ähnlichkeitsfunktion für die Positionsattributfunktion des Knotens normalisiert und dient als Maß für die Positionsattribute des Knotens.

Zu den primären Standortattributen eines Knotens gehören die Bewegungsgeschwindigkeit des Knotens, der Winkel zum Zielknoten und die Entfernung. Die Stand eines Knotens wird definiert als

$$\text{loc}(x, d) = \frac{vx \text{ ÿ } |\cos A|}{D(x,d)}, \quad \textbf{(2)}$$

wobei vx **die Bewegungsgeschwindigkeit des Knotens x bezeichnet. Angenommen, die Koordinaten des Knotens** x **und des Zielknotens** d **sind (ax,bx) bzw.** (ad ,bd). **Dann ist** cosA = ÿ $\frac{\text{axÿa} \quad d+\text{bxÿbd}}{\sqrt{\text{ÿ}^2 \text{x+b xÿ a}} \quad \sqrt{d+b_D^2}}$,

wobei 0 ÿ A **ÿ ÿ und** A den Winkel zwischen Knoten x **und Zielknoten darstellt .** D(x, d) = (ax **ÿ** ad) **zwischen Knoten** x **und Zielknoten.** ^{2} + (bx **ÿ** bd) ^{2}**, wobei** D(x, d) **die euklidische Distanz darstellt**

Die Gaußsche Ähnlichkeitsfunktion [28] wird verwendet, um Gleichung (2) zu normalisieren , und die Positionsrelevanz der Knoten wird berechnet als

$$g(x) = e^{\frac{ÿ(\text{loc}(x,d))2}{2ÿ^2}}, \quad \textbf{(3)}$$

wobei ÿ = **1 der Skalenparameter ist. Wenn die Standortattributfunktion des Knotens** x **kleiner wird,** nähert sich **der Wert von** g(x) **1, was bedeutet, dass die Zeit, die Knoten** x braucht , **um das Ziel zu erre**

Machine Translated by Google

Knoten ist kürzer. Im Preismodell ist Knoten x bereit, einen höheren Preis für den **Weiterleitungsdienst zu zahlen, um die Nachricht schnell weiterzuleiten.**

3.1.3. Beteiligungsbeitrag der Knoten

In einem Netzwerk spiegelt der Beteiligungsbeitrag eines Knotens seinen Grad der Beteiligung an der Nachrichtenweiterleitung und seinen Beitrag zur Gesamteffizienz des Netzwerks wider. Dies ist ein wesentlicher Faktor, der die Preisgestaltung eines Knotens beeinflusst. Ein höherer Beteiligungsbeitrag ist ein Hinweis auf die Bereitschaft eines Knotens, Dienste für andere Knoten bereitzustellen und so durch seine Weiterleitungsaktivitäten die Nachrichtenzustellungsrate innerhalb des Netzwerks zu erhöhen. Daher führt der DFIA-Algorithmus den Kooperationsgrad und den durchschnittlichen Beitragsgrad von Knoten ein, um ihren Be

Ob ein Knoten zur Kooperation bereit ist, zeigt sich an seiner Bereitschaft, seine Ressourcen zu verbrauchen, um von anderen Knoten generierte Nachrichten weiterzuleiten. Die Anzahl der von einem Knoten weitergeleiteten nicht-lokalen Nachrichten ist direkt proportional zu seiner Beteiligung. Somit ist der Kooperationsgrad Xx(h) **des Knotens** x **zum Zeitpunkt** h **definiert als**

$$Xx(h) = \frac{\text{Effekt}}{Mx}. \quad \textbf{(4)}$$

wobei Fx **die Gesamtzahl der nicht lokalen Nachrichten bezeichnet, die von Knoten x weitergeleitet werden. Jedes Mal, wenn Knoten** x **eine nicht lokale Nachricht erfolgreich weiterleitet, gilt** Fx + **1.** Mx **stellt die Anzahl der Begegnungen dar, die Knoten** x **mit anderen Knoten hat. Jedes Mal, wenn Knoten** x **auf einen anderen Knoten trifft, gilt** Mx + **1.**

Wenn die Nachricht, die ein Knoten weiterleitet, erfolgreich an den Zielknoten zugestellt wurde , weist dieses Dokument dem Knoten je nach seiner Position im Weiterleitungspfad Beitragswerte zu, um eine aktive Teilnahme an der Netzwerkkommunikation zu fördern. Dem ersten Relaisknoten im Weiterleitungspfad wird der höchste Beitragswert zugewiesen, da er die Richtung der Nachrichtenübermittlung bestimmt. Darüber hinaus wird dem letzten Relaisknoten ein relativ hoher Beitragswert zugewiesen, da er die eigentliche Nachrichtenübermittlung durchführt und die Aufgabe erfüllt. Die anderen Relaisknoten im Weiterleitungspfad erhalten je nach ihrer Position im Weiterleitungspfad abnehmende Beitragswerte. Der **von Knoten** i im n-ten Pfad zugewiesene **Beitragswert** C(i, n) **wird wie folgt dargestellt:**

$$C(ich, n) = \begin{cases} (1 \ddot{y} r)^{ich \ddot{y}1} \ddot{y} r, & \textbf{wenn } 0 \ddot{y} i < L \\ (1 \ddot{y} r)^{ich \ddot{y}1}, & \textbf{wenn } i = L \end{cases}, \quad \textbf{(5)}$$

$$\text{DurchschnittCx} = \frac{\ddot{y}_{t=1}^{N} Ct(x, n)\, t}{}, \quad \textbf{(6)}$$

wobei r **(0** < r < 1) **den Zuteilungsfaktor für den Beitragswert bezeichnet und** L **die Gesamtzahl der Relaisknoten auf dem Pfad darstellt .** Nach erfolgreicher Zustellung einer Nachricht an den **Zielknoten wird der Beitragswert des Knotens** x **gemäß dem oben genannten Zuteilungsschema akkumuliert. Danach wird der durchschnittliche Beitragswert des Knotens x, avgCx , gemäß Gleichung (6) berechnet.**

Daher kann der Beteiligungsbeitrag PCx **des Knotens** x durch **die Gleichungen (5) und (6) wie folgt ausgedrückt werden :**

$$PCx(h) = \ddot{y} \ddot{y} Xx(h) + (1 \ddot{y} \ddot{y})\ddot{y} avgCx , \quad \textbf{(7)}$$

wobei $\ddot{y}$ **der Gewichtungsfaktor ist, mit dem** Xx(h) **und** avgCx Einfluss auf PCx **haben und** $\ddot{y}$ **ÿ [0,** 1].

3.1.4. Vermögensstatus der Knoten

Knoten im Netzwerk verfügen zunächst über eine bestimmte Menge an virtueller Währung, die in gewissem Maße ihre Weiterleitungsfähigkeit widerspiegelt. Basierend auf der Menge an virtueller Währung, die einem Knoten zur Verfügung steht, kann er in drei verschiedene virtuelle Währungszustände eingeteilt werden: ausreichender Zustand, allgemeiner Zustand und unzureichender Zustand. Wenn Vx(h) > Vmax ist, **befindet sich der Knot**

Machine Translated by Google

Wohlstandszustand, zu welchem Zeitpunkt Cx(h) > **1. Wenn** Vx(h) < Vmin, **befindet sich der Knoten in einem ungenügenden Wohlstandszustand, zu welchem Zeitpunkt** Cx(h) < **1. Wenn** Vmin **ÿ** Vx(h) **ÿ** Vmax, **befindet sich der Knoten in einem allgemeinen Wohlstandszustand, zu welchem Zeitpunkt** Cx(h) = **1. Die entsprechende Berechnung lautet wie fo**

$$Cx(h) = \begin{cases} \ddot{y}\ \dfrac{Vx(h)}{\mathbf{Vmax}}, & \textbf{Adäquater Zustand} \\ \mathbf{1}, & \textbf{Allgemeiner Zustand} \\ „\ \dfrac{Vx(h)}{\mathbf{Vmin}}, & \textbf{Unzureichender Zustand} \end{cases}. \quad \textbf{(8)}$$

wobei Vx(h) **die Menge an virtueller Währung darstellt, die dem Knoten** x **zum Zeitpunkt h gehört.**
Die Schwellenwerte für virtuelle Währungen Vmax **und** Vmin **stellen jeweils die maximalen und minimalen Mengen an virtueller Währung dar, die ein Knoten besitzen kann.**

3.1.5. Grad der sozialen Korrelation zwischen Knoten

In SAN beschreiben die sozialen Korrelationen zwischen Knoten das Ausmaß, in dem eine Art Verbindung oder gegenseitiger Einfluss zwischen Knoten besteht. Soziale Korrelation wird verwendet um die Interaktion und Bindung zwischen Knoten im Netzwerk widerzuspiegeln, indem die gemeinsamen Nachbarn und historischen Begegnungsaufzeichnungen zwischen Knoten analysiert werden. Daher bewertet dieses Dokument die soziale Korrelation zwischen Knoten umfassend anhand der Dimensionen der sozialen Überlappung von Knoten und des Kontaktgrads zwischen Knoten.

Der Grad der Überlappung zwischen zwei Knoten wird durch die Anzahl gemeinsamer Nachbarn quantifiziert, die sie im SAN haben. Diese Knoten mit gemeinsamen Nachbarn sind die Brücke, die zwei Knoten verbindet, und sie stellen die Ähnlichkeit zwischen Knoten dar. Somit wird die soziale Überlappung xy **zwischen Knoten** x **und Knoten** y **definiert als**

$$\text{Überlappung}_{xy} = \frac{\text{bequem}}{\text{Nachbar}_x + \text{Nachbar}_y \ddot{y} \text{Nachbar}_{xy}}, \quad \textbf{(9)}$$

Dabei bezeichnet comxy **die Anzahl der gemeinsamen Nachbarknoten zwischen den Knoten** x **und y;** neighx **und** neighy **stellen jeweils die Gesamtzahl der Nachbarn der Knoten** x **und y dar.**

Der Kontaktgrad zwischen Knoten quantifiziert die Intimität des Kontakts zwischen diesen Knoten durch die Analyse der Häufigkeit und Dauer von Begegnungen. Wenn zwei Knoten häufig miteinander kommunizieren, häufig und intensiv interagieren und der Kontaktgrad zwischen ihnen hoch ist, dann sind die beiden Knoten, wenn man die Definition eines Freundes in der realen Gesellschaft heranzieht, miteinander befreundet. Der Kontaktgrad contactxy **zwischen den Knoten** x **und** y **wird wie folgt berechnet:**

$$\text{Kontaktxy} = \frac{\text{Nmcxy } \ddot{y} \text{ Tmxy}}{\text{N } \ddot{y} \text{ genauT}}. \quad \textbf{(10)}$$

Wie oben erwähnt, bezeichnet Nmcxy **die Anzahl der Begegnungen zwischen den Knoten** x **und** y **bisher,** Tmxy **bezeichnet die Summe aller Begegnungszeiten zwischen den Knoten** x **und** y **bisher,** N **bezeichnet die Anzahl der Knoten im aktuellen Netzwerk und** accurentT **repräsentiert die aktuelle Moment.**

Basierend auf den Gleichungen (9) und (10) wird **die soziale Korrelation** SCxy **zwischen Knoten berechnet wie folgt:**

$$SCxy = \ddot{y}\ \ddot{y}\ \text{overlapxy} + (\mathbf{1}\ \ddot{y}\ \ddot{y})\ \ddot{y}\ \text{contactxy}, \quad \textbf{(11)}$$

wobei $\ddot{y}$ **der Gewichtungsfaktor ist, der den Einfluss von** overlapxy **und** contactxy **auf** SCxy bestimmt, **und** $\ddot{y}$ **ÿ [0,** 1].

3.2. Transaktionsmodell

In SAN wird die Nachrichtenweiterleitung zwischen Knotenpunkten als Transaktionsprozess abstrahiert , bei dem der Weiterleitungsdienst mit virtueller Währung ausgehandelt und gehandelt wird. In diesem Prozess fungiert der Knoten, der die Nachricht vor der Weiterleitung übermittelt, als Käufer, während der Knoten, der die Nachricht empfängt, als Verkäufer fungiert und der Nachrichtenweiterleitungsdienst a

Machine Translated by Google

die Ware. Beide Parteien sind verpflichtet, ihr Vermögen und ihre Ressourcen sowie ihre Angebote für die Speditionsdienstleistung auf der Grundlage ihrer jeweiligen Erwägungen offenzulegen.

3.2.1. Ausgabenfunktion des Käufers Um

sicherzustellen, dass die Nachricht reibungslos an den nächsten Knoten zugestellt werden kann, muss Käufer x **seine eigene Vermögensreserve, die Zustellungsdringlichkeit und die übertragenen Nachrichtenattribute berücksichtigen, um einen angemessenen Ausgabenpreis als Entschädigung für die Bereitstellung von Weiterleitungsdiensten an den nächsten Knoten zu bestimmen. Wenn die Vermögensreserve von x ausreicht, wird der Käufer seine eigene Zustellungsdringlichkeit und Vermögensreserve zu Wenn die Vermögensreserve von x nicht ausreicht, verkürzt sich mit der Zeit die Überlebenszeit der Nachricht allmählich, der verbleibende Cache und die Energie des Knotens werden allmählich aufgebraucht und die Dringlichkeit der Zustellung durch den Käufer wird höher. Um den Druck zu verringern, bietet der Käufer daher einen höheren Transaktionspreis an. Wenn der** Beteiligungsbeitrag **von x höher ist, was auf eine bessere Serviceleistung hinweist, wird** x **beim Kauf des Weiterleitungsdienstes davon profitieren. Wenn der Knoten eine größere Nachricht überträgt, steigen auch die für die Weiterleitung verbrauchte Energie und der Cache, und auch die Dringlichkeit der Zustellung durch x wird steigen, was wiederum den Gebotspreis für die Nachricht beeinflusst. Darüber hinaus muss** x seine eigene Knotenpositionskorrelation berücksichtigen . **Wenn seine Positionskorrelation am Zielknoten der Nachricht hoch ist, bedeutet dies,** dass die Zeit, die x **benötigt, um den Zielknoten zu erreichen, kürzer ist und er den Preis leicht erhöhen wird, um die Nachricht so schnell wie möglich weiterzuleiten. Zusammenfassend ist die Ausgabenfunktion von Käufer** x **zum Weiterleiten der Nachricht** m **zum Zeitpunkt** h **definiert als**

$$BNpx(h) = Msize \cdot (Q + \cdot_{1}^{H} \cdot g(x)) \cdot Cx(h) \cdot PCx(h). \quad \textbf{(x,m)} \qquad \textbf{(12)}$$

wobei Msize **die Größe von m bezeichnet,** g(x) **die Positionsrelevanz von x bezeichnet, ÿ1 das Gewicht ist, das die Positionsrelevanz eines Knotens einnimmt und** $\ddot{y} + \ddot{y} + \ddot{y} + \ddot{y}1 = 1$ **erfüllt,** Cx(h) den Reichtum von x **zum Zeitpunkt** h **bezeichnet und** PCx(h) **den Beteiligungsbeitrag des Knotens** x **zum Zeitpunkt h bezeichnet.**

3.2.2. Einkommensfunktion des Verkäufers

Beim Empfang von von x übermittelten Nachrichten muss **der Verkäufer** y **bei der Preisgestaltung auch seine eigenen Vermögensreserven, die Zustellungsdringlichkeit und die Eigenschaften der übermittelten Nachricht berücksichtigen, um seine Einnahmen zu maximieren. Wenn ys Vermögensreserven ausreichen, verringert sich die Bereitschaft des Knotens, Nachrichten weiterzuleiten, sodass diese mit der Zustellungsdringlichkeit von y** kombiniert werden muss , **um eine angemessene Preisgestaltung zu erzielen. Wenn ys Vermögensreserven nicht ausreichen, muss er virtuelle Währung woanders zu einem niedrigeren Preis beschaffen, und seine Bereitschaft, Nachrichten weiterzuleiten, steigt. Die Nachrichtenweiterleitung ist günstiger, wenn** y **einen höheren Beteiligungsbeitrag hat. Die Größe der Nachricht ist direkt proportional zu der dafür benötigten Energie und dem Cache; wenn die Nachricht größer ist, steigt auch der Energie- und Cacheverbrauch für** y , **das einen Weiterleitungsdienst bereitstellt, was wiederum die Preisge Wenn die Korrelation zwischen den Knoten hoch ist, d. h.** SCxy $\geq$ SCxymin, **dann zeigt dies an, dass die Beziehung zwischen zwei Knoten besser ist. Knoten** y **ist stärker bereit, die Nachricht weiterzuleiten , und der Gebotspreis wird relativ niedrig sein, wie in Gleichung (13) gezeigt. Im Gegenteil, wenn die Korrelation zwischen den Knoten niedrig ist, d. h.** SCxy < SCxymin, wird **der weiterleitende Knoten relativ hoch bieten, wie in Gleichung (14) gezeigt.**

$$SNpy(h) = Msize \cdot (Q \cdot \cdot_{2}^{H} \cdot (1 \cdot SCxy)) \cdot \cdot (1 \cdot PCy(h)) \cdot \frac{1}{Cy(h)}, \quad \textbf{(y,m)} \qquad \textbf{(13)}$$

$$SNpy(h) = Msize \cdot (Q \quad \textbf{(y,m)} \quad + \cdot_{2}^{H} \cdot SCxy) \cdot \cdot \frac{1}{(1 \cdot PCy(h))}. \qquad \textbf{(14)}$$

wobei Msize **die Größe von m bezeichnet,** SCxy **die soziale Korrelation zwischen** x **und y bezeichnet und ÿ2 den Gewichtungswert bezeichnet, den die soziale Korrelation zwischen** x **und y einnimmt, und zwar unter Erfüllu**

Machine Translated by Google

ÿ + *ÿ* + *ÿ* + **ÿ2** = **1.** SCxymin **bezeichnet die Mindestschwelle der sozialen Korrelation zwischen** x **und y,** Cy(h) **bezeichnet den Vermögensstatus von** y **zum Zeitpunkt h und** PCy(h) **bezeichnet den Beteiligungsbeitrag von** y **zum Zeitpunkt h.**

3.2.3. Strategie der erzwungenen

Weiterleitung Angesichts der begrenzten Energie- und Cacheressourcen, die den Knoten im Netzwerk zur Verfügung stehen, neigen sie dazu, der Weiterleitung ihrer eigenen Nachrichten den Vorzug zu geben oder die Weiterleitung nicht lokaler Nachrichten zu verweigern, um sicherzustellen, dass ihre generierten Nachrichten erfolgreich zugestellt werden. Dieses Verhalten schadet den Gesamtinteressen des Netzwerks erheblich. Um die kollektiven Um Knoten mit geringer Kooperationsbereitschaft zu motivieren, an Kooperationen teilzunehmen, stellt dieser Beitrag eine Strategie zur erzwungenen Weiterleitung vor, um das Verhalten jener Knoten einzuschränken, die nur an ihren eigenen Interessen interessiert sind und nicht bereit sind, im Netzwerk zu kooperieren. Diese Strategie fördert die Kooperation zwischen Knoten und optimiert die Netzwerkleistung.

Die Größe des erzwungenen Weiterleitungswerts CFx **von Knoten** x **wird durch den Kollaborationsgrad** Xx(h) **und den Beitragswert** C(i, n) **von Knoten x bestimmt, der wie folgt berechnet wird:**

$$CFx = \ddot{y}10 \ddot{y} (0,3 \ddot{y} Xx(h)) \ddot{y} C(i, n)\ddot{y}. \quad (15)$$

Durch die erzwungene automatische Aktualisierung des Weiterleitungswerts in einer festen Frequenz wird sichergestellt, dass das Weiterleitungsverhalten der Knoten zu bestimmten Zeitpunkten überwacht und reguliert werden kann . Wenn CFx > **0 ist, muss** Knoten x CFx- **Nachrichten weiterleiten, die von anderen Knoten generiert wurden, bevor er seine eigenen generierten Nachrichten weiterleiten kann. Wenn Knoten** x **in der Weiterleitung eingeschränkt ist, kann er, um seine eigenen Nachrichten reibungslos weiterzuleiten, den Grad der Zusammenarbeit nur verbessern, indem er an der Weiterleitung von Nachrichten teilnimmt, die von anderen Knoten generiert wurden, und dann seinen eigenen erzwungenen Weiterleitungswert schrittweise reduzieren, bis er schließlich die normale Nachrichtenweiterleitungsfunktion wiederherstellt .**

3.2.4. Transaktionsregeln

Wenn im SAN Käufer x **auf Verkäufer y trifft, müssen beide Parteien den Preis des Datenweiterleitungsdienstes nach bestimmten Regeln aushandeln. Die spezifischen Transaktionsregeln des Algorithmus lauten wie folgt:**

(1) Sowohl Käufer als auch Verkäufer werden vollständig über die Preis- und Vermögenslage der anderen Vertragspartei informiert.

(2) Nach Beginn der Transaktion ist das Vermögen des Käufers **der vom Käufer** x **als** BNpx(h) **angegebene Ausgabepreis und der vom Verkäufer** y **als** SNpy(h) angegebene Preis. **Wenn** BNpx(h) **ÿ** SNpy(h), **dann wurde Nachricht** m **erfolgreich weitergeleitet und die nächste Nachricht kann weitergeleitet werden. Wenn** BNpx(h) < SNpy(h), **schlägt die erste Verhandlungsweiterleitung fehl.**

(3) Wenn die erste Verhandlung zwischen Käufer und Verkäufer scheitert, beginnt die zweite Verhandlungsphase . Durch Vergleich des Vermögens beider Parteien werden die jeweiligen Preise angepasst. Der Käufer passt BNpx(h) **unter Berücksichtigung seines Kooperationsgrads, seines durchschnittlichen Beitrags und seiner Positionsrelevanz an, und der Verkäufer passt** SNpy(h) **basierend auf seinem Kooperationsgrad, seinem durchschnittlichen Beitrag und seiner sozialen Korrelation zwischen Knoten an. Die Anpassungsfaktoren für die zweite Verhandlung sind k1 bzw. k2 . Die zw**

- **Wenn** Cx(h) > Cy(h), **dann** bleibt SNpy(h) **unverändert und** BNpx(h) wird **gemäß der Gleichung** BNpx(h) = BNpx(h) **ÿ (1** + k1) **angepasst , wobei k1 wie folgt berechnet wird:**

$$k1 = \ddot{y}1 \ddot{y} Xx(h) + \ddot{y}2 \ddot{y} avgCx + \ddot{y}3 \ddot{y} g(x), \text{ wobei } Xx(h) \quad (16)$$

den Kooperationsgrad von x, avgCx **den durchschnittlichen Beitragsgrad von x und** g(x) **die Positionsrelevanz von Knoten x bezeichnet. Die Gewichte ÿ1, ÿ2 und ÿ3 bezeichnen jeweils den Einfluss des Kooperationsgrads, des durchschnittlichen Beitragsgrads und der Positionsrelevanz des Knotens auf den Anpassungsfaktor k1 , und die Summe der drei ergibt 1. Hier werden die Gewichte dynamisch mithilfe einer adaptiven Gewichtungsmethode angepasst, und der Wert des Gewichts wird mithilfe von Gleichung (17) berechnet:**

Machine Translated by Google

$$\begin{aligned} \text{ÿ ÿ1} &= \frac{\text{Xx(h)}}{\text{Xxh + Durchschnitt Cx + g (x)}} \\ \text{ÿÿ ÿ2} &= \frac{\text{DurchschnittCx}}{\text{Xxh + Durchschnitt Cx + g (x)}} \\ \text{ÿÿ ÿ3} &= \frac{\text{g(x)}}{\text{Xxh + Durchschnitt Cx + g (x)}} \end{aligned} \quad \textbf{(17)}$$

- **Wenn** Cy(h) **ÿ** Cx(h), **dann** bleibt BNpx(h) **unverändert und** SNpy(h) **wird angepasst gemäß die Formel** SNpy(h) = SNpy(h) **ÿ (1 ÿ** k2), **wobei k2 wie folgt berechnet wird:**

$$\textbf{k2} = \textbf{ÿ1 ÿ} \text{ Xy(h)} + \textbf{ÿ2 ÿ} \text{ avgCy} + \textbf{ÿ3 ÿ} \quad \textbf{(18)}$$

SCxy, **wobei** Xy(h) **den Kooperationsgrad von y,** avgCy **den durchschnittlichen Beitragsgrad von y und** SCxy **die soziale Korrelation zwischen den Knoten** x **und y bez**

Die Gewichte ÿ1, ÿ2 und ÿ3 repräsentieren den Einfluss des Grads der Zusammenarbeit, des Durchschnitts

Beitragsgrad und soziale Korrelation zwischen Knoten auf den Anpassungsfaktor k2,

und die Summe dieser drei ist 1. Die dynamische Anpassung dieser

Auch die Berechnung der Gewichte erfolgt nach der gleichen Methode wie in Gleichung (17).

(4) Wenn beide Knoten **nach der zweiten Verhandlung zu einem Konsens gelangen, dass** BNpx(h) **ÿ** SNpy(h) , wird die Nachricht wie in Schritt (2) weitergeleitet. Wenn nicht, wird die Transaktion abgebrochen.

und die nächste Transaktion wird fortgesetzt. Der Nachrichtentransaktionsfluss wird in Figur 2.

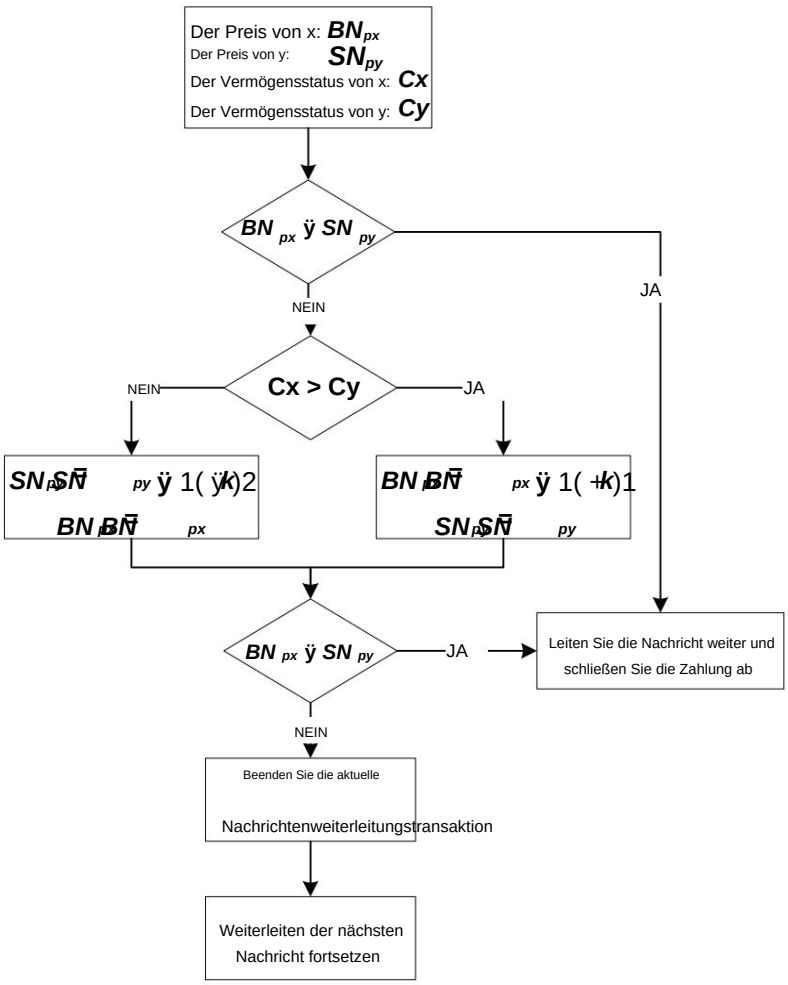


Abbildung 2. Flussdiagramm der Nachrichtentransaktion.

Der detaillierte Ablauf des normalen Nachrichtentransaktionsalgorithmus ist in Algorithmus 1 dargestellt , der den vollständigen Verhandlungsprozess für die Nachrichtenweiterleitu

Machine Translated by Google

Dienst, wenn sich zwei Knoten treffen, ohne dass eine erzwungene Weiterleitungsstrafe auftritt. Der detaillierte Weiterleitungsprozess ist wie folgt: Wenn sich die Knoten x **und** y treffen, **geben** x **und** y ihre **jeweiligen Erstgebote für die Weiterleitungsdienstleistung entsprechend den Gleichungen (12)–(14) . Wenn** BNpx **ÿ** SNpy, **dann wird die Nachricht weitergeleitet und der Transaktionspreis wird auf den Durchschnitt von** BNpx **und** SNpy. **Und dann wird der Transaktionsdatensatz in die Blockchain gepackt. Im Gegenteil, gemäß den zuvor beschriebenen Transaktionsregeln wird der Preis angepasst basierend auf dem Vermögensstatus von** x **und** y , **um ein neues Angebot zu machen. Wenn** BNpx1 **ÿ** SNpy1, **dann Die Hälfte der Summe aus** BNpx1 **und** SNpy1 **ist der Transaktionspreis. Und der Transaktionsdatensatz ist nach der erfolgreichen Nachrichtenweiterleitung in die Blockchain verpackt. Wenn die Verhandlung immer noch fehlschlägt, nachdem beide Parteien ihre Preise neu angepasst haben, geben sie die Weiterleitung dieser Nachricht auf m, und die nächste Nachricht wird zur Weiterleitung verarbeitet.**

Algorithmus 1 Normaler Nachrichtentransaktionsalgorithmus

Eingabe: BNpx, SNpy, Cx, Cy
Ausgabe: die erfolgreiche Weiterleitungsnachricht m

1: wenn BNpx >= SNpy **dann**
2: Kostenm = (BNpx + SNpy)/2
3: x sendet Nachricht m an y, y empfängt Nachricht m
4: wenn die Menge an Währung, die x derzeit besitzt, größer ist als Costm , **dann**
5: Transaktion erfolgreich. ÿ **NeuenBlock erstellen**
6: Der Block wird am Ende der Blockchain hinzugefügt.
7: Aktualisieren Sie den Reichtum von x und y
Rückkehr m
8: 9: Ende, wenn
10: sonst
11: Berechnen Sie die Kosten von m und die angepassten Gebote beider Parteien neu.
12: BNpx1 **= getPriceTwice_x(BNpx,SNpy** Cx, Cy)
13: SNpy1 **= getPriceTwice_y(BNpx,SNpy** Cx, Cy)
14: wenn BNpx1 >= SNpy1 **dann**
15: Kostenm = (BNpx1 + **SNpy1)/2**
16: x sendet Nachricht m an y, y empfängt Nachricht m
17: wenn der Betrag an Währung, den x derzeit besitzt, größer ist als Costm , **dann**
18: Transaktion erfolgreich. ÿ **NeuenBlock erstellen**
19: Der Block wird am Ende der Blockchain hinzugefügt.
20: Aktualisieren Sie den Reichtum von x und y
21: Rückkehr m
22: Ende, wenn
23: anders
24: Fahren Sie mit der nächsten Transaktion fort.
25: Ende, wenn
26: Ende, wenn

3.3. Spielprozess zur Nachrichtenweiterleitung

In SAN kann die Nachrichtenweiterleitung zwischen Knoten abstrahiert werden als dynamischer Spielprozess. Die Spielteilnehmer in diesem Prozess sind der Käufer x **und der Verkäufer y. Es gibt zwei Hauptphasen: ob der Käufer** x **eine Anfrage zur Nachrichtenweiterleitung sendet und ob der Verkäufer** y **zu diesem Zeitpunkt eine Anfrage zur Nachrichtenweiterleitung erhält. Das Verhalten von Beide Parteien in den beiden Phasen können entweder Kooperation oder Kooperationsverweigerung sein. Es ist zu erwartet, dass die rationalen Knoten im Netzwerk eine Strategie wählen, die ihre eigenen Vorteile. In diesem Artikel verwenden wir Rückwärtsinduktion, um die beiden Phasen des Spiels zu analysieren Prozess und zeigen Sie, dass eine wechselseitige Zusammenarbeit für beide Seiten die optimale Strategie ist.**

Angenommen, zum Zeitpunkt h **beträgt der Gebotspreis für** x für die Nachricht m BNpx(h) **und die Wenn der Gebotspreis von** y SNpy(h) **ist , dann ist der Transaktionspreis für die Nachricht** Costm = $\frac{BNpx(h)+SNpy(h)}{2}$.

Machine Translated by Google

(1) Analysieren Sie zunächst die zweite Phase des Spielprozesses. Ob der Verkäufer y die **Anfrage zur Nachrichtenweiterleitung erhält** oder nicht, und sein Verhalten bestimmt den Gewinn des Knotens. Wenn y **eine Anfrage zur Nachrichtenweiterleitung von x akzeptiert, also ein kooperatives Verhalten wählt, wird der aktuelle Gewinn** Ey **von** y **wie folgt berechnet:**

$$Ey = SNpy(h) \ddot{y} Kostenm. \quad (19)$$

Wenn y **sich weigert, eine Anfrage zur Nachrichtenweiterleitung von x zu empfangen, also ein kooperatives Verhalten ablehnt , gibt es keinen Gewinn. Da** y außerdem die Nachricht **bei der Begegnung** mit x **nicht akzeptiert hat ,** ist die Anzahl der von y **weitergeleiteten nichtlokalen Nachrichten geringer, während die Anzahl der Begegnungen zunimmt und der Grad der Zusammenarbeit von** y **abnimmt, was dazu führt, dass auch die Beteiligung am Beitrag abnimmt. Aus den Gleichungen (13) und (14) ist ersichtlich, dass ys nächstes Gebot für andere Nachrichten höher wird, was zum Scheitern der Transaktion führen kann. Im Gegensatz dazu wird** y bereit sein, Nachrichten zu empfangen.

(2) Als nächstes analysieren wir, ob der Käufer x **Nachrichtenweiterleitungsanfragen sendet oder nicht. Wenn** x **zum Zeitpunkt h Nachrichtenweiterleitungsanfragen sendet, also das kooperative Verhalten wählt,** dann **kann die Währung, die er spart ,** ausgedrückt werden als

$$Beispiel = BNpx(h) \ddot{y} Kostenm, \quad (20)$$

Wenn x **zum Zeitpunkt h keine Anfrage zur Nachrichtenweiterleitung sendet, also die Zusammenarbeit verweigert, analysieren Sie zunächst die Kosten, die** x **aus der Sicht der weiterzuleitenden Nachricht zu zahlen hat.**

Zum Zeitpunkt h + **1 ist die Überlebenszeit der Nachricht abgelaufen, und um m erfolgreich weiterzuleiten, können die zusätzlichen Kosten H1 , die** x **zahlen muss, wie folgt ausgedrückt werden:**

$$\mathbf{H1} = MGröße \ddot{y} Cx(h + 1) \ddot{y} PCx(h + 1) \ddot{y} \ddot{y}1 \ddot{y} \frac{Tre(h) \ddot{y} Tre(h + 1)}{TTL}, \quad (21)$$

wobei Tre(h + 1) **die verbleibende TTL von** m **zum Zeitpunkt** h + **1 bezeichnet und die Bedeutungen der anderen Symbole denen in Gleichung (12) entsprechen.**

Darüber hinaus werden die von x **zu zahlenden Kosten** im Hinblick auf **die Energie und den Cache von Käufer x analysiert . Da die Energie von** x **mit zunehmender Zeit strikt abnimmt und** x **die Zusammenarbeit verweigert, ist sein Cache auch teilweise durch die Nachricht m belegt . Wenn** x **also zum Zeitpunkt** h + 1 **eine Weiterleitungsanfrage senden möchte und sein vorgeschlagener Preis ebenfalls steigt, dann kann der zusätzliche Preis H2 , den** x **zahlen muss, wie folgt ausgedrückt werden:**

$$\mathbf{H2} = \ddot{y}2 \ddot{y} \left(\frac{Größe\ M}{Brx\ (h + 1)} + \frac{MGröße}{Erx\ (h + 1)} \right) \ddot{y} Lx \ddot{y} Cx(h + 1) \ddot{y} PCx(h + 1). \quad (22)$$

wobei Lx die Länge der von x **zum Zeitpunkt** h + **1 weiterzuleitenden Nachricht bezeichnet ,** Brx (h + 1) **das verbleibende Pufferverhältnis von** x **zum Zeitpunkt** h + **1 bezeichnet,** Erx (h + 1) **das verbleibende Energieverhältnis von** x **zum Zeitpunkt** h + **1 ist und die Bedeutungen der anderen Symbole denen in Gleichung (1**

Zusammenfassend sind die zusätzlichen Kosten H , die Käufer x **für die Ablehnung zu zahlen hat, Die Kooperation zum Zeitpunkt** h **ist mindestens**

$$H = \mathbf{min\{H1, H2\}}. \quad (23)$$

Offensichtlich ist Ex > **0** > ÿH, **also** muss x **sich dafür entscheiden, eine Anfrage zur Nachrichtenweiterleitung zu senden.**

Nach der obigen Analyse werden sich sowohl Käufer als auch Verkäufer für das kooperative Verhalten entscheiden, unter der Prämisse, dass beide den maximalen Nutzen daraus ziehen, d.ÿh. (Kooperation, Kooperation) ist die optimale Strategie.

3.4. Sicherheitsmodul **In**

diesem Dokument wird für die Gestaltung des Algorithmus eine auf virtueller Währung basierende Anreizstrategie verwendet. Daher ist es unvermeidlich, dass im Verlauf jeder Transaktion falsche Notierungen der Knoten auftreten, was die normale Reihenfolge der Transaktion ernsthaft untergräbt.

Machine Translated by Google

Um zu verhindern, dass Knoten den Währungsbetrag illegal manipulieren, verwendet dieses Dokument die Blockchain-Technologie, um ein Sicherheitsmodul zu erstellen, das offene und transparente Nachrichtenweiterleitungstransaktionen zwischen Knoten garantiert. Blockchain ist ein dezentrales digitales Hauptbuch für wirtschaftliche Transaktionen mit den Vorteilen von Transpar Im Netzwerk werden alle erfolgreichen Transaktionen in einer Blockchain-Struktur aufgezeichnet, normalerweise in sequenzieller Reihenfolge, und kryptografisch garantiert, dass sie nicht manipuliert und nicht fälschbar sind. Jeder Knoten in der Kette ist unabhängig und gleich und verlässt sich auf einen Konsensmechanismus, um die Datenkonsistenz sicherzustellen, mit den Merkmalen Dezentralisierung, Offenheit und Transparenz sowie Manipulationssicherheit [33].

3.4.1. Modulerstellung

Beim Speichern von Transaktionsdaten generiert jeder Knoten im Netzwerk mithilfe eines Public-Key-Verschlüsselungsalgorithmus (z. B. RSA) einen öffentlichen Schlüssel PK **und einen geheimen Schlüssel** SK . **Der** SK **wird dabei vom Knoten sicher aufbewahrt, während der** PK **im Netzwerk öffentlich verfügbar Um die Sicherheit des Transaktionsprozesses zu gewährleisten und falsche Notierungen zu vermeiden, führt dieses Dokument den folgenden Prozess für den zwischen dem Initiator** A **und dem Akzeptanten** B **einer Transaktion generierten Transaktionsdatensatz durch, der die Transaktionsparteien** A **und B, den Transaktionspreis, die Kosten**

usw. enthält: (1) Blockerstellung: Wenn die Nachrichtentransaktion zwischen A **und** B **erfolgreich abgeschlossen wurde ,** wird mit der Hash-Funktion (SHA-256) der Hash-Wert jedes Elements im **Transaktionsdatensatz berechnet .** Anschließend wird **A's** SK verwendet, um den Hash-Wert H(Kosten) **des Transaktionspreises zu signieren , wodurch** SigSK(H(Kosten)) generiert wird. **Anschließend werden der Transaktionsdatensatz, der Hash-Wert und die Signatur in einen neuen Block gepackt und an das gesamte Netzwerk gesendet, damit andere Knoten dies überprüfen können.**

(2) Verifizierung: B **kann A's** PK verwenden , **um die digitale Signatur zu verifizieren und den Hashwert Hÿ** (Kosten) **des Transaktionspreises neu zu berechnen. Stimmen diese beiden Hashwerte überein, ist der** Verifizierungsprozess erfolgreich, was darauf hinweist, dass die Transaktionsdaten **während der Übertragung preislich nicht manipuliert wurden .** Andere Knoten im Netzwerk **prüfen nach Erhalt des neuen Blocks, ob die Transaktion im Block gültig ist.**

(3) Konsenserreichung: Wenn eine ausreichende Anzahl von Knoten bestätigt, dass die Wenn die Transaktion gültig ist und ein Konsens erreicht wurde, gilt der Block als gültig.

(4) Blockchain-Update: Sobald ein neuer Block netzwerkweiten Konsens erlangt hat, wird er am Ende der Blockchain hinzugefügt, wie in Abbildung 3 dargestellt. Dieser Aktualisierungsprozess stellt Konsens und Konsistenz des Transaktionsverlaufs im gesamten Netzwerk

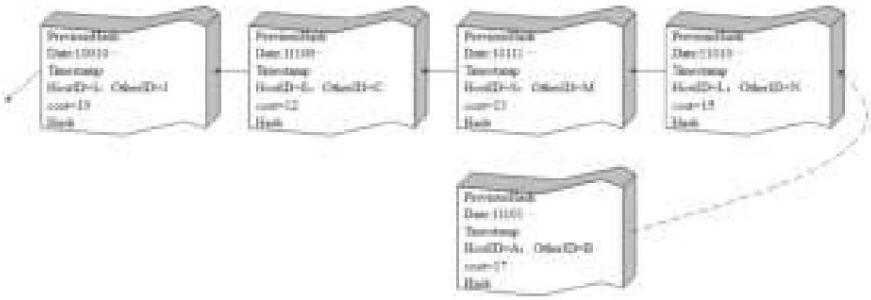


Abbildung 3. Schematische Darstellung der Blockchain.

Die Manipulation von Transaktionsdaten wird durch die Verwendung von Hash-Algorithmen in den durch den obigen Prozess erstellten Modulen sichergestellt. Jeder neue Block enthält den Hash-Wert des vorherigen Blocks, und jede Manipulation der Transaktionsdaten ändert seinen Hash-Wert, wodurch die Kontinuität der gesamten Blockchain gestört wird und unehrliches Verhalten von anderen Knoten im Netzwerk leicht erkannt und abgelehnt werden kann. Darüber hinaus sind die Transaktionsdaten transparent und vertrauenswürdig, da die verifizierten Transaktionsdatensätze in Blöcken geschrieben und der Blockchain hinzugefügt werden, was wiederum Transparenz und Vertrauenswürdigkeit der Transaktionsdaten erreicht. Jeder am Netzwerk teilnehmende Knoten kann auf diese Transaktionsdatensätze zugreifen und sie verifi

Machine Translated by Google

falsches Anführungszeichenverhalten. Alle Knoten folgen dem Prinzip der längsten Kette, was bedeutet, dass, wenn Es gibt mehrere konkurrierende Zweigketten, Knoten werden sich entscheiden, der längsten gültigen Kette beizutreten, da die Hauptkette, die die Robustheit und Sicherheit der Blockchain gewährleistet. Somit ist die Fairness und die Sicherheit des Nachrichtenübertragungsprozesses im Netzwerk wird durch die Verwendung gewährleistet der Blockchain.

3.4.2. Weiterleitungsschema

Der vollständige Prozess des Nachrichtenweiterleitungsdienstes in DFIA wird in Algorithmus 2 dargestellt. Angenommen, Knoten x **trifft Knoten y,** dann muss x die Nachricht m **an y weiterleiten ,** wobei CFx **die erzwungener Weiterleitungswert von x.** CFx **wird vor der Nachrichtenweiterleitungstransaktion beurteilt. Wenn** CFx > **0 ist,** tritt x **in die erzwungene Weiterleitungsphase ein; es wird gezwungen, nicht lokale Nachrichten weiterzuleiten. Für jede erfolgreiche Transaktion des Nachrichtenweiterleitungsdienstes wird der Wert von** CFx **wird um eins dekrementiert und umgekehrt; es wird mit der Weiterleitung der nächsten Nachricht fortgefahren. Wenn** CFx **ÿ 0 ist,** führt x **die normale Nachrichtenweiterleitungstransaktion gemäß Algorithmus 1. Während der Nachrichtenweiterleitungsdiensttransaktion wird der Weiterleitungsdienst Die Preise werden nach den Gleichungen (12)–(14) festgelegt, und die Verhandlungen werden nach spezifischen Transaktionsregeln. Wenn die Verhandlung erfolgreich ist und die Nachricht erfolgreich weitergeleitet wird, Die Transaktionsinformationen einschließlich der beiden Parteien** x **und y, des Transaktionspreises usw. werden gepackt in den Block und nach der Verifizierung zur Blockchain hinzugefügt. Andernfalls wird die Transaktion schlägt fehl und die nächste Nachricht wird weitergeleitet.**

Algorithmus 2 DFIA-Algorithmus

Eingabe: CFx, **x,** y
Ausgabe: die erfolgreiche Weiterleitungsnachricht m
wenn CFx > **0 dann**
Wenn der Quellknoten der Nachricht nicht der Knoten x ist , **dann**
wenn BNpx >= SNpy **dann**
Kostenm = (BNpx + SNpy)**/2**
Wenn der Betrag der Währung, die x derzeit besitzt, größer ist als Costm , **dann**
Transaktion erfolgreich. Der Block wird am Ende ÿ **NeuenBlock erstellen**
der Blockchain hinzugefügt.
Aktualisieren Sie den Reichtum von x und y
CFx = CFx **ÿ 1**
Rückkehr m
Ende, wenn
sonst
Berechnen Sie die Kosten von m und die angepassten Gebote beider Parteien neu.
BNpx1 **= getPriceTwice_x(BNpx,SNpy** Cx, Cy)
SNpy1 **= getPriceTwice_y(BNpx,SNpy** Cx, Cy)
wenn BNpx1 >= SNpy1 **dann**
Kostenm = (BNpx1 + **SNpy1)/2**
Wenn der Betrag der Währung, die x derzeit besitzt, größer ist als Costm , **dann**
Transaktion erfolgreich. Der Block wird am Ende der ÿ **NeuenBlock erstellen**
Blockchain hinzugefügt.
Aktualisieren Sie den Reichtum von x und y
CFx = CFx **ÿ 1**
Rückkehr m
Ende, wenn
sonst
Weiter mit der nächsten Transaktion
Ende, wenn
Ende, wenn
Ende, wenn
sonst
ÿ **Normaler Nachrichtentransaktionsalgorithmus**
Ende, wenn
Funktion CREATENEWBLOCK (HostX, andere, **m,** Kosten)
vorherigerHash **ÿ blockchain.getLastBlock().getHash()**
Zeitstempel **ÿ SimClock.getTime()**
Signieren Sie den Hash-Wert der Transaktionskosten mit dem privaten Schlüssel SK von A.
neuerBlock **ÿ neuer** Block(HostX, anderer, **m,** Kosten, vorherigerHash, Zeitstempel, **0)**
blockHash **ÿ neuerBlock.calculateHash()**
newBlock.setHash(blockHash)
neuen Block **zurückgeben**
Funktion beenden

Machine Translated by Google

4. Leistungsbewertung

4.1. Simulationsumgebung und Szenario

Die in diesem Kapitel beschriebenen Experimente wurden auf einem PC durchgeführt.
mit einer CPU von 3,60 GHz, 16,0 GB RAM und dem Betriebssystem Windows 11
der Opportunistic Network Environment Simulator [34] (ONE) durch die IntelliJ IDEA
2021.2.1 Softwarebetrieb. In den Experimenten dieser Studie wurde Spray and Wait Routing verwendet
als zugrunde liegender Routing-Algorithmus und verglichen DFIA mit RSIA in [26], STIA in [25],
und EEIS in [24] zur Simulation. Die folgenden fünf Simulationsszenarien wurden entworfen:

(1) S + egoistische Knoten;

(2) S + egoistische Knoten + DFIA;

(3) S + egoistische Knoten + RSIA;

(4) S + egoistische Knoten + STIA.

(5) S + egoistische Knoten + EEIS.

Die spezifischen experimentellen Parametereinstellungen sind in Tabelle 1 aufgeführt.

Tabelle 1. Parametereinstellungen für Simulationsexperimente.

Parameter	Wert
Simulationszeit	12 h
Anzahl der Knoten	180
Simulationsbereich	4500 m × 3400 m
Übertragungsgeschwindigkeit	250 kB/s
Sendereichweite	10 m
Bewegungsmodell	KürzesterPfadKartenbasierteBewegung
Routing-Protokoll	Sprühen und warten
Puffergröße	10 M
Nachrichten-TTL	6 h
Energie übertragen	4 Watt
Anfangsenergie	2000 J
Nachrichtenerstellungsintervall	25 s–35 s
Nachrichtengröße	500 kB–1 MB
$\ddot{y}$	0,1
ÿ, ÿ, ÿ1 , ÿ2	0,3
$\ddot{y}$	0,6
$\ddot{y}$	0,45

4.2. Leistungsparameter

Die Simulationsexperimente verwenden hauptsächlich die folgenden drei Parameter als Bewertungs-
tionskriterien zur Beurteilung der Wirksamkeit und Praktikabilität des Anreizalgorithmus.

Wahrscheinlichkeit der Nachrichtenzustellung: Dies ist das Verhältnis der Gesamtzahl der
erfolgreich an den Zielknoten zugestellten Nachrichten zur Gesamtzahl der erstellten Nachrichten.
Eine höhere Nachrichtenübermittlungsrate weist auf eine höhere Erfolgsrate des Knotens bei der Weiterleitung der
Nachricht an den Zielknoten. Die Wahrscheinlichkeit der Nachrichtenübermittlung ist ein äußerst wichtiger
Parameter zur Bewertung der Netzwerkleistung.

Durchschnittliche Latenz: Dies ist die durchschnittliche Zeit, die eine Nachricht von ihrem Beginn an benötigt
bis zur erfolgreichen Weiterleitung zum Zielknoten. Eine geringere Latenz bedeutet weniger Netzwerk
Ressourcen werden bei der Nachrichtenweiterleitung verbraucht und die Nachricht benötigt weniger
Zeit, um das Ziel zu erreichen.

Netzwerk-Overhead: Dieser wird dargestellt durch das Verhältnis der Differenz zwischen
Anzahl der von Knoten im Netzwerk weitergeleiteten Nachrichten und die Anzahl der erfolgreich
weitergeleiteten Nachrichten auf die Anzahl der erfolgreich weitergeleiteten Nachrichten. Eine Reduzierung der
Der Netzwerk-Overhead ist ein Hinweis auf eine Reduzierung des Energieverbrauchs des Knotens während
Nachrichtenweiterleitung.

Machine Translated by Google

4.3. Simulationsergebnisse

In dieser Studie wurden Evaluierungsexperimente in Bezug auf unterschiedliche Prozentsätze egoistischer Knoten und unterschiedliche TTLs entworfen. In der ersten Evaluierung beobachteten wir die experimentellen Ergebnisse, indem wir den Prozentsatz egoistischer Knoten auf 0 %, 20 %, 40 %, 60 %, 80 % un Die Abbildungen 4–6 veranschaulichen die vergleichende Leistung jedes Algorithmus in Bezug auf Nachrichtenübermittlungswahrscheinlichkeit , durchschnittliche Latenz und Netzwerk-Overhead in drei Dimensionen, wenn in den oben genannten fünf Szenarien unterschiedliche Prozentsätze egoistischer Knoten vorhanden sind. Bei der zweiten Bewertung wurde der Prozentsatz egoistischer Knoten auf 60 % und die Nachrichten-TTL auf 2 h, 4 h, 6 h, 8 h bzw. 10 h festgelegt. Die Abbildungen 7–9 zeigen die Änderungen der Nachrichtenübermittlungswahrscheinlichkeit , der durchschnittlichen Latenz und des Netzwerk-Overheads jedes Algorithmus für die fünf oben genannten Szenarien unter unterschiedlichen Nachrichten-TTLs. Für jeden Satz von Bewertungsexperimenten wurden insgesamt 30 Versuche durchgeführt , und der Mittelwert der Daten aus diesen Versuchen wurde als endgültiges Versuchsergebnis verw

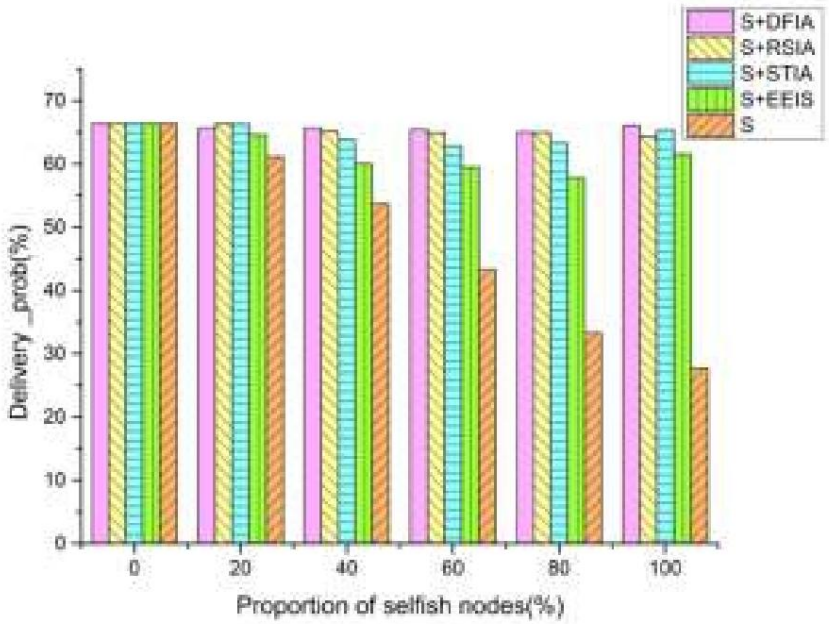


Abbildung 4. Wahrscheinlichkeit der Nachrichtenübermittlung mit unterschiedlichen Prozentsätzen egoistischer Knoten.

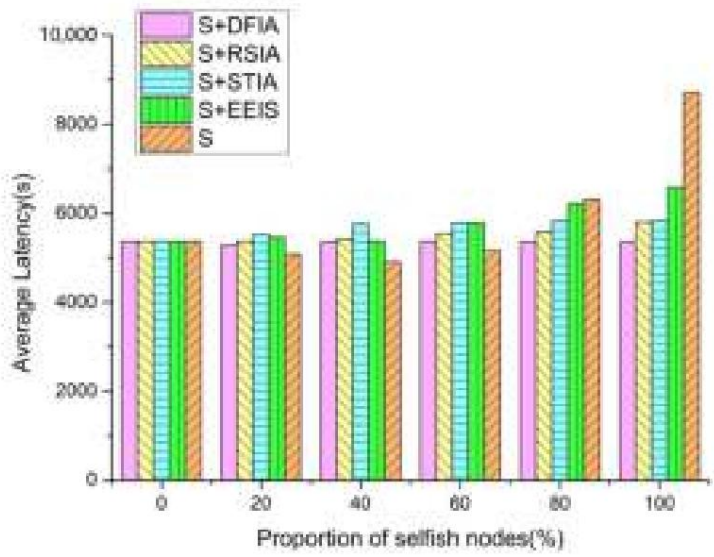


Abbildung 5. Durchschnittliche Latenz mit unterschiedlichen Prozentsätzen egoistischer Knoten.

Machine Translated by Google

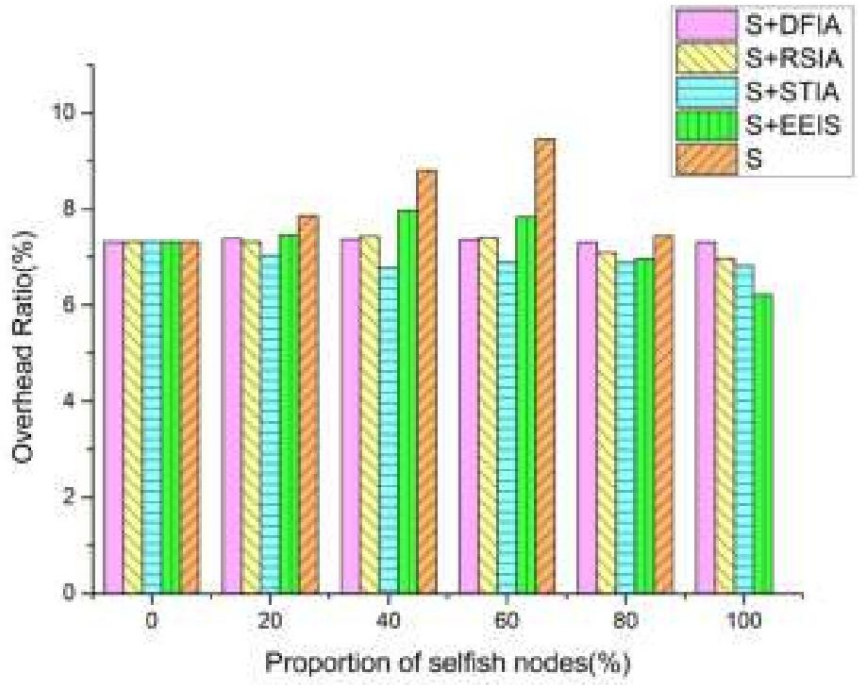


Abbildung 6. Netzwerk-Overhead für unterschiedliche Prozentsätze egoistischer Knoten.

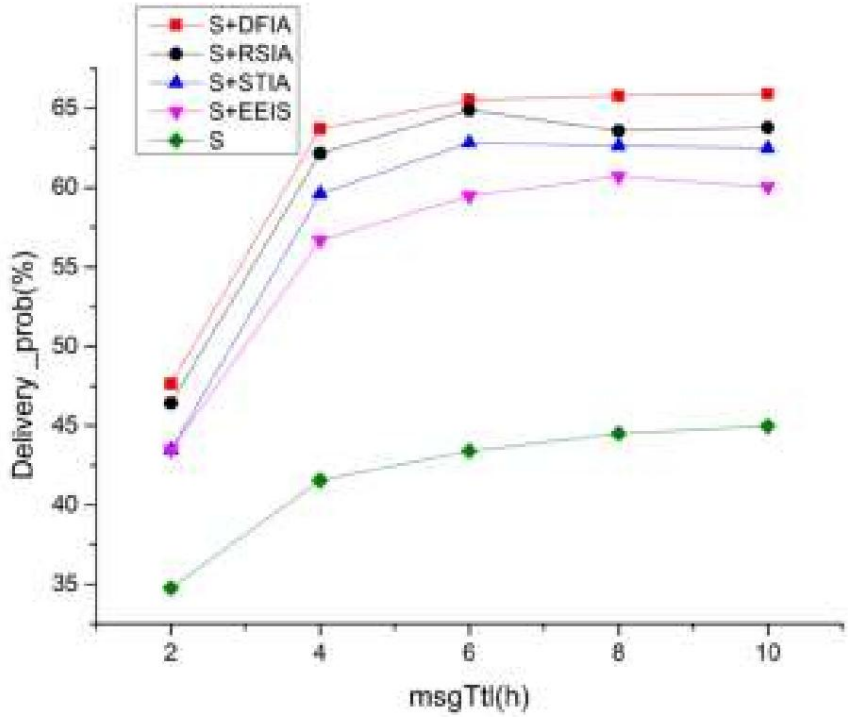


Abbildung 7. Wahrscheinlichkeit der Nachrichtenzustellung bei unterschiedlichen Nachrichten-TTLs.

Machine Translated by Google

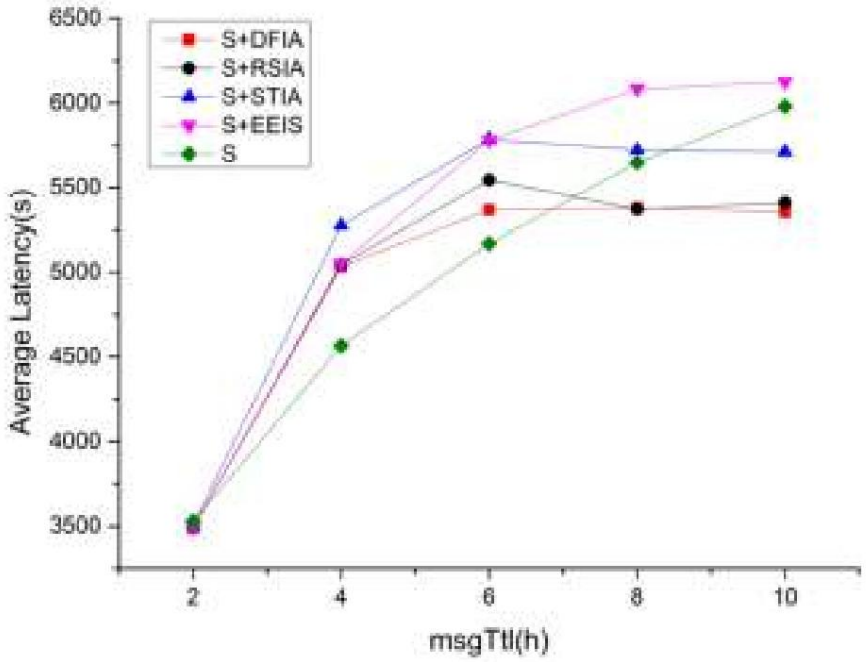


Abbildung 8. Durchschnittliche Latenz bei verschiedenen Nachrichten-TTLs.

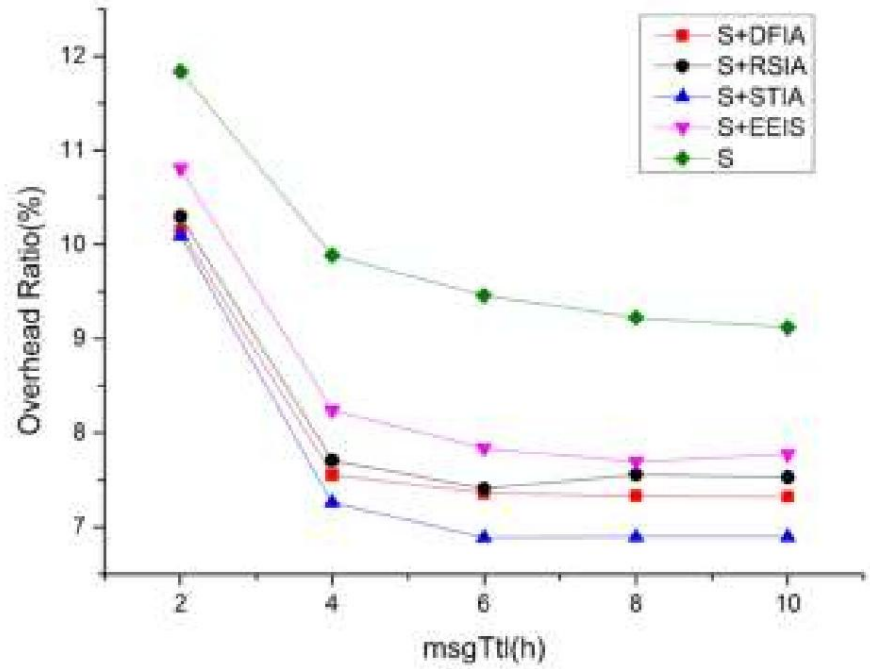


Abbildung 9. Netzwerk-Overhead bei unterschiedlichen Nachrichten-TTLs.

4.3.1. Unterschiedliche Prozentsätze für egoistische Knoten

Wie in Abbildung 4 dargestellt, ist die Wahrscheinlichkeit der Nachrichtenübermittlung für die fünf Szenarien mit unterschiedlichen Prozentsätzen egoistischer Knoten dargestellt. Der Rückgang der Prozentsätze der Nachrichtenübermittlung kann darauf zurückgeführt werden, dass egoistische Knoten nicht am kooperativen Weiterleitungsverhalten teilnehmen, was die Erfolgsrate der Nachrichten, die den Zielknoten erreichen, verringert . Der allgemeine Trend zeigt, dass das Hinzufügen egoistischer Knoten dazu führt, dass DFIA, RSIA und STIA im Fall von Knoten mit egoistischem Verhalten als Anreiz wirken. DFIA ist jedoch im Vergleich zu RSIA und STIA effektiver. S weist einen ausgeprägteren Trend zur Verringerung der Nachrichtenübermittlungsrate auf, da es keine Anreize verwendet. RSIA konzentriert sich auf die Ressourcenbeschränkungen und das mögliche Vorh

Machine Translated by Google

und die möglichen Auswirkungen von bevorzugtem Egoismus auf das Netzwerk. STIA berücksichtigt bei Entscheidungen über das Weiterleitungsverhalten von Knoten die Qualität des Weiterleitungsdienstes der Knoten sowie die möglichen Auswirkungen kurzfristiger Währungsknappheit. EEIS verwendet die Preisgestaltung für Knotenressourcen und übernimmt den Modus des gleichwertigen Austauschs, um Transakti Sie berücksichtigen jedoch nicht den Einfluss der Standortkorrelation zwischen Knoten und den Einfluss starker und schwacher sozialer Korrelationen zwischen Knoten. DFIA berücksichtigt umfassend die Merkmale des Knotenstandorts und der sozialen Korrelationen. Obwohl die Wahrscheinlichkeit der Nachrichtenübermittlung 0,68 % niedriger ist als bei RSIA, wenn 20 % egoistische Knoten vorhanden sind, ist die Zustellungswahrscheinlichkeit von DFIA mit der kontinuierlichen Zunahme egoistischer Knoten höher als bei anderen Algorithmen. Basierend auf der Auswertung der Gesamtwirkung des Experiments zeigt DFIA die optimale Leistung in Bezug auf die Wahrscheinlichkeit der Nachrichtenübermittlung.

Abbildung 5 zeigt die durchschnittliche Latenz für die fünf Szenarien mit unterschiedlichen Anteilen egoistischer Knoten. Die Erhöhung der durchschnittlichen Verzögerung bedeutet, dass das Hinzufügen egoistischer Knoten zu einer längeren Zeit für die Nachrichtenweiterleitung zum Ziel führt. Die in Abbildung 5 dargestellten Daten zeigen, dass mit der Erhöhung des Anteils egoistischer Knoten von 0 % auf 100 % die durchschnittliche Latenz in allen Szenarien mit unterschiedlichem Grad der Erhöhung zunimmt. S weist die stärkste Erhöhung der durchschnittlichen Latenz auf, insbesondere wenn der Anteil egoistischer Knoten 60 % erreicht; sein durchschnittlicher Latenzwert steigt dramatisch an, da keine Anreizmaßnahmen ergriffen werden. In den verbleibenden vier Szenarien ist die Erhöhung der durchschnittlichen Latenz relativ gleichmäßig und scheint nicht stark anzusteigen. Unter ihnen ist die durchschnittliche Latenz von DFIA die stabilste und ihre durchschnittliche Latenz ist niedriger als die von RSIA, STIA und EEIS. Dies deutet darauf hin, dass die Einführung einer DFIA-Anreizstrategie dazu beiträgt, die Auswirkungen egoistischer Knoten zu mildern, und dass sie gleichzeitig die Zustellungswahrscheinlichkeit der Nachricht verbessert und die durchschnittliche Latenz verringert.

Abbildung 6 zeigt den Netzwerk-Overhead in den fünf Szenarien bei unterschiedlichen Prozentsätzen egoistischer Knoten. Aus den Daten in der Abbildung ist ersichtlich, dass mit zunehmendem Anteil egoistischer Knoten der Netzwerk-Overhead der anderen vier Szenarien, mit Ausnahme von S ohne Anreize, im Allgemeinen abnimmt und sich stabiler verändert. Im S-Szenario nimmt der Netzwerk-Overhead stark ab, wenn der Anteil egoistischer Knoten 60 % übersteigt, und er sinkt direkt auf 0, wenn der Anteil egoistischer Knoten im Netzwerk 100 % beträgt , da es immer mehr egoistische Knoten im Netzwerk gibt und die Knoten die Nachrichten nicht weiterleiten oder die Nachrichten ohne Anreizstrategien erfolgreich an den Zielknoten übermitteln können . Wenn der Anteil egoistischer Knoten auf 60 % steigt, ist der Gesamtnetzwerk-Overhead von DFIA im Vergleich zu RSIA, STIA, EEIS und S ohne Anreize hoch. Dies liegt daran, dass DFIA die Blockchain-Technologie durch die Verwendung von RSA zur Verschlüsselung der Transaktionsdatensätze einführt, was zu einem höheren Gesamtnetzwerk-Overhead auf Kosten eines höheren Netzwerk-Overheads bei der Erhöhung der Nachrichtenübermittlungswahrscheinlichkeit und der Gewährleistung der Transaktionssicherheit führt.

4.3.2. Verschiedene TTLs

Abbildung 7 veranschaulicht die Nachrichtenübermittlungswahrscheinlichkeit für die oben genannten Szenarien unter verschiedenen Nachrichten-TTLs. Es ist offensichtlich, dass die Nachrichtenübermittlungswahrscheinlichkeit mit jeder Erhöhung der Nachrichten-TTL zunimmt. Die Erhöhung der Nachrichten-TTL bedeutet, dass die Knoten beim Weiterleiten von Nachrichten nicht durch zu viele zeitliche Einschränkungen eingeschränkt werden. Es bleibt genügend Zeit, um die Nachrichten weiterzuleiten, wodurch das Verlieren von Nachrichten und die daraus resultierende geringere Leistung des Netzwerks vermieden werden können. Im Vergleich zu den anderen vier Algorithmen ist die Nachrichtenübermittlungswahrscheinlichkeit von DFIA eindeutig optimal, was darauf zurückzuführen ist, dass DFIA die Faktoren berücksichtigt, die in vielen Aspekten zum Egoismus der Knoten führen, was wiederum die egoistischen Knoten motiviert, an der Zusammenarbeit teilzunehmen. Im Gegensatz dazu berücksichtigen RSIA, STIA und EEIS die Auswirkungen der Standortrelevanz von Knoten auf Knoten nicht und unterscheiden offensichtlich nicht zwischen der Stärke der sozialen Korrelation, was zu weniger wirksamen Anreizen führt. Es ist ersichtlich, dass DFIA effektiv verhindern kann, dass

Machine Translated by Google

Abbildung 8 veranschaulicht die durchschnittliche Latenz der vier Szenarien unter verschiedenen Nachrichten -TTLs. Es ist klar, dass die durchschnittliche Latenz für alle Szenarien mit zunehmender Nachrichten-TTL zunimmt. Eine Erhöhung der Nachrichten-TTL bedeutet, dass die Nachrichten mit größerer Wahrscheinlichkeit über eine größere Entfernung im Netzwerk zugestellt werden, was wiederum zu einer Erhöhung der Anzahl von Paketen in den Knotenüberlastungen bei der Datenweiterleitung und einer daraus resultierenden Erhöhung der Latenz des Netzwerks führt. Die durchschnittliche Verzögerung dieser fünf Szenarien weist mit zunehmender Nachrichten-TTL einen deutlichen Anstiegstendenz auf, und die maximale Erhöhung der du
Im Vergleich dazu ist bei DFIA, RSIA, STIA und EEIS ein geringerer Anstieg der durchschnittlichen Latenz zu verzeichnen.
Unter diesen ist die durchschnittliche Latenz von DFIA im Vergleich zu RSIA, STIA und EEIS geringer. Das liegt daran, dass STIA zwar die Knotenressourcen und die Servicequalität berücksichtigt , aber den sozialen Korrelationsfaktor der Knoten ignoriert, wodurch die durchschnittliche Latenz der Nachricht bis zu einem gewissen Grad steigt. RSIA berücksichtigt nur die egoistischen Gründe der Knoten, berücksichtigt aber nicht die Positionsbeziehungen zwischen ihnen und unterscheidet zwischen starken und schwachen sozialen Korrelationen, was zu einer höheren durchschnittlichen Latenz führt. EEIS wendet quadratische Verhandlungen an und berücksichtigt nicht, dass die Nachricht **aufgrund von Paketverlusten der Knoten nicht rechtzeitig übermittelt werden kann ,** was die durchschnittliche Latenz beeinflusst. DFIA berücksichtigt die **Positionsbeziehungen zwischen den Knoten und führt den sozialen Korrelationsfaktor ein, wodurch die Zustellrate der Nachricht verbessert und die durchschnittliche Latenz reduziert wird.**

Abbildung 9 veranschaulicht den Vergleich des Netzwerk-Overheads. Mit zunehmender Nachrichten-TTL nimmt der Netzwerk-Overhead in allen Szenarien allmählich ab. Da S keine **Anreizmaßnahmen ergreift ,** ist sein Netzwerk-Overhead höher als der von DFIA, RSIA, STIA und EEIS.
STIA berücksichtigt die Servicequalität der Knoten und reduziert die Nachrichteninteraktionen während des Verhandlungsprozesses, wodurch die Begegnungszeit für die Nachrichtenübermittlung optimiert und der Netzwerk-Overhead reduziert wird. EEIS berücksichtigt nur die Preisgestaltung der Knotenressourcen und verwendet den Kreditwert für die Geldleihe und berücksichtigt nicht das Paketverlustverhalten, das durch den Rückgang der Transaktionskapazität aufgrund übermäßiger Kreditaufnahme und Nichtrückzahlung verursacht wird, was zu einem hohen Netzwerk-Overhead führt. DFIA berücksichtigt Knotenressourcen, Standortattribute und soziale Korrelationen für das Verhandlungsspiel, wodurch sein Anreizeffekt bedeutender ist als der von RSIA, und es hat auch einen geringeren Netzwerk-Overhead.
Aufgrund der Berücksichtigung mehrerer Faktoren während der Verhandlungen in DFIA werden mehr Informationsinteraktionen generiert und der Netzwerk-Overhead ist im Vergleich zu STIA höher. Aber DFIA erhöht die Wahrscheinlichkeit der Nachrichtenübermittlung und reduziert die durchschnittliche Latenz, und der allgemeine Anreizeffekt ist besser als der von STIA.

5. Schlussfolgerungen

In diesem Artikel schlagen wir einen ressourcenbasierten dynamischen Preis- und Zwangsweiterleitungsanreizalgorithmus (DFIA) vor. DFIA verhandelt den von Knoten benötigten Weiterleitungsdienst, indem es die durch den Ressourcenstatus des Knotens und die Standortrelevanz bestimmte Dringlichkeit der Zustellung integriert und zwischen starken und schwachen sozialen Korrelationen unterscheidet. Während des Transaktionsprozesses soll die Zwangsweiterleitungsstrategie den Grad der Zusammenarbeit der Knoten verbessern. Um ein falsches Angebotsverhalten der Knoten zu verhindern, verwendet DFIA Blockchain-Technologie zur Verschlüsselung, um Transparenz und Glaubwürdigkeit der Transaktionsdaten zu erreichen . Die Ergebnisse von Simulationsexperimenten zeigen, dass DFIA die durchschnittliche Latenz reduziert und gleichzeitig die Wahrscheinlichkeit der Nachrichtenzustellung erhöht und egoistische Knoten effektiv zur Teilnahme an der Zusammenarbeit anregt. Der Algorithmus verwendet jedoch RSA, um die Transaktionsdatensätze zu verschlüsseln, was zu einem höheren Gesamtnetzwerk-Overhead führt, der in zukünftigen Arbeiten angegangen und optimiert werden muss . Als Nächstes werden wir uns mit den Sicherheitsmaßnahmen befassen, die für mögliche böswillige Verhaltensweisen wie Knoten, die miteinander kollaborieren und **Nachrichten in einem SAN manipulieren, ergriffen werd**

Machine Translated by Google

Beiträge der Autoren: Konzeptualisierung, XZ, YL (Yuan Li) und ZX; Methodik, YL (Yuan Li); Software, YL (Yuan Li); Validierung, ZX, YL (Yuan Li) und XZ; Ressourcen, XZ, YL (Yuan Li) und ZX; Datenkuratierung, YL (Yuan Li); Schreiben – Vorbereitung des Originalentwurfs, YL (Yuan Li) und DH; Schreiben – Überprüfung und Bearbeitung, ZX, YL (Yanchao Liu) und SW; Projektverwaltung, ZX. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Forschung wurde von der National Natural Science Foundation of China (Zuschussnummer Nr. 61972136) und der Outstanding Youth Scientific Innovation Team Support Foundation des Bildungsministeriums der Provinz Hubei (Zuschussnummern Nr. T201410, T2020017) finanziert.

Datenverfügbarkeitserklärung: Die Daten sind im Artikel enthalten.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Verweise

1. **Wang, Z.; Liao, J.; Cao, Q.; Qi, H.; Wang, Z. Erreichen der k-Barriere-Abdeckung in hybriden Richtungssensornetzwerken.** IEEE Trans. 2014 , 13, **1443–1455. [CrossRef]**
2. **Tsugawa, S.; Ohsaki, H. Gemeinschaftsstruktur und Interaktionslokalität in sozialen Netzwerken.** J. Inf. Process. **2015,** 23, **402–410. [Querverweis]**
3. **Ning, Z.; Liu, L.; Xia, F.; Jedari, B.; Lee, I.; Zhang, W. CAIS: Ein anpassbares Anreizsystem für gemeindebasierte, sozial Bewusste Vernetzung.** IEEE Trans. Veh. Technol. **2017,** 66, **3406–3419. [CrossRef]**
4. **Xu, F.; Xu, Q.; Xiong, Z.; Xiao, N.; Xie, Y.; Deng, M.; Hao, H. Intelligentes verteiltes Routing-Schema basierend auf sozialer Ähnlichkeit für mobile soziale Netzwerke.** Future Gener. Comput. Syst. **2019,** 96, **472–480. [CrossRef]**
5. **Zhu, K.; Li, W.; Fu, X. SMART: Eine soziale und mobile Routing-Strategie für störungstolerante Netzwerke.** IEEE Trans. Veh. Technol. **2014,** 63, **3423–3434 .**
6. **Annalisa, S.; Antonio, C.; Floriano, DR; Peppino, F. Routing in mobilen opportunistischen sozialen Netzwerken mit egoistischen Knoten.** Wire Kommun. Mob. Comput. **2019,** 2019, **6359806.**
7. **Xia, F.; Liu, L.; Jedari, B.; Das, SK PIS: Ein mehrdimensionales Routing-Protokoll für Socially Aware Networking.** IEEE Trans. Mob. 2016 , 15, **2825–2836. [CrossRef]**
8. **Galati, A.** Delay Tolerant Network; **LAP Lambert Academic Publishing: Köln, Deutschland, 2010.**
9. **Luciana, P.; Andrea, P.; Marco, C. Opportunistische Vernetzung: Datenweiterleitung in getrennten mobilen Ad-hoc-Netzwerken.** IEEE Commun. Mag. **2006,** 44, **134–141.**
10. **Xia, F.; Liu, L.; Li, J.; Mohammed, AA; Tianruo, YL; Jianhua, M. BEEINFO: Interessenbasierte Weiterleitung mit künstlichen Bienen Kolonie für sozial bewusstes Networking.** IEEE Trans. Veh. Technol. **2015,** 64, **1188–1200. [CrossRef]**
11. **Wu, J.; Yu, G.; Guan, P. Methode zur Vorhersage der Wahrscheinlichkeit von Interessenmerkmalen in sozialen opportunistischen Netzwerken.** IEEE Access **2019, 7, 59002–59012. [CrossRef]**
12. **Elnaz, N.; Nahideh, D. FTR: Features Tree-basiertes Routing in mobilen sozialen Netzwerken.** Wirel. Netw. **2020,** 26, **283–291.**
13. **Mao, Y.; Zhou, C.; Qi, J.; Zhu, X. Ein fairer, kreditbasierter Anreizmechanismus für das Routing in DTN-basierten Sensornetzwerken mit Knoten Egoismus.** EURASIP J. Wirel. Commun. Netw. **2020,** 2020, **232. [CrossRef]**
14. **Loudari, SE; Benamar, N. Auswirkungen von Egoismus auf den Energieverbrauch in opportunistischen Netzwerken: Eine Leistungsbewertung . In Proceedings der 2019 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS), Fez, Marokko, 3.–4. April 2019; S. 1–7.**
15. **Li, Q.; Gao, W.; Zhu, S.; Cao, G. Ein Routing-Protokoll für sozial egoistische, verzögerungstolerante Netzwerke.** Ad Hoc Netw. **2012,** 10, **1619–1632. [Querverweis]**
16. **Zhang, X.; Deng, H.; Xiong, Z.; Liu, Y.; Rao, Y.; Lyu, Y.; Li, Y.; Hou, D.; Li, Y. Sichere Routing-Strategie basierend auf attributbasiertem Vertrauensbasierte Zugriffskontrolle in sozialen Netzwerken.** J. Signal Process. Syst. **2024,** 96, **153–168. [CrossRef]**
17. **Chang, J.; Pang, Z.; Xu, W.; Wang, H.; Yin, G. Ein anreizkompatibler Reputationsmechanismus für P2P-Systeme.** J. Supercomput. **2014,** 69, **1382–1409. [CrossRef]**
18. **Roy, A.; Acharya, T.; Das Bit, S. Sozialbasierte, reputationsbewusste Datenweiterleitung für verbesserte Multicast-Zustellung in Gegenwart von egoistischen Knoten in DTNs.** Int. J. Commun. Syst. **2020,** 33, **e4235. [CrossRef]**
19. **Bounouni, M.; Bouallouche-Medjkoune, L. Bestätigungsbasiertes Bestrafungs- und Stimulationsschema für mobile Ad-hoc -Netzwerke.** J. Supercomput. **2018,** 74, **5373–5398. [CrossRef]**
20. **Wang, EK; Chen, CM; Yiu, SM; Hassan, MM; Alrubaian, M.; Fortino, G. Anreiz-Evolutionsspielmodell für opportunistische soziale Netzwerke.** Future Gen. Comput. Syst. **2020,** 102, **14–29. [CrossRef]**
21. **Wu, J.; Mo, W.; Yin, X.; Bai, G. Forschung zu opportunistischen Routing-Anreizmechanismen basierend auf wiederholtem Spiel in WMNs.** Comput. Eng. Appl. **2018,** 54, **5.**
22. **Haq, A.; Faheem, Y. Ein auf Peer-to-Peer-Kommunikation basierendes Inhaltsverteilungsprotokoll für anreizbewusste, verzögerungstolerante Netzwerke.** Wirel. Netw. **2020,** 26, **583–601. [CrossRef]**
23. **Li, L.; Zhong, X.; Jiang, Y. IAR-GT: Ein anreizbewusstes Routing basierend auf der Spieltheorie für egoistische opportunistische Netzwerke.** KSII Trans. Internet Inf. Syst. **2019,** 13, **152–171.**

Machine Translated by Google

24. Xiong, Z.; Xiao, N.; Xu, F.; Zhang, X.; Xu, Q.; Zhang, K.; Ye, C. Ein gleichwertiges, auf Austausch basierendes Anreizsystem für die Datenweiterleitung für sozial bewusste Netzwerke. J. Signal Process. Syst. **2021,** 93, **249–263.** [CrossRef]

25. Xiong, Z.; Li, X.; Zhang, X.; Zhu, S.; Xu, F.; Zhao, X.; Wu, Y.; Zeng, M. Ein auf Servicepreisen basierender zweistufiger Anreizalgorithmus für sozial bewusste Netzwerke. J. Signal Process. Syst. **2022,** 94, **1227–1242.**

26. Zhang, X.; Rao, Y.; Xiong, Z.; Deng, H.; Xu, F.; Yuan, L. Ressourcenbeschränkter und auf sozialem Egoismus basierender Anreizalgorithmus für sozial bewusste Netzwerke. J. Signal Process. Syst. **2023,** 95, **1439–1453.**

27. Spyropoulos, T.; Psounis, K.; Raghavendra, CS Spray and wait: Ein effizientes Routing-Schema für zeitweise verbundene Mobilfunknetze . In Proceedings des 2005 ACM SIGCOMM Workshop on Delay-Tolerant Networking, New York, NY, USA, 26. August 2005; Seiten 252–259.

28. Cui, J.; Liu, S.; Chang, Y.; Liu, Q.; Wu, Q. Effizienter Routing-Algorithmus basierend auf virtuellen Währungstransaktionen in DTN. Comput. Sci. **2024,** 51, **300–308.**

29. Li, X.; Jiang, P.; Chen, T.; Luo, X.; Wen, Q. Eine Untersuchung zur Sicherheit von Blockchain-Systemen. Future Gener. Comput. Syst. **2020,** 107, **841–853.** [CrossRef]

30. Deng, M.; Lyu, Y.; Yang, C.; Xu, F.; Ahmed, M.; Yang, N.; Xu, Z.; Ke, C. Leichtgewichtiges Vertrauensmanagement-Schema basierend auf Blockchain in ressourcenbeschränkten intelligenten IoT-Systemen. IEEE Internet Things J. **2024,** 11, **25706–25719.** [CrossRef]

31. Matsumoto, S.; Reischuk, RM IKP: Eine PKI mit dezentralen automatisierten Anreizen umkrempeln. In Proceedings of the 2017 IEEE-Symposium für Sicherheit und Datenschutz (SP), San Jose, CA, USA, 22.–26. Mai 2017; S. 410–426.

32. He, Y.; Li, H.; Cheng, X.; Liu, Y.; Yang, C.; Sun, L. Ein Blockchain-basierter wahrheitsgetreuer Anreizmechanismus für verteilte P2P-Anwendungen. IEEE Access **2018, 6, 27324–27335.** [CrossRef]

33. Zhang, W.; Wang, H.; Hua, H.; Wang, Q. Egoistisches Mining und Verteidigungsstrategien im Bitcoin. Intell. Autom. Soft Comput. **2022,** 34, **1861–1875.** [CrossRef]

34. Keränen, A.; Ott, J.; Kärkkäinen, T. Der ONE-Simulator zur DTN-Protokollauswertung. In Proceedings der 2. Internationalen Konferenz zu Simulationswerkzeugen und -techniken, Rom, Italien, 2.–6. März 2009; ICST: Brüssel, Belgien, 2010; S. 1–10.

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

FPGA-Implementierung der säulenbasierten Objektklassifizierung für Autonomer mobiler Roboter

Chaewoon-Park [1], **Seongjoo Lee** [2,3] **und Yunho Jung** [1,4,*]

[1] Fakultät für Elektronik und Informationstechnik, Korea Aerospace University, Goyang-si 10540, Republik Korea; pcw0201@kau.kr
[2] Fakultät für Elektrotechnik, Sejong-Universität, Seoul 05006, Republik Korea; seongjoo@sejong.ac.kr
[3] Fakultät für Konvergenztechnik für intelligente Drohnen, Sejong-Universität, Seoul 05006, Republik Korea
[4] Fakultät für intelligente Luftmobilität, Korea Aerospace University, Goyang-si 10540, Republik Korea * Korrespondenz: yjung@kau.ac.kr; Tel.: +82-2-300-0133

Zusammenfassung: Mit dem Fortschritt der künstlichen Intelligenz werden autonome mobile Roboter in verschiedenen Anwendungen eingesetzt. In Szenarien des autonomen Fahrens ist die Objektklassifizierung für die Roboternavigation von entscheidender Bedeutung. Um diese Aufgabe zu erfüllen, werden Lichterkennungs- und Entfernungsmesssensoren (LiDAR), die Tiefen- und Höheninformationen erfassen und eine höhere Auflösung als Funkerkennungs- und Entfernungsmesssensoren (Radar) haben, Kamerasensoren vorgezogen. Die säulenbasierte Methode verwendet einen Pillar Feature Encoder (PFE), um 3D-LiDAR-Punktwolken in 2D-Bilder zu kodieren, wodurch eine Hochgeschwindigkeitsinferenz mithilfe von 2D-Convolutional Neural Networks ermöglicht wird. Obwohl die säulenbasierte Methode eingesetzt wird, um die Echtzeitreaktionsfähigkeit autonomer Fahrsysteme sicherzustellen, wird nicht aktiv an der Beschleunigung des PFE geforscht, obwohl der PFE eine erhebliche Menge an Rechenzeit innerhalb des Systems verbraucht. Daher schlägt dieses Papier einen PFE-Hardwarebeschleuniger und ein säulenbasiertes Objektklassifizierungsmodell für autonome mobile Roboter vor. Das vorgeschlagene Objektklassifizierungsmodell wurde mithilfe von 2971 Datensätzen aus acht Klassen trainiert und getestet und erreichte eine Klassifizierungsgenauigkeit von 94,3 %. Der PFE-Hardwarebeschleuniger wurde in einem Field Programmable Gate Array (FPGA) durch ein Register-Transfer-Level-Design implementiert, das eine 40-fache Beschleunigung im Vergleich zur Firmware für die ARM Cortex-A53-Mikroprozessoreinheit erreichte; das Objektklassifizierungsnetzwerk wurde mithilfe des FINN-Frameworks im FPGA implementiert. Durch die Integration des PFE und des Objektklassifizierungsnetzwerks haben wir ein säulenbasiertes Echtzeit-Obje

Schlüsselwörter: feldprogrammierbares Gate-Array; autonomer mobiler Roboter; Objektklassifizierung; Säulenmerkmal- Encoder; FINN

check for updates

Zitat: Park, C.; Lee, S.; Jung, Y. FPGA-Implementierung der säulenbasierten Objektklassifizierung für autonome Mobiler Roboter. Elektronik **2024,** 13, **3035. https://doi.org/10.3390/Elektronik13153035**

Wissenschaftliche Redaktion: Nicola Lusardi

Empfangen: 14. Juni 2024
Überarbeitet: 29. Juli 2024
Akzeptiert: 30. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

Autonome mobile Roboter bewegen sich, indem sie ihre Umgebung erkennen und physische Aufgaben ausführen, die früher von Menschen erledigt wurden. Fortschritte in den Bereichen Computer Vision, Robotersteuerung und künstliche Intelligenz (KI) haben die Fähigkeit autonomer mobiler Roboter verbessert, ihre Umgebung wahrzunehmen und darauf zu reagieren. Folglich werden diese Roboter in mehreren Bereichen des menschlichen Lebens eingesetzt, darunter in der Fertigung, im Transportwesen, im Gesundheitswesen und im täglichen Leben [1,2]. Die Objektklassifizierung ist für autonome mobile Roboter von entscheidender Bedeutung, um Objekte zu identifizieren und Fahrwege zu planen [3].

Zur Erfassung der Umgebung in autonomen Fahrsituationen werden verschiedene Sensortypen wie Kameras und Radarsensoren eingesetzt . **Kamerasensoren können keine Tiefeninformationen erfassen, wodurch die Entfernung zu einem Objekt nicht genau bestimmt werden kann. Dies kann zu schwerwiegenden Problemen führen, wenn sich die beiden Objekte überlappen [4]. Darüber hinaus sind Kamerasensoren helligkeitsempfind**

Machine Translated by Google

Sichtfeld, was zu betrieblichen Einschränkungen führt [3]. Radarsensoren können die Höheninformationen von Objekten nicht bestimmen und haben eine sehr niedrige Winkelauflösung, was eine präzise Beobachtung von Objekten verhindert **[5]. Im Gegensatz dazu enthalten von Light Detection and Ranging (LiDAR)-Sensoren erfasste Punktwolken dreidimensionale physikalische Koordinateninformationen, die die Erfassung von Tiefen- und Höheninformationen zu Objekten ermöglichen. Darüber hinaus gibt es weniger betriebliche Einschränkungen und die hohe Auflösung ermöglicht eine präzise Beobachtung von Objekten [3]. Daher wird die Verwendung von von LiDAR-Sensoren erfassten Punktwolken für das autonome Fahren bevorzugt.**

Für die Verarbeitung von Punktwolken stehen verschiedene Deep-Learning-Algorithmen zur Verfügung. MVCNN [6] projiziert eine Punktwolke in mehrere Richtungen und klassifiziert Objekte mithilfe eines Convolutional Neural Network (CNN) auf Grundlage der projizierten Bilder. Obwohl es eine einfache Methode zur Umwandlung von 3D-Daten in 2D-Bilder verwendet und ein CNN nutzt, kommt es während des **Projektionsprozesses zu Informationsverlusten [7]. PointNet [8] ist ein permutationsinvariantes Netzwerk, das mithilfe einer Symmetriefunktion zur Verarbeitung ungeordneter Punktmengen entwickelt wurde. VoxelNet [9] kodiert Merkmale mithilfe von PointNet Voxel für Voxel und führt Inferenzen mithilfe eines Region Proposal Network durch. Da die Inferenz auf einem 3D-CNN basiert, ist eine lange Inferenzzeit von 225 ms erforderlich [10]. SECOND [11] erreichte eine verbesserte Inferenzgeschwindigkeit von 50 ms durch Anwendung eines Sparse-CNN auf VoxelNet, die Komplexität der 3D-CNN-Berechnung blieb jedoch hoch. PointPillars [12] schlug eine säulenbasierte Methode zur Kodierung von 3D-Punktwolken in 2D-Bilder vor. Im Gegensatz zu MVCNN, das nur Informationen vom höchsten Element verwendet, verwendet die säulenbasierte Methode Informationen von allen Punkten innerhalb der Säule. Durch die Durchführung einer 2D-CNN-basierten Inferenz auf einem kodierten Bild können säulenbasierte M**

Autonome mobile Roboter müssen in Echtzeit reagieren können, um Unfälle zu vermeiden. Unter diesen Einschränkungen ist die Ausführung von säulenbasierten Methoden auf Basis von Mikroprozessoreinheiten (MPU) extrem zeitaufwändig [17]. Grafikprozessoren (GPUs) können Inferenzen auf Basis von 2D-CNNs beschleunigen; die Verwendung von GPUs auf Edge-Geräten ist jedoch im Hinblick auf den Stromverbrauch ineffizient **[18]. Daher wurde in den Referenzen [19–21] ein Hardwarebeschleuniger für säulenbasierte Methoden unter Verwendung eines feldprogrammierbaren Gate-Arrays (FPGA) implementiert.** In **[19] wurde jedoch ein Säulenmerkmalsencoder (PFE) auf einer MPU berechnet , der 71,2 ms benötigte und 19 % der gesamten Reaktionszeit ausmachte. Ebenso wurde in [20] ein PFE auf einer MPU in 9,51 ms** berechnet , **was 22 % der gesamten Reaktionszeit ausmachte.** Zusätzlich wurde in **[21] ein Teil des PFE beschleunigt , wobei für die Berechnung 37,7 ms benötigt wurden, was 59 % der gesamten Reaktionszeit ausmachte . Da das PFE einen erheblichen Teil der gesamten Reaktionszeit der säulenbasier**

Die säulenbasierte Methode kann Inferenzen auf ein durch den PFE codiertes Bild mit einem Netzwerk basierend auf einem tiefenweise trennbaren CNN (DSCNN) durchführen. Das in [22–24] verwendete DSCNN **wurde aufgrund seiner geringen Rechenkosten und der Anzahl der Parameter in neueren CNN-Modellen weithin übernommen. Aufgrund dieser Vorteile werden DSCNNs häufig in ressourcenbeschränkten oder latenzempfindlichen Anwendungen** verwendet **[25]. Darüber hinaus haben sie trotz ihrer geringen Komplexität eine erfolgreiche Leistung in Bildklassifizierungsmodellen gezeigt, indem sie die Darstellungseffizienz** verbesserten **[26].**

In dieser Studie schlagen wir ein effizientes, säulenbasiertes Objektklassifizierungsmodell für autonome mobile Roboter unter Verwendung eines PFE und DSCNN vor, **das durch Training und Validierung mit unter Verwendung von Ouster OS1-32 erfassten Datensätzen eine Klassifizierungsgenauigkeit von 94,3 % erreicht [27]. Darüber hinaus schlagen wir ein FPGA-basiertes Beschleunigungssystem für den Echtzeitbetrieb des vorgeschlagenen Objektklassifizierungsmodells vor. Das vorgeschlagene Beschleunigungssystem umfasst einen PFE-Beschleuniger und einen Klassifizierungsnetzwerkbeschleuniger. Wir haben den PFE-Beschleuniger so konzipiert, dass er eine Rechenzeit von 0,23 ms erreicht, und haben eine 40-fache Beschleunigungsleistung im Vergleich zur MPU bestätigt. Ein Klassifizierungsnetzwerkbeschleuniger wurde unter Verwendung von FINN auf einem FPGA implementiert. Für den Betrieb in Edge-Umgebungen haben wir das vorgeschlagene Beschleunigu**

Der Rest dieses Dokuments ist wie folgt strukturiert. Abschnitt 2 bietet einen Überblick über das vorgeschlagene Beschleunigungssystem, einschließlich der PFE- und DSCNN-basierten Klassi

Machine Translated by Google

Netzwerke, verwendeter Datensatz, Modellkonfiguration und Evaluierungsergebnisse. Das Design des vorgeschlagenen Beschleunigungssystems wird in Abschnitt 3 beschrieben. Abschnitt 4 stellt die Implementieru Ergebnisse und Beschleunigungseffekte. Abschließend fasst Abschnitt 5 die Studie zusammen.

2. Vorgeschlagenes System

Abbildung 1 zeigt einen Überblick über das vorgeschlagene Beschleunigungssystem. Die Punktwolke Die Eingaben in das System werden vom PFE in drei Schritten in ein Bild kodiert: (a) Konvertieren die Punktwolke in eine Reihe von Säuleneinheiten, (b) Hinzufügen von handgefertigten Features (HCFs) und (c) Erweiterung der Merkmale durch punktweise (PW) Faltung, Komprimierung durch eine Max-Pool-Operation und streuen sie zurück an ihre ursprüngliche Position auf der x–y- Ebene, um ein Bild erzeugen. Das erzeugte Bild wird dann in ein Klassifizierungsnetzwerk eingespeist, um zu klassifizieren die Objekte. Das Klassifizierungsnetzwerk besteht aus einem DSCNN-basierten Netzwerk.

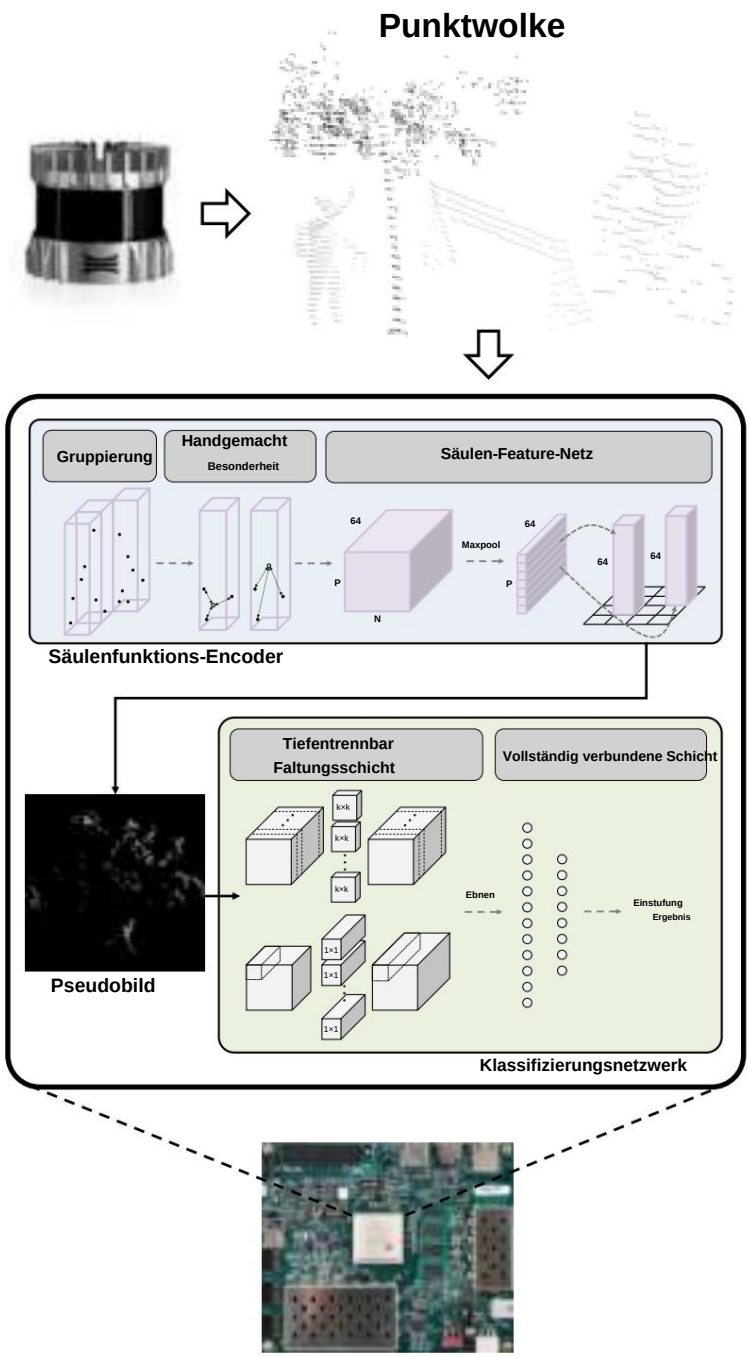


Abbildung 1. Übersicht über das vorgeschlagene Beschleunigungssystem.

Machine Translated by Google

2.1. PFE

Der PFE, der 3D-Punktwolken in 2D-Bilder kodiert, um 2D-CNN-basierte Inferenz zu ermöglichen, wurde in [12] vorgeschlagen. Zunächst wird durch Diskretisierung der **x–y** -Ebene **ein Gitter erzeugt , und dann wird die Punktwolke in eine Menge von Säuleneinheiten umgewandelt. Die Punktwolkenmenge g** wird **durch Gleichung (1) definiert .**

$$\mathbf{g} = \mathbf{[p1 , p2 , . . .} \text{ pi , . . . pP } \mathbf{]} \quad \textbf{(1)}$$

Wir setzen g als eine Menge von Pi , **wobei** P **die maximale Anzahl von Säulen ist, die Punkte enthalten. Wir setzen** Pi wie in Gleichung **(2) definiert .**

$$\text{pi} = \mathbf{[n1, n2, . . .} \text{ ni , . . . nN]} \quad \textbf{(2)}$$

Wir setzen pi **als eine Menge von** nj , **wobei** N **die maximale Anzahl von Punkten ist, die eine Säule enthalten kann. Die Menge** nj **ist wie in Gleichung (3) definiert.**

$$\text{ni} = [\text{x, } \mathbf{y,} \text{ z,r}] \quad \textbf{(3)}$$

Der Satz nj **umfasst die x-, y-,** z- **Koordinaten und Reflektivitätswerte der Punkte. Schließlich wird die Punktwolke durch einen Tensor der Größe** (P, **N, 4) dargestellt, wobei** P **und** N Hyperparameter **sind . Wenn die Anzahl der Punkte in einer Säule größer oder gleich N ist, wird eine Zufallsstichprobe verwendet; wenn sie kleiner als N ist, wird Null-Padding verwendet.**

Nachdem die Punktwolke in einen Satz von Säuleneinheiten umgewandelt wurde, wird dem Satz nj ein HCF hinzugefügt . **Der HCF besteht aus den Differenzen** xm, ym , zm **zwischen den arithmetischen Mittelkoordinaten der Punkte innerhalb der Säule und den x-, y-,** z- **Koordinaten jedes Punkts sowie den Differenzen** xc, yc **zwischen den Mittelpunktskoordinaten der Säule und den x-,** y- **Koordinaten jedes Punkts. Die Einstellung** nj **mit dem angehängten HCF ergibt sich aus Gleichung (4).**

$$\text{ni} = [\text{x, } \mathbf{y,} \text{ z, r, xm, ym, zm, xc, yc}] \quad \textbf{(4)}$$

Ein Tensor der Größe (P, **N, 9) mit einem hinzugefügten HCF wird einer PW-Faltung, Batch-Normalisierung und einer ReLU-Schicht unterzogen, um ihn in einen Tensor der Größe** (P, **N, 64) umzuwandeln. Anschließend wird Maximum Pooling entlang der zweiten Dimension angewendet, um einen Tensor d Schließlich wird jede Säule zurück in die ursprüngliche x–y -Ebene gestreut, wodurch die 2D-Bilddaten mit 64 Kanälen vervollständigt werden.**

2.2. Klassifizierungsnetzwerk

Das vorgeschlagene Klassifizierungsnetzwerk basiert auf dem DSCNN. Abbildung 2 zeigt, wie das DSCNN funktioniert. Das DSCNN umfasst die tiefenweisen (DW) und PW-Methoden. Bei DW sind die Anzahl der Kanäle im Eingabebild, die Anzahl der Filter und die Anzahl der Kanäle im Ausgabebild gleich. Jeder Filter hat nur einen Kanal und führt eine Faltungsoperation an nur einem Kanal im Eingabebild durch, um ein einkanaliges Ausgabebild zu erstellen. Beispielsweise wird **bei der in Abbildung 2 gezeigten DW-Faltung an einem Kanal des Eingabebilds, der durch eine gestrichelte Linie getrennt ist, eine Faltungsoperation mit einem Filter durchgeführt, um einen Kanal des Ausgabebilds zu erzeugen, der ebenfalls durch eine gestrichelte Linie getrennt ist. Lassen Sie** iC **die Anzahl der Kanäle im Eingabebild und** oC **die Anzahl der Kanäle im Ausgabebild sein; PW verwendet** oC **Filter der Größe 1 × 1 mit** iC **Kanälen. Jeder Filter führt eine Faltungsoperation an allen Kanälen im Eingabebild durch, um einen Kanal des Ausgabebilds zu erstellen. Beispielsweise vervollständigt bei der in Abbildung 2 dargestellten PW-Faltung ein Pixel des Eingabebildes, das durch eine gestrichelte Linie getrennt ist, ein Pixel des Ausgabebildes, das durch eine gestrichelte Linie getrennt ist, indem alle Filt**

Machine Translated by Google

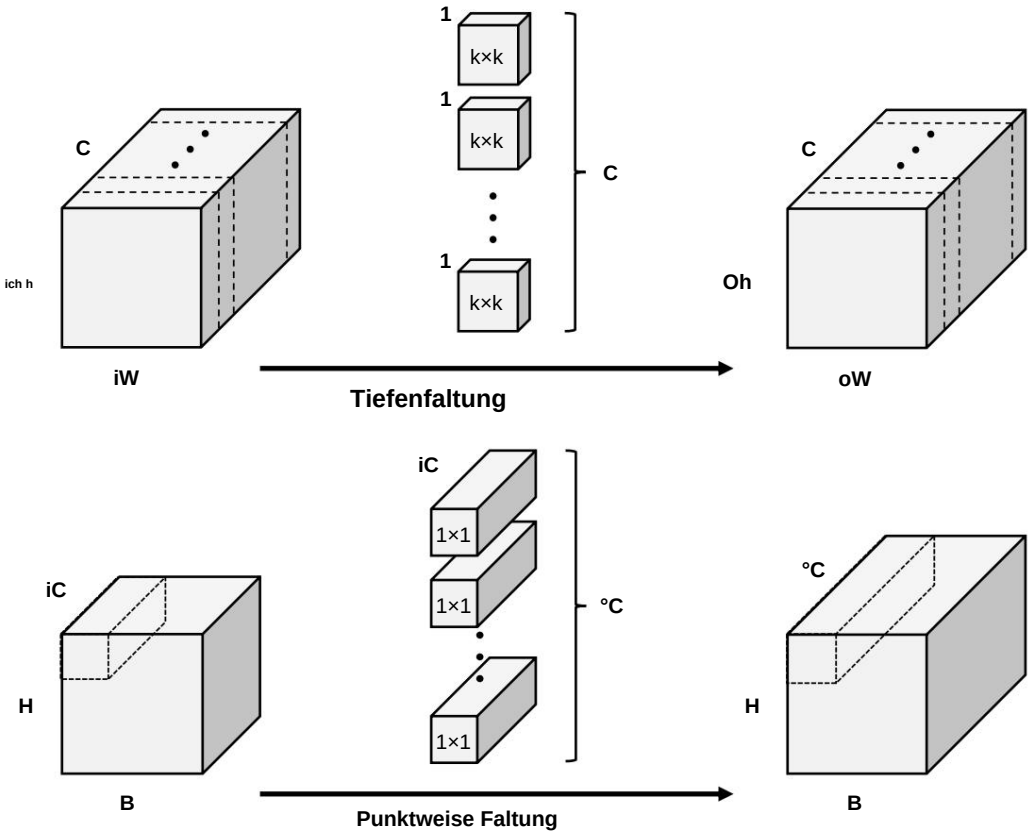


Abbildung 2. Funktionsmechanismus von DSCNN.

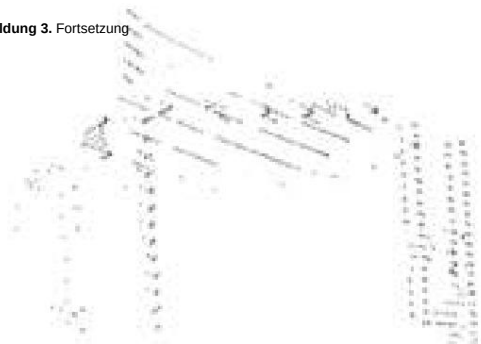
2.3. Datensatz

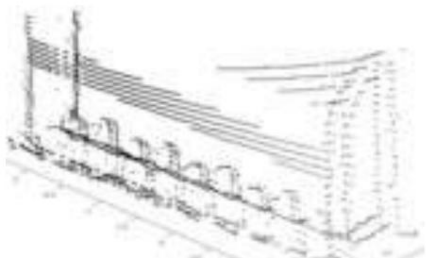
Um das vorgeschlagene Objektklassifizierungsmodell zu trainieren, haben wir Punktwolken mithilfe eines Ouster OS1-32-Sensors erzeugt. Abbildung 3 zeigt Beispiele aus einem Datensatz mit Objekten, denen autonome mobile Roboter typischerweise in ihrer Fahrumgebung begegnen. Um die Trainingsqualität zu verbessern, haben wir den Datensatz mit Beispielen aus dem nuScenes- Datensatz **[28]** erweitert . **Der Datensatz enthält Klassen wie Gebäude, Baum, Fahrzeug, Fahrrad, Hindernis, Grünfläche , Person und städtische Einrichtung (Straßenlaterne, Fahrradständer, Ampel usw.), wie in Abbildung 4 dargestellt, die die Konfiguration des Datensatzes zeigt. Von insgesamt 2971 Datenpunkten wurden 2617 für das Training**





Abbildung 3. Fortsetzung





Machine Translated by Google

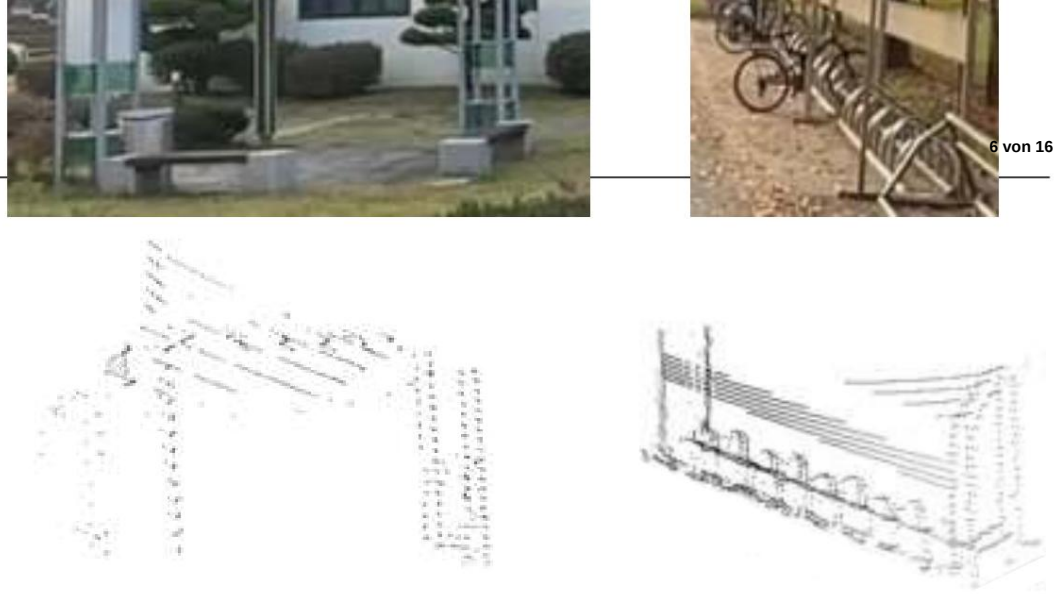


Abbildung 3. Beispiele für Datensätze.

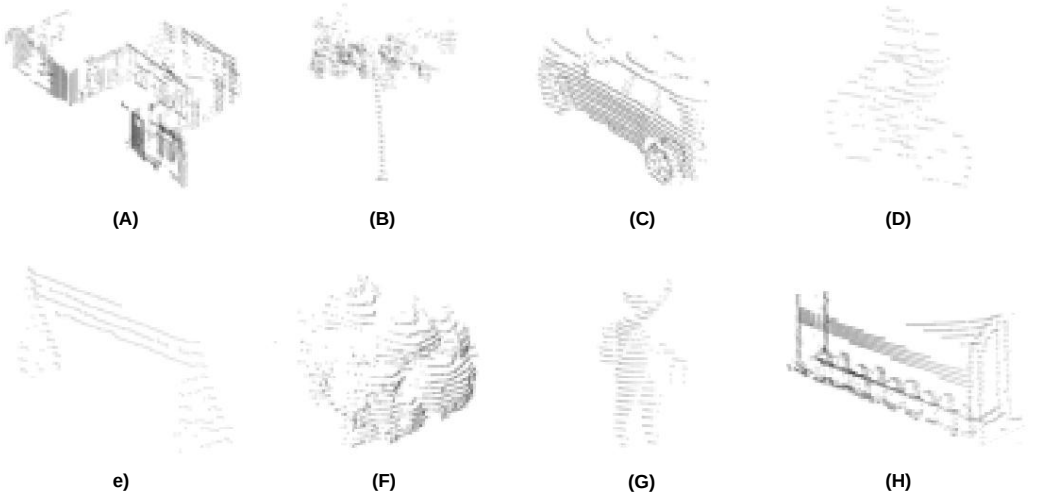


Abbildung 4. Konfiguration der Datensatzklassen: (a) Gebäude; (b) Baum; (c) Fahrzeug; (d) Fahrrad; (e) Hindernis; (f) Grünanlage; (g) Person; (h) städtisches Bauwerk.

2.4. Leistungsbewertung

Im PFE sind die Bildgröße, die Anzahl der Säulen und die Anzahl der Punkte pro Säule Hyperparameter, die sowohl die Leistung als auch die Komplexität des Modell. Die Bildgröße bestimmt die Auflösung und Abmessungen des codierten Bildes. Eine größere Bildgröße führt zu Bildern mit höherer Auflösung; allerdings erhöht sie auch die Rechenleistung des CNN. Ebenso erhöht sich die Anzahl der Säulen und Punkte pro Säule ermöglicht die Erfassung detaillierterer Informationen aus der Punktwolke. Dies erhöht auch die Rechenlast des PFE aufgrund der erhöhten Anzahl von Berechnungen, die zur Verarbeitung einer größeren Datenmenge erforderlich sind. Daher ist es wichtig, eine ausgewogene Die Kombination dieser Hyperparameter ist entscheidend für die Erzielung hoher Leistung und effiziente Berechnung in säulenbasierten Objektklassifizierungsmodellen.

Abbildung 5 zeigt die Architekturen der zur Leistungsbewertung verwendeten Netzwerke , die alle auf dem DSCNN basieren. Tabelle 1 zeigt die Genauigkeiten entsprechend auf verschiedene Hyperparameter und Netzwerkstrukturen, wobei die Gesamtgenauigkeit als eine Bewertungsmetrik für die Objektklassifizierung. Wir trainierten 300 Epochen mit einer Batchgröße von 16, und die Ergebnisse der Epochen mit der besten Leistung sind in Tabelle 1 dargestellt. Das Training wurde mit einer Lernrate von 0,001 und einer Gewichtsabnahmerate von 0,0001 durchgeführt. Erstens Es wurde beobachtet, dass die Netzwerke 1 und 5 eine geringere Leistung zeigten als die Netzwerke 2, 3, und 4. Unter den Netzwerken 2, 3 und 4 sticht Netzwerk 4 durch seine ausgewogene Leistung hervor, zeigte eine vergleichbare Genauigkeit wie die Netzwerke 2 und 3, zeichnete sich jedoch durch eine geringere Rechenkomplexität und weniger Parameter aus. Darüber hinaus zeigte die Bildgröße von 12 herausragende Leistung, während die Leistung vergleichbar war, wenn die Anzahl der Säulen waren 1024 und 512. Die beste Leistung wurde erreicht, wenn 16 Punkte innerhalb waren

Machine Translated by Google

jede Säule. Folglich wurde für das vorgeschlagene Objektklassifizierungsmodell eine Bildgröße von 128 × 128, 512 Säulen und 16 Punkten innerhalb jeder Säule, mit dem Klassifizierungsnetzwerk strukturiert nach Netzwerk 4.

Tabelle 1. Genauigkeit des Modells durch verschiedene Hyperparameter und Netzwerke.

Hyperparameter			Netzwerk				
Bildgröße	Nummer von Säulen	Anzahl der Punkte pro Säule	1	2	3	4	5
256 × 256	2048	32	91,8 %	93,5 %	93,5 %	92,9 %	89,7 %
		16	92,9 %	92,9 %	92,4 %	92,4 %	88,0 %
		8	92,9 %	92,9 %	92,9 %	92,4 %	89,7 %
	1024	32	92,4 %	92,4 %	91,8 %	92,9 %	89,1 %
		16	92,4 %	92,9 %	92,4 %	92,9 %	90,2 %
		8	93,5 %	93,5 %	92,4 %	92,4 %	89,7 %
	512	32	92,4 %	92,9 %	93,5 %	93,5 %	89,7 %
		16	92,9 %	94,0 %	92,4 %	93,5 %	89,7 %
		8	94,6 %	92,9 %	92,4 %	91,8 %	90,8 %
128 × 128	1024	32	90,8 %	91,8 %	92,4 %	94,6 %	90,8 %
		16	90,8 %	94,6 %	92,9 %	92,4 %	91,3 %
		8	90,8 %	91,8 %	91,8 %	94,0 %	91,8 %
	512	32	90,8 %	94,0 %	91,8 %	93,5 %	92,4 %
		16	91,3 %	92,9 %	93,5 %	94,6 %	92,9 %
		8	90,8 %	93,5 %	91,8 %	92,9 %	92,4 %
	256	32	91,8 %	93,5 %	92,9 %	91,8 %	90,8 %
		16	90,8 %	92,9 %	92,4 %	91,3 %	90,8 %
		8	90,8 %	94,0 %	94,0 %	92,4 %	91,3 %
64 × 64	512	32	89,7 %	89,7 %	92,9 %	91,8 %	91,3 %
		16	88,6 %	90,2 %	93,5 %	91,8 %	91,8 %
		8	88,6 %	90,8 %	92,9 %	92,4 %	92,4 %
	256	32	88,6 %	90,2 %	93,5 %	90,8 %	90,8 %
		16	87,5 %	89,7 %	93,5 %	91,8 %	91,8 %
		8	87,5 %	90,8 %	93,5 %	91,8 %	90,8 %
	128	32	88,6 %	89,1 %	89,7 %	90,8 %	89,7 %
		16	89,1 %	90,2 %	91,3 %	90,2 %	90,8 %
		8	86,4 %	92,4 %	91,3 %	89,1 %	91,8 %

Tabelle 2 zeigt einen Leistungsvergleich zwischen dem vorgeschlagenen Klassifizierungsnetzwerk und den anderen Klassifizierungsnetzwerken in dieser Studie. Das vorgeschlagene Netzwerk erreichte die höchste Klassifizierungsgenauigkeit von 94,6% basierend auf dem DSCNN, mit einem geringe Anzahl von Multiplikations-Akkumulationsoperationen (MAC) von 37 M und eine geringe Anzahl von Parameter von 50,8 K.

Machine Translated by Google

Tabelle 2. Leistungsvergleich mit anderen Netzwerken.

Netzwerk	Genauigkeit	Anzahl der MACs	Anzahl der Parameter
LeNet	92,2 %	787,6 M	2,0 M
ResNet18	83,8 %	1,4 G	11,4 M
ResNet34	82,3 %	2,0 G	21,5 M
MobileNet	85,7 %	128,3 M	3,2 M
Unsere	94,6 %	37,0 M	50,8 T

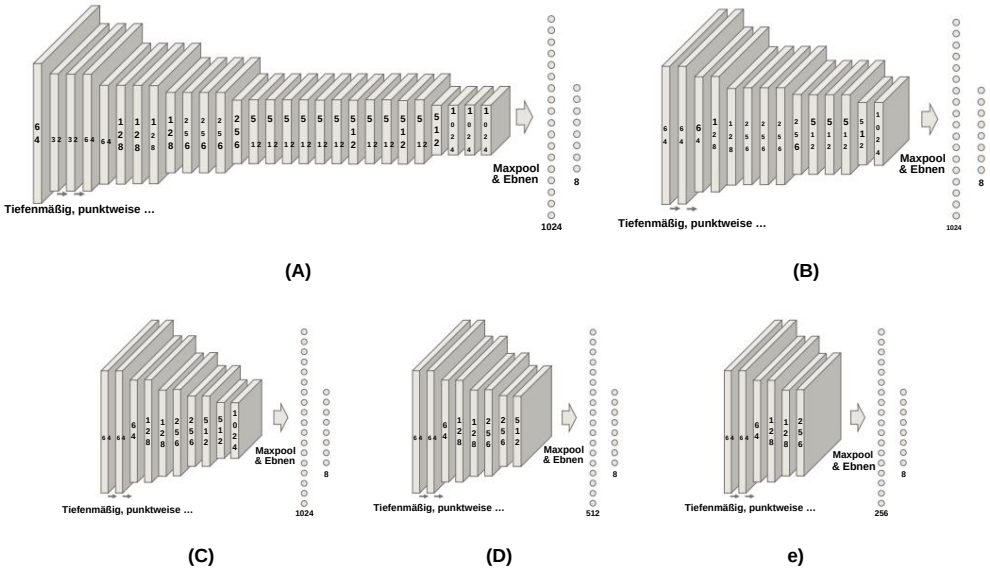


Abbildung 5. Architektur des DSCNN-basierten Objektklassifizierungsnetzwerks: (a) Netzwerk 1; (b) Netzwerk 2; (c) Netzwerk 3; (d) Netzwerk 4; (e) Netzwerk 5.

3. Implementierung des Beschleunigungssystems

Das vorgeschlagene Beschleunigungssystem wurde auf einem FPGA implementiert, indem das geistige Eigentum (IP) des PFE-Hardwarebeschleunigers und das IP des Hardwarebeschleunigers des Klassifizierungsnetzwerks mit einer MPU integriert wurden. Abschnitt 3.1 stellt die Bitbreite jedes IP dar, die durch Quantisierungsexperimente erhalten wurde. Abschnitt 3.2 beschreibt die Struktur und den Betrieb des vorgeschlagenen Beschleunigungssystems. Abschnitt 3.3 beschreibt die Struktur und das Verhalten des PFE-Hardwarebeschleunigers und Abschnitt 3.4 beschreibt das FINN, das zur Implementierung des Hardwarebeschleunigers des Objektklassifizierungsnetzwerks verwendet wird.

3.1. Quantisierung **In**

dieser Studie haben wir das vorgeschlagene Objektklassifizierungsmodell quantisiert, um das vorgeschlagene Beschleunigungssystem in der Hardware zu implementieren. Um die Leistungseinbußen aufgrund von durch die Quantisierung verursachten Fehlern zu minimieren, haben wir ein quantisierungsbewusstes Tr Die Ergebnisse der Experimente mit unterschiedlichen Bitformaten für PFE und Klassifizierungsnetzwerke sind in Tabelle 3 aufgeführt. PFE ist mit 16 und 32 Bit genauer als mit 8 Bit, aber es gibt keinen Unterschied in der Genauigkeit zwischen 16 und 32 Bit. Daher haben wir uns entschieden, ein 16-Bit-Format für PFE zu verwenden. Das Objektklassifizierungsnetzwerk zeigte eine höhere Genauigkeit mit 4 und 8 Bit, aber es gab keinen Unterschied in der Leistung; daher haben wir 4 Bit verwendet.

Machine Translated by Google

Tabelle 3. Genauigkeit nach Bitformat von PFE- und Klassifizierungsnetzwerken.

PFE	Klassifizierungsnetzwerk		
	2	4	8
8	90,3 %	91,7 %	92,3 %
16	89,7 %	94,3 %	94,3 %
32	88,7 %	92,7 %	94,3 %

3.2. Systemarchitektur

Abbildung 6 zeigt die Struktur des vorgeschlagenen Beschleunigungssystems. Zur Validierung des System in einer FPGA-Umgebung wurden die Beschleuniger über eine MPU mit einer Advanced Extensible Interface (AXI) Interconnect und in einem FPGA implementiert. Jede Beschleuniger- IP ist ein Slave der MPU, und die MPU sendet die Start- und End-DRAM-Adressen, die IP-Adresse, auf die zugegriffen wird, die Anzahl der aus dem DRAM zu lesenden Wörter, AXI-Protokollparameter, und Start- und Stoppsignale an die Slave-Register des IP. Der IP fungiert als Master für DRAM, Generieren von Lese- und Schreibadressen über die Master-Schnittstelle, um die erforderlichen Daten abzurufen aus dem DRAM während der Berechnung und schreiben Sie die Ergebnisse zurück in den DRAM.

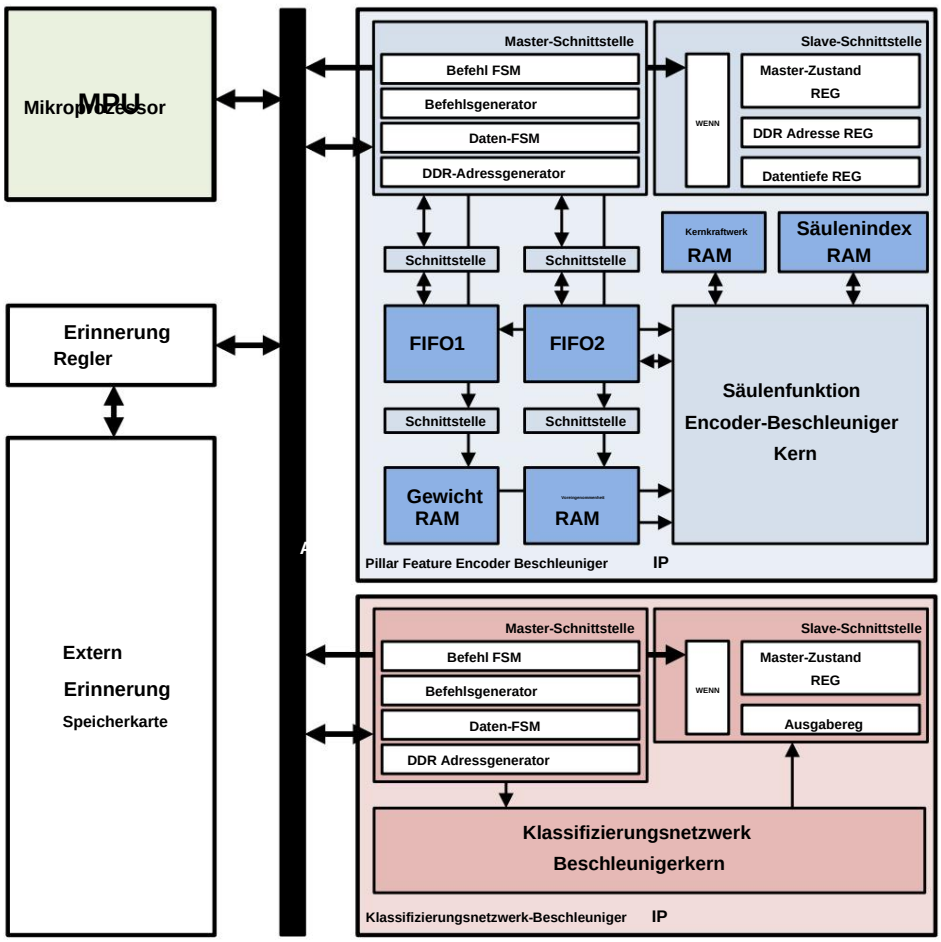


Abbildung 6. Architektur des vorgeschlagenen Beschleunigungssystems auf FPGA.

Die MPU setzt zuerst das Slave-Register des PFE-Beschleuniger-IP, das mit dem Lesen beginnt die Punktwolke vom DRAM über die Master-Schnittstelle, um die Punktwolke in Säulen. Der PFE-Beschleuniger prüft die Koordinaten der Punktwolke und speichert die Punkt Punktwolke nacheinander in FIFO1. Um die Punktwolke Säule für Säule im DRAM zu speichern, generiert die DRAM-Adressen und speichert sie in FIFO2. Wenn alle DRAM-Adressen

Machine Translated by Google

für die Punktwolke werden generiert und in FIFO2 gespeichert, die Masterschnittstelle verweist auf FIFO1 und FIFO2, um die Punktwolke an einem geeigneten Ort im DRAM zu speichern.

Der PFE-Beschleuniger organisiert die gesamte Punktwolke in Säulen, indem er sie entsprechend im DRAM speichert, liest die Punktwolke dann zurück und speichert sie in FIFO1 und Außerdem liest es die Gewichte und Bias zur Durchführung der PW-Faltung aus dem DRAM und speichert sie im Gewichtsspeicher (WM) und im Bias-Speicher (BM). Nach dem Hinzufügen des HCF zur Punktwolke wird eine PW-Faltung durchgeführt, um die Merkmale zu erweitern, und es wird eine Max-Pool-Operation ausgeführt. Somit wird eine Säule mit 64 Merkmalen codiert, was einem Pixel mit 64 Kanälen in einem 2D-Bild entspricht. Das codierte Pixel wird unter Berücksichtigung seiner Position im DRAM gespeichert.

Nachdem der PFE-Beschleuniger alle Punktwolken in 2D-Bilder kodiert hat, setzt die MPU das Slave-Register des Klassifizierungsnetzwerk-Beschleuniger-IP und startet den Beschleuniger, um Bilder über die Master-Schnittstelle aus dem DRAM zu lesen. Nachdem der Klassifizierungsnetzwerk-Beschleuniger alle Bilder gelesen und alle Berechnungen durchgeführt hat, schreibt er die Klassifizierungsergebnisse in das Slave-Register. Die MPU schließt den Systembetrieb ab, indem sie auf das Slave- Register des Klassifizierungsnetzwerk-Beschleuniger-IP zugre

3.3. PFE-Hardwarearchitektur

Abbildung 7 zeigt ein Blockdiagramm des PFE-Beschleunigers. Der vorgeschlagene PFE-Beschleuniger besteht aus einer Gruppierungseinheit (GU), einer HCF-Einheit (HU), einer PFN-Einheit (PU) und ac Die Speicher sind FIFO1, FIFO2, Anzahl der Punkte pro Säule (NPPP)-Puffer (NB), Säulenindexpuffer (PB), HCF-Puffer1 (HB1), HB2, WM und BM. HB1 und HB2 werden in einem Ping-Pong-Schema von HU und PU innerhalb einer Pipeline-Struktur verwendet. Während der HU arbeitet, schreibt er Ergebnisse in HB1, während die PU gleichzeitig auf Daten von HB2 zugreift. Anschließend wechselt die HU zum Schreiben von Berechnungsergebnissen in HB2, während die PU die **in dieser Phase in HB1 gespeicherten Daten verwendet .**

Die Punktwolken werden der Reihe nach in FIFO1 gespeichert und der Gitterfinder in der GU prüft die Koordinaten der Punkte, um die Koordinaten der Säulen zu erhalten, die die Punkte entha Wenn beispielsweise die Breite einer Gitterzelle ein ganzzahliger Wert von 25 ist , **, die Koordinaten der kann die Säule, die einen Punkt enthält, bestimmt werden, indem die Koordinaten des Punkts durch 25 . Das** geteilt **werden. Die Divisionsoperation wird durch eine 5-Bit-Rechtsverschiebungsoperation ersetzt. Die Koordinaten der Säulen werden als Leseadresse des NB verwendet, um den NPPP zu bestimmen. Wenn der NPPP 0 ist, wird eine neue Säule erstellt; daher wird der Säulenzähler um eins erhöht und der Wert wird in den PB geschrieben, um als Index der Säule verwendet zu werden. Wenn der NPPP kleiner als 16 ist, wird er um 1 erhöht und in den NB zurückgeschrieben; andernfalls bleibt er unverändert. Die Adresse wird generiert, indem der Säulenindex multipliziert mit 16 (dem maximalen NPPP) als Offset verwendet und dann der NPPP zu diesem Offset addiert wird. Die generierte Adresse wird in FIFO2 gespeichert. Mit der Punktwolke in FIFO1 und der DRAM-Adresse in FIFO2 wird die Punktwolke an der richtigen Stelle im DRAM gespeichert.**

Nachdem alle Punktwolken in den Säulen im DRAM gespeichert sind, werden sie in FIFO1 und FIFO2 zurückgelesen. In FIFO2 wird die DRAM-Schreibadresse während des Gruppierungsvorgangs gespeichert , während die DRAM-Lesedaten zu anderen Zeiten gespeichert werden. Der PFE-Beschleuniger führt dann die Berechnung für alle Punktwolken säulenweise durch. Die Mittelwerteinheit (AU) der HU addiert die x-, y- und z- **Koordinaten aller Punkte in der Säule und wird durch das NPPP geteilt, um die arithmetischen Mittelkoordinaten der Punkte zu erhalten. Die Offseteinheit in der HU berechnet** xm, ym , zm, **indem sie die Mittelpunktskoordinaten der von der AU erhaltenen Punkte von den x-, y-,** z- **Koordinaten jedes Punkts abzieht, und** xc, yc, **indem sie die Mittelpunktskoordinaten der Säule von den x-,** y- **Koordinaten jedes Punkts abzieht. Die x-, y-, z-, r-,** xm-, ym- , zm-, xc- und yc **-Werte der Punkte werden d**

Machine Translated by Google

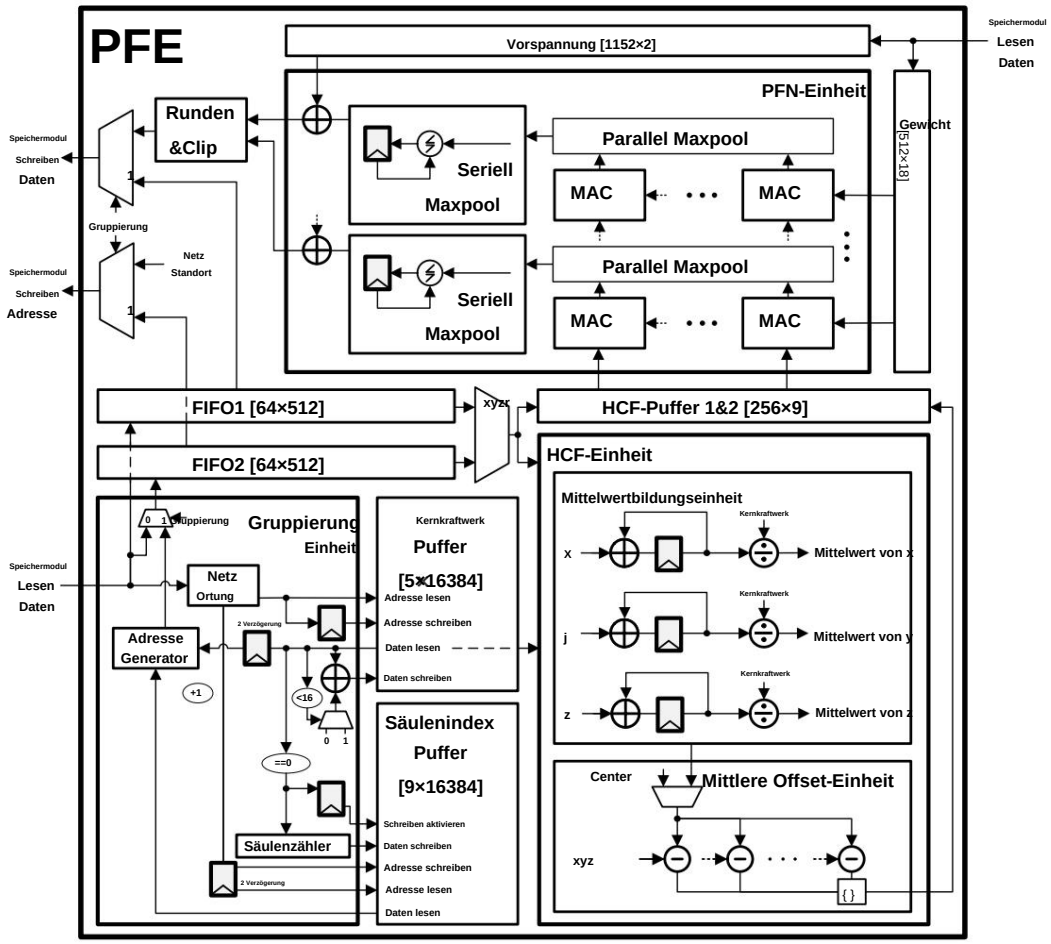


Abbildung 7. Blockdiagramm des PFE-Beschleunigers.

Abbildung 8a zeigt die von der PU ausgeführte Operation auf Tensorebene. Die PU führt PW- Faltung, Batch-Normalisierung, Max-Pool und ReLU-Operationen auf dem HCF aus. Die Gewichte und Biase der Batch-Normalisierung können mit denen der PW-Faltungsschicht kombiniert werden . Bei der PW-Faltung werden nur die Gewichte verwendet und die Biase werden nach einer verzerrten ReLU-Operation hinzugefügt, die die Schwellenwerte bei -bias statt bei 0 setzt.

Abbildung 8b stellt die Operation in Abbildung 8a auf Matrixebene dar. Die PU führt die Multiplikation der HCF-Matrix der Größe (16, 9) und der Gewichtsmatrix der Größe (9, 64) mithilfe eines 8 × 32 MAC-Arrays durch. Dazu wird die HCF-Matrix in zwei Matrizen der Größe (8, 9) und die Gewichtsmatrix in zwei Matrizen der Größe (9, 64) unterteilt, wobei die gesamte Matrixmultiplikation in vier Teilen durchgeführt wird. Beispielsweise ergibt die Matrixmultiplikation von HCF-Kachel 2 und Gewichtskachel 1 Ergebniskachel 2, und diese Kacheln werden maximal gepoolt, um maximal gepoolte Kachel 1 zu erstellen.

Abbildung 8c veranschaulicht den Betrieb der PU auf Hardwareebene in chronologischer Reihenfolge. Die PU liest HCF-Kachel 1 und Gewichtskachel 1 jeweils aus HB und WM. Die PU liest HCF-Kachel 1 Spalte für Spalte und Gewichtskachel 1 Zeile für Zeile über 9 Zyklen und speist sie in das MAC-Array ein. Das MAC-Array verarbeitet die HCF- und Gewichtseingaben über 9 Zyklen und gibt Ergebniskachel 1 der Größe (8, 32) aus. Ergebniskachel 1 wird dann von der parallelen Max-Pool-Einheit verarbeitet, um Max-Pool-Operationen durchzuführen, wodurch Zwischenkachel 1 der Größe (1, 32) erzeugt wird. Da die parallele Max-Pool-Einheit sofort Max-Pool-Operationen am MAC-Array- Ausgang durchführt, spart sie Zyklen und Pufferspeicherplatz, indem sie Ergebniskachel 1 nic Zwischenkachel 1 wird mit dem -bias-Wert in der seriellen Max-Pool-Einheit verglichen und der größere Wert wird im Register gespeichert. Durch Ausführen von Bias-ReLU-Operationen in der seriellen Max-Pool- Einheit könnte die ReLU-Operation in die Max-Pool-Operation integriert werden. Nach dem Erstellen von Zwischenkachel 1 liest die PU HCF-Kachel 2 und Gewichtskachel 1, um Z

Machine Translated by Google

2. Zwischenkachel 2 wird Max-Pool-Operationen unterzogen, wobei Zwischenkachel 1 im Register der seriellen Max-Pool-Einheit gespeichert wird, was zu Max-Pool-Kachel 1 führt. Max-Pool-Kachel 1 ist ein halbes Pixel, wenn die Verzerrung hinzugefügt und gerundet und abgeschnitten wird. Dieses halbe Pixel wird basierend auf seinen Koordinaten im Bild an der richtigen Stelle im DRAM gespeichert. Um die DRAM-Schreibadresse zu erstellen , wird die x- **Koordinate des Pixels multipliziert mit der Breite des Bildes als Offset verwendet, und die** y- **Koordinate wird zu diesem Offset addiert. Wenn die PU sowohl Max-Pool-Kacheln 1 als auch 2 erstellt hat , befindet sich ein vollständiges Pixel im DRAM**

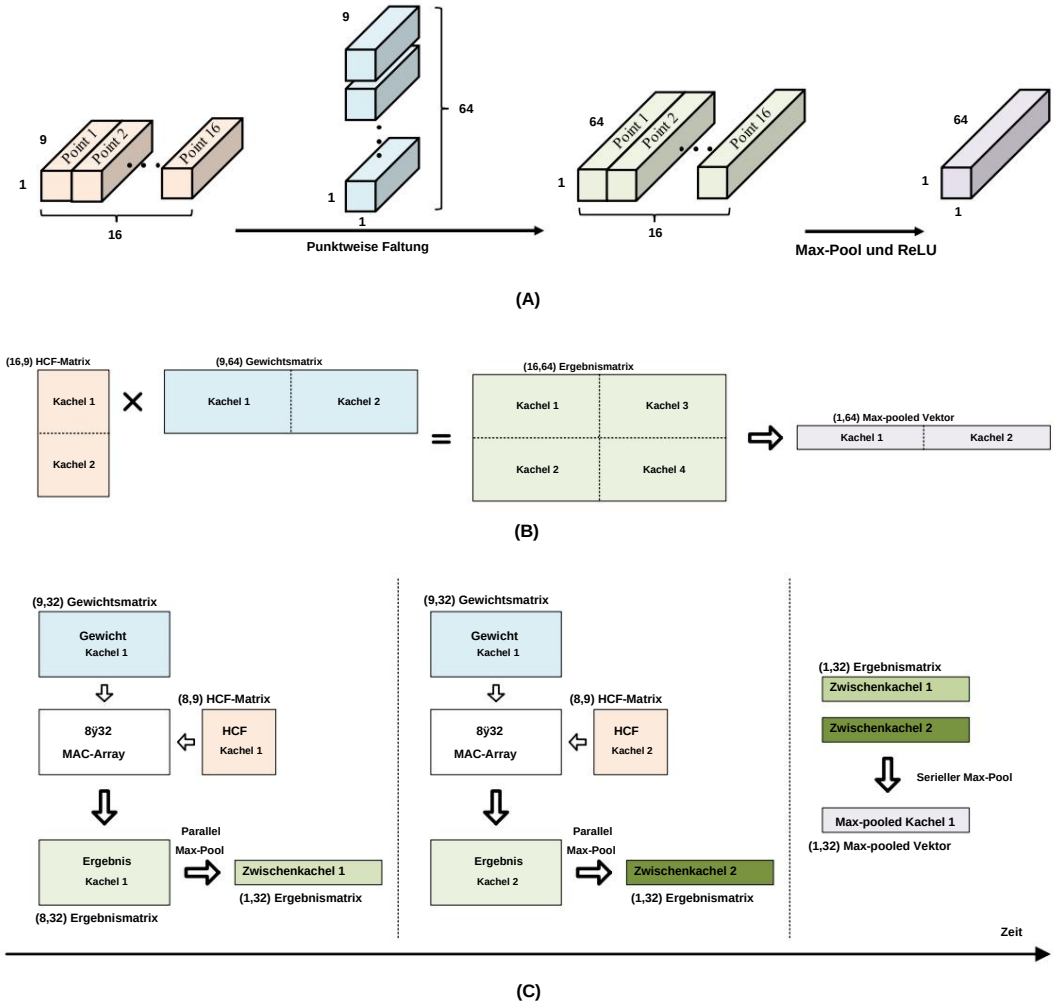


Abbildung 8. Funktionsweise der PU: (a) Funktionsweise auf Tensorebene; (b) Funktionsweise auf Matrixebene; (c) Funktionsweise auf Hardwareebene.

3.4. Klassifizierungsnetzwerkbeschleuniger mit FINN

Xilinx bietet verschiedene Tools zur Implementierung von Deep-Learning-Netzwerken. Vitis AI ist eine Plattform zur Implementierung und Optimierung von Anwendungen für künstliche Intelligenz und maschinelles Lernen auf FPGAs und adaptiven System-on-Chip-Plattformen. Vitis AI unterstützt verschiedene Frameworks und Bibliotheken für künstliche Intelligenz und maschinelles Lernen und generiert eine Deep- Learning-Processing-Unit (DPU)-IP. FINN ist ein Open-Source-Framework für FPGA-basierte quantisierte Deep-Learning-Netzwerkinferenz. FINN verwendet ein offenes neuronales Netzwerkaustauschmodell als Eingabe, optimi

In dieser Studie haben wir FINN verwendet, um eine Hardwareversion eines Objektklassifizierungsnetzwerks zu implementieren . Während die DPU als Allzweckbeschleuniger für Deep-Learning-Operationen dient , optimiert FINN die Hardwarearchitektur und den Datenfluss speziell für das Deep- Learning-Netzwerk. Folglich bietet FINN geringere Latenzzeiten, ermöglicht Echtzeit-Reaktionsgeschwindigkeiten und kann durch Optimierung der Pipeline einen hohen Durchsatz errei

Machine Translated by Google

Werke mit geringer Komplexität, wie das vorgeschlagene Klassifizierungsnetzwerk, können deutlich Reduzierung der Hardwareressourcennutzung. Darüber hinaus unterstützt FINN 4-Bit-Ganzzahlen und bietet eine Vorteil gegenüber DPUs. Unter Ausnutzung der Vorteile von FINN wurde in [19] ein Backbone Schicht von PointPillars mit FINN.

4. Ergebnisse der Umsetzung

Das vorgeschlagene Beschleunigungssystem wurde auf einem Xilinx Zynq Ultrascale+ implementiert. ZCU104 [30]. Tabelle 4 fasst die Hardware-Ressourcennutzung und Betriebsfrequenz zusammen des vorgeschlagenen Beschleunigungssystems. Das vorgeschlagene Beschleunigungssystem nutzte 75.548 konfigurierbare Logikblock-(CLB)-Look-Up-Tabellen (LUTs), 56.931 CLB-Register, 374 digitale Signal Prozessoren (DSPs) und 65 Block-RAMs. Die Betriebsfrequenz betrug 187,5 MHz, was ermöglicht einen Hochgeschwindigkeitsbetrieb und die Reaktionszeit der vorgeschlagenen Objektklassifizierung System betrug 6,41 ms.

Tabelle 4. Hardwareressourcen des vorgeschlagenen Objektklassifizierungssystems.

Einheit	CLB LUTs	CLB Register	DSPs	Block RAM	Frequenz (MHz)
PFE	35.567	28.513	302	10.5	
Klassifizierungsnetzwerk	34.788	21.625	72	54,5	
AXI-Verbindung	5193	6793	0	0	
Gesamt	75.548	56.931	374	65	187,5

Tabelle 5 vergleicht die Antwortzeiten der Firmware-basierten PFE-Implementierung auf ein ARM Cortex A53 (Basislinie) mit 1,2 GHz und die vorgeschlagene PFE-Hardware Beschleuniger (vorgeschlagen) in Bezug auf die Ausführungszeit für jede Operation. Die Gruppierung Betrieb, verantwortlich für die Umwandlung einer Punktwolke in eine Reihe von Säuleneinheiten, erfahrene eine bemerkenswerte Beschleunigung von 23,67 mal, von 0,71 ms auf 0,03 ms. Besonders bemerkenswert ist die Die HCF- und PFN-Operationen erreichten eine signifikante Beschleunigung um das 42,6-fache, von 8,52 ms auf 0,20 ms. Diese bemerkenswerte Beschleunigung wurde durch die Nutzung der parallele Rechenleistung des MAC-Arrays im PFE-Beschleuniger und der pipelined Struktur der HCF- und PFN-Operationen. Die Gesamtausführungszeit für PFE-Operationen von 9,23 ms auf 0,23 ms reduziert, was einer 40-fachen Beschleunigung im Vergleich zu mit der Firmware-basierten Implementierung.

Tabelle 5. Vergleich der PFE-Reaktionszeit mit dem Ausgangswert.

Betrieb	Baseline (Firmware)	Vorgeschlagen (Hardware)
Gruppierung	0,71 ms	0,03 ms
HCF	0,20 ms	0,20 ms
PFN	8,32 ms	
Gesamt	9,23 ms	0,23 ms

Tabelle 6 präsentiert einen Vergleich zwischen dem vorgeschlagenen Objektklassifizierungssystem und säulenbasierte Methoden, die in früheren Studien implementiert wurden. In [19–21] wurde ein heterogenes System zwischen MPU und FPGA wurde konstruiert, während unser System eine End-to- End-Hardwareimplementierung aller Operationen auf einem FPGA darstellt. In [19] wurde die PFE-Berechnung wurde auf einer MPU durchgeführt, und die Berechnung des Deep-Learning-Netzwerks erfolgte mit ein Beschleuniger, der über FINN implementiert wurde. Ref. [20] berechnete die PFE auf einer MPU und ein Deep-Learning-Netzwerk auf einem Beschleuniger ausgeführt, der mit einem Registertransfer implementiert wurde Level (RTL) Design. Ref. [21] verwendete sowohl eine MPU als auch eine DPU für die PFE-Berechnung und eine DPU für Deep-Learning-Netzwerkberechnungen. Wir haben den PFE mit einem Beschleuniger berechnet

Machine Translated by Google

implementiert durch ein RTL-Design und Deep-Learning-Netzwerk mit einem Beschleuniger umgesetzt durch FINN.

Unsere Methode erreichte die schnellste Reaktionszeit von 6,41 ms, weil sie die einzige war, Eines davon beschleunigte die PFE, und das vorgeschlagene Objektklassifizierungsmodell wurde entwickelt durch verschiedene Experimente eine geringe Rechenkomplexität zu erreichen. Darüber hinaus Experimente führten zu einem Objektklassifizierungsmodell mit weniger Parametern. Unsere Methode wurden zwar mehr DSPs als in [19] verwendet, dafür aber deutlich weniger LUTs, Register und Speicher. Ref. [19] verwendete die gleiche FPGA-Plattform für Edge-Umgebungen wie wir, die ZCU104, einen niedrigen Stromverbrauch zu erreichen. Allerdings verbrauchte es immer noch mehr Strom als unseres, die ein effizientes Deep Learning Netzwerk implementierte. Im Vergleich zur Methode in [20], Unsere Methode verbrauchte insgesamt weniger Ressourcen. Besonders auffällige Unterschiede gab es in DSPs und Speichernutzung. Ref. [20] verbrauchte auch wenig Strom durch die Implementierung eines effizienten Deep-Learning-Netzwerk. Darüber hinaus wurde in Ref. [20] das Deep-Learning-Netzwerk entwickelt Beschleuniger auf RTL-Ebene, verbrauchte es sogar weniger Strom als unseres, das entworfen wurde basierend auf HLS. Darüber hinaus verwendete unsere Methode mehr LUTs, aber weniger Register und DSPs, und wesentlich weniger BRAM und URAM als in [21]. Darüber hinaus ist unsere Methode verbrauchte deutlich weniger Strom als in [21]. Ref. [21] verwendete die U280 FPGA-Plattform auf Serverumgebungen ausgerichtet, die einen viel höheren Stromverbrauch haben und nicht geeignet für Edge-Umgebungen.

Tabelle 6. Vergleich mit anderen säulenbasierten Implementierungen.

	[19]	[20]	[21]	Unsere
Plattform	ZCU104	ZC706	U280	ZCU104
Berechnung Gerät	MPU und FPGA (FINN)	MPU und FPGA (RTL)	MPU und FPGA (DPU)	FPGA (RTL & FINN)
Ausführung Zeit (ms)	377.1	43,26	64.1	6.41
CLB LUTs	189.074	128.721	48.006	75.548
CLB-Register	150.187	111.118	85.801	56.931
DSPs	88	883	525	374
Block-RAM	159	382,5	131,5	65
Ultra-RAM	-	-	64	1
Frequenz (MHz)	150	150	300	187,5
Leistung (W)	6.5	3.6	73,8	4.4

5. Schlussfolgerungen

In dieser Studie schlagen wir ein Deep-Learning-Objektklassifizierungsmodell vor, das auf einem säulenbasierte Methode für Echtzeitreaktionen autonomer mobiler Roboter. Das Objekt Das Klassifizierungsnetzwerk des Deep-Learning-Modells basiert auf dem DSCNN, das über eine Geringe Komplexität und wenige Parameter, wodurch es für Edge-Geräteumgebungen geeignet ist. Das Objektklassifizierungsmodell wurde mit einem Datensatz trainiert, der die Betriebsdaten berücksichtigte. Umgebung autonomer mobiler Roboter. Um die Leistungseinbußen durch Zur Quantisierung haben wir das Modell quantisiert und mit QAT getestet und erreichten eine Klassifizierung Genauigkeit von 94,3 %. Da die Berechnung des PFE in säulenbasierten Objektklassifizierungsmodellen zeitaufwändig war, schlugen wir einen Hardwarebeschleuniger für den PFE vor. Der Der PFE-Beschleuniger wurde durch RTL-Design implementiert und erreichte eine Reaktionszeit von 0,23 ms, was 40 mal schneller war als das Firmware-Programm. Der Beschleuniger für das Objekt Das Klassifizierungsnetzwerk des vorgeschlagenen Deep-Learning-Modells wurde implementiert durch FINN-Framework und erreichte eine Ausführungszeit von 6,18 ms. Wir schlugen auch ein Objekt vor Klassifizierungssystem, das den vorgeschlagenen PFE-Beschleuniger und die Objektklassifizierung integriert Netzwerkbeschleuniger über eine AXI-Verbindung. Das vorgeschlagene Objektklassifizierungssystem

Machine Translated by Google

wurde auf einem FPGA implementiert und verifiziert und erreichte eine Echtzeit-Reaktionszeit von 6,41 ms.

Der vorgeschlagene PFE-Beschleuniger kann mit anderen Deep-Learning-Netzwerkbeschleunigern kombiniert werden, um vielfältigere Deep-Learning-Modelle zu beschleunigen. In dieser Studie haben wir jedoch nur die Objektklassifizierungsmodelle beschleunigt. In zukünftigen Arbeiten planen wir die Implementierung eines Systems, das die Beschleunigung vielfältigerer und anspruchsvollerer Deep- Learning-Modelle unterstützt.

Autorenbeiträge: CP hat das vorgeschlagene Beschleunigungssystem entworfen und implementiert, das Experiment und die Auswertung durchgeführt und das Papier geschrieben. SL hat das vorgeschlagene Beschleunigungssystem ausgewertet und dieses Manuskript überarbeitet. YJ hat die Forschung konzipiert und geleitet, die experimentellen Ergebnisse analysiert und das Papier geschrieben. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen un

Finanzierung: Diese Arbeit wurde durch ein Stipendium des Institute of Information & Communications Technology Planning & Evaluation (IITP) der koreanischen Regierung (MSIT) (Nr. 2022-0-00960) gefördert, und die CAD-Tools wurden von IDEC unterstützt.

Erklärung des Institutional Review Board: Nicht zutreffend.

Einverständniserklärung: Nicht zutreffend.

Datenverfügbarkeitserklärung: Datensatz auf Anfrage bei den Autoren erhältlich.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Verweise

1. **Varlamov, O. „Gehirne“ für Roboter: Anwendung der Mivar-Expertensysteme zur Implementierung autonomer intelligenter Roboter.** Big Data Res. **2021,** 25, **100241. [CrossRef]**
2. **Liu, Y.; Li, Z.; Liu, H.; Kan, Z. Skill Transfer Learning für autonome Roboter und Mensch-Roboter-Kooperation: Eine Untersuchung.** Roboter. Auton. Syst. **2020,** 128, **103515. [CrossRef]**
3. **Yoshioka, M.; Suganuma, N.; Yoneda, K.; Aldibaja, M. Echtzeit-Objektklassifizierung für autonome Fahrzeuge mit LIDAR. In Proceedings der 2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), Okinawa, Japan, 24.–26. November 2017; S. 210–211.**
4. **Gao, H.; Cheng, B.; Wang, J.; Li, K.; Zhao, J.; Li, D. Objektklassifizierung mittels CNN-basierter Fusion von Vision und LIDAR in Autonome Fahrzeugumgebung.** IEEE Trans. Ind. Inform. **2018,** 14, **4224–4231. [CrossRef]**
5. **Zhou, Y.; Liu, L.; Zhao, H.; López-Benítez, M.; Yu, L.; Yue, Y. Auf dem Weg zur tiefen Radarwahrnehmung für autonomes Fahren: Datensätze, Methoden und Herausforderungen.** Sensors **2022,** 22, **4208. [CrossRef] [PubMed]**
6. **Su, H.; Maji, S.; Kalogerakis, E.; Learned-Miller, E. Multi-view Convolutional Neural Networks für 3D-Formerkennung. In Proceedings der IEEE International Conference on Computer Vision, Santiago, Chile, 7.–13. Dezember 2015; S. 945–953.**
7. **Hoang, L.; Lee, SH; Lee, EJ; Kwon, KR GSV-NET: Ein multimodales Deep Learning-Netzwerk zur Klassifizierung von 3D-Punktwolken.** Appl. Sci. **2022,** 12, **483. [CrossRef]**
8. **Qi, CR; Su, H.; Mo, K.; Guibas, LJ Pointnet: Deep Learning auf Punktmengen für 3D-Klassifizierung und -Segmentierung. In Proceedings der IEEE-Konferenz zu Computer Vision und Mustererkennung, Honolulu, HI, USA, 21.–26. Juli 2017; S. 652–660.**
9. **Zhou, Y.; Tuzel, O. Voxelnet: End-to-End-Lernen für 3D-Objekterkennung auf Basis von Punktwolken. In Proceedings der IEEE- Konferenz zu Computer Vision und Mustererkennung, Salt Lake City, UT, USA, 18.–22. Juni 2018; S. 4490–4499.**
10. **Bhanushali, D.; Relyea, R.; Manghi, K.; Vashist, A.; Hochgraf, C.; Ganguly, A.; Kwasinski, A.; Kuhl, ME; Ptucha, R. LiDAR-Kamera Fusion zur 3D-Objekterkennung.** Electron. Imaging **2020,** 32, **1–9. [CrossRef]**
11. **Yan, Y.; Mao, Y.; Li, B. Zweitens: Sparsely Embedded Convolutional Detection.** Sensors **2018,** 18, **3337. [CrossRef] [PubMed]**
12. **Lang, AH; Vora, S.; Caesar, H.; Zhou, L.; Yang, J.; Beijbom, O. Pointpillars: Schnelle Encoder zur Objekterkennung aus Punktwolken. In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung, Long Beach, CA, USA, 15.–20. Juni 2019; S. 12697–12705.**
13. **Lis, K.; Kryjak, T. PointPillars Backbone-Typauswahl für schnelle und genaue LiDAR-Objekterkennung. In Proceedings der International Conference on Computer Vision and Graphics, Warschau, Polen, 19.–21. September 2022; Springer: Berlin/Heidelberg, Deutschland, 2022; S. 99–119.**
14. **Shu, X.; Zhang, L. Forschung zum PointPillars-Algorithmus basierend auf einem Feature-Enhanced Backbone Network.** Electronics **2024,** 13, **1233. [Querverweis]**
15. **Wang, Y.; Han, X.; Wei, X.; Luo, J. Instance Segmentation Frustum–PointPillars: Ein leichtgewichtiger Fusionsalgorithmus für die Kamera- LiDAR-Wahrnehmung beim autonomen Fahren.** Mathematics **2024,** 12, **153. [CrossRef]**
16. **Agashe, P.; Lavanya, R. Objekterkennung mit PointPillars mit modifiziertem DarkNet53 als Backbone. In Proceedings der 2023 IEEE 20th India Council International Conference (INDICON), Hyderabad, Indien, 14.–17. Dezember 2023; S. 114–119.**

Machine Translated by Google

17. **Choi, Y.; Kim, B.; Kim, SW Leistungsanalyse von PointPillars auf CPU- und GPU-Plattformen. In Proceedings der 36. Internationalen Technischen Konferenz für Schaltkreise/Systeme, Computer und Kommunikation (ITC-CSCC) 2021, Jeju, Republik Korea, 27.–30. Juni 2021; S. 1–4.**

18. **Silva, A.; Fernandes, D.; Névoa, R.; Monteiro, J.; Novais, P.; Girão, P.; Afonso, T.; Melo-Pinto, P. Ressourcenbeschränkte Onboard- Inferenz der 3D-Objekterkennung und -lokalisierung in Punktwolken für selbstfahrende Anwendungen.** Sensors **2021,** 21, **7933.** **[Querverweis] [PubMed]**
19. **Stanisz, J.; Lis, K.; Gorgon, M. Implementierung des Pointpillars-Netzwerks zur 3D-Objekterkennung in reprogrammierbaren Heterogene Geräte mit FINN.** J. Signal Process. Syst. **2022,** 94, **659–674.** **[CrossRef]**
20. **Li, Y.; Zhang, Y.; Lai, R. TinyPillarNet: Auf winzigen Säulen basierendes Netzwerk zur Erkennung von 3D-Punktwolkenobjekten am Rand.** IEEE Trans. Circuits Syst. Video Technol. **2023,** 34, **1772–1785.** **[CrossRef]**
21. **Latotzke, C.; Kloeker, A.; Schoening, S.; Kemper, F.; Slimi, M.; Eckstein, L.; Gemmeke, T. FPGA-basierte Beschleunigung der Verarbeitung und Erkennung von Lidar-Punktwolken am Rand. In Proceedings des 2023 IEEE Intelligent Vehicles Symposium (IV), Anchorage, AK, USA, 4.–7. Juni 2023; S. 1–8.**
22. **Tan, M.; Le, Q. Efficientnet: Neue Überlegungen zur Modellskalierung für Convolutional Neural Networks. In Proceedings of the International Konferenz zum maschinellen Lernen. PMLR, Long Beach, CA, USA, 9.–15. Juni 2019; S. 6105–6114.**
23. **Chollet, F. Xception: Deep Learning mit tiefenweise trennbaren Faltungen. In Proceedings der IEEE-Konferenz zu Computer Vision und Mustererkennung, Honolulu, HI, USA, 21.–26. Juli 2017; S. 1251–1258.**
24. **Howard, AG; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Effiziente Convolutional Neural Networks für mobile Vision-Anwendungen.** arXiv **2017, arXiv:1704.04861.**
25. **Lu, G.; Zhang, W.; Wang, Z. Optimierung tiefenweise trennbarer Faltungsoperationen auf GPUs.** IEEE Trans. Parallel Distrib. Syst. **2021,** 33, **70–87.** **[CrossRef]**
26. **Kaiser, L.; Gomez, AN; Chollet, F. Tiefenseparierbare Faltungen für neuronale maschinelle Übersetzung.** arXiv **2017, arXiv:1706.03059.**
27. **Ouster. Ouster OS1 Lidar-Sensor. Online verfügbar: https://ouster.com/products/hardware/os1-lidar-sensor (aufgerufen am 22 Mai 2024).**
28. **Caesar, H.; Bankiti, V.; Lang, AH; Vora, S.; Liong, VE; Xu, Q.; Krishnan, A.; Pan, Y.; Baldan, G.; Beijbom, O. nuscenes: Ein multimodaler Datensatz für autonomes Fahren. In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung , Seattle, WA, USA, 14.–19. Juni 2020; S. 11621–11631.**
29. **Häufig gestellte Fragen zur Vitis™ AI-Dokumentation von Xilinx Inc. Online verfügbar: https://xilinx.github.io/Vitis-AI/3.0/ html/docs/reference/ faq.html#what-is-the-difference-between-the-vitis-ai-integrated-development-environment-and-the- finn -workflow (abgerufen am 22. Mai 2024).**
30. **AMD. UltraSclae+ ZCU104. Online verfügbar: https://www.xilinx.com/products/boards-and-kits/zcu104.html (aufgerufen am 22. Mai 2024).**

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Swin-Fake: Ein Transformator für Konsistenzlernen Deepfake-Video-Detektor

Liang Yu Gong *,†,‡ , Xue Jun Li *,†,‡ und Peter Han Joo Chong‡

Fakultät für Elektro- und Elektroniktechnik, Auckland University of Technology, Auckland 1010, Neuseeland; peter.chong@aut.ac.nz *
Korrespondenz: liangyu.gong@autuni.ac.nz (LYG); xuejun.li@aut.ac.nz (XJL)
† **Aktuelle Adresse: 6 Saint Paul Street, Auckland 1010, Neuseeland.**
‡ **Diese Autoren haben zu gleichen Teilen zu dieser Arbeit beigetragen.**

Zusammenfassung: Deepfake ist **in den letzten Jahren zu einer aufstrebenden Technologie geworden, die mit ihren illegalen Anwendungen die Cybersicherheit beeinträchtigt .** Die meisten Deepfake-Detektoren verwenden CNN-basierte Modelle wie das Xception **Network, um echte von gefälschten Medien zu unterscheiden. Ihre Leistung bei datensatzübergreifenden Verfahren ist jedoch nicht optimal, da sie im aktuellen Stadium unter Überanpassung leiden. Daher wird in diesem Artikel eine Methode zum Lernen räumlicher Konsistenz vorgeschlagen, um dieses Problem in dreierlei Hinsicht zu lösen. Erstens haben wir die Auswahl der Datenerweiterungsmethoden auf 5 erhöht, was mehr ist als die Datenerweiterungsmethoden unserer vorherigen Studie . Insbesondere haben wir mehrere gleiche Videobilder eines Videos aufgenommen und zufällig fünf verschiedene Datenerweiterungen ausgewählt, um unterschiedliche Datenansichten zu erhalten und die Eingabevielfalt zu bereichern. Zweitens haben wir Swin Transformer als Merkmalsextraktor anstelle eines CNN-basierten Backbones gewählt, was bedeutet, dass unser Ansatz es nicht für nachgelagerte Aufgaben nutzte und diese Daten mithilfe eines End-to-End-Swin Transformer kodieren konnte , um die Korrelation zwischen verschiedenen Bildausschnitten zu lernen. Schließlich wurde dies in unserer Studie mit Konsistenzlernen kombiniert, und Konsistenzlernen konnte mehr Datenbeziehungen bestimmen als überwachte Klassifizierung. Wir untersuchten die Konsistenz der Merkmale von Videobildern, indem wir ihre Kosinusdistanz berechneten, und wendeten den traditionellen Kreuzentropieverlust an, um diesen Klassifizierungsverlust zu regulieren. Umfangreiche In-Dataset- und Cross-Dataset-Experimente zeigten, dass** Swin-Fake bei einigen Open-Source-Deepfake-Datasets, darunter **FaceForensics++, DFDC, Celeb-DF und FaceShifter, relativ gute Ergebnisse erzielen konnte .** Durch den Vergleich unseres Modells mit mehreren Benchmark- **Modellen zeigt unser Ansatz eine**

Schlüsselwörter: Konsistenzlernen; Swin Transformer; Konsistenzverlust; Deepfake-Erkennung

check for updates

Zitat: Gong, LY; Li, XJ; Chong, PHJ Swin-Fake: Ein auf einem Konsistenzlerntransformator basierender Deepfake-Videodetektor. Electronics **2024,** 13, **3045. https://doi.org/10.3390/Elektronik13153045**

Wissenschaftliche Herausgeber: Petia Georgieva, Krasimir Tonchev und Agata Manolova

Erhalten: 1. Juli 2024
Überarbeitet: 26. Juli 2024
Akzeptiert: 29. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https://creativecommons.org/licenses/by/4.0/).

1. Einleitung

Sowohl die Generierung als auch die Erkennung von Deepfakes wurden in einem Angreifer-Verteidiger-Spiel von 2019 bis 2021 weiterentwickelt . Mit der rasanten Entwicklung von GAN-basierten Methoden **[1,2] und Diffusionsmodellen (DMs) [3] sorgen die leistungsstarken generativen Fähigkeiten dieser Modelle mit erweiterter Nachbearbeitung, wie z. B. Farbmischung, dafür, dass synthetische Gesichter re Daher ist es heutzutage schwieriger, die Authentizität eines einzelnen Videobilds in einem Deepfake-Video mit dem menschlichen Auge zu erkennen.**

Obwohl Detektoren auf Basis von Convolutional Neural Networks (CNN) [4–6] einen bemerkenswerten Einfluss auf die Arbeit zur Klassifizierung von Deepfakes hatten, leiden sie immer noch unter der großen Abhängigkeit von Trainingsdaten, was dazu führt, dass die Genauigkeit oder der Bereich unter der Kurve (AUC) bei Experimenten mit mehreren Datensätzen erheblich abnimmt. Daher werden Deepfake-Detektoren in den letzten Jahren eher als Konsistenzlernmethoden entwickelt, um die räumliche oder zeitliche Konsistenz, Selbstkonsistenz oder Inkonsistenz zwischen Darstellungen oder Patches zu untersuchen. Darüber hinaus haben sich Vision Transformers [7,8] als leistungsstarke Anwendungen erwiesen, insbesondere bei der Bildklassifizierung . Transformer-basierte Deepfake-Detektoren werden in diesem Bereich zum Mainstream, da sie über einzigartige Aufmerksam

Machine Translated by Google

zeitliche Merkmale. Im Vergleich zu anderen Vision Transformers [9,10] kann Swin Transformer die Klassifizierungsleistung durch die Einstellung von Fensterpartitionierung, zyklischer Verschiebung und relativen Positionsbias-Mechanismen erheblich verbessern.

In diesem Artikel schlagen wir ein einfaches Konsistenzlern-Framework vor, das Swin Transformer [11] als Merkmalsextraktor verwendet , **um die Konsistenz zwischen aufgenommenen Videobildern explizit zu extrahieren und zu lernen. Im Vergleich zu herkömmlichen Deepfake-Detektoren enthält unser Ansatz auch einen Datenvorverarbeitungsblock und ein Backbone-Modell zur Merkmalsextraktion, aber unser Verlust kombiniert den Konsistenz- und Klassifizierungsverlust der Bilder. Was die Datenvorverarbeitung betrifft, haben wir ein Video in mehrere Bilder als Gruppe aufgeteilt, um Deepfake-Videos zu klassifizieren, was als Ensemble-Technik angesehen werden kann. Im Gegensatz zu anderen Ensemble-Lerntechniken haben wir die Merkmalsextraktion erweitert, um dem Modell zu helfen, verschiedene Fälschungsmerkmale mit zufälligen und unterschiedlichen Datenerweiterungsmethoden zu lernen. Darüber hinaus verwenden herkömmliche Deepfake-Detektoren immer einige unveränderte Datenerweiterungsmethoden wie grundlegende Transformation und Normalisierung; für unseren Datenvorverarbeitungsblock haben wir jedoch mehrere erweiterte Datenerweiterungsmethoden gewählt, um die Zufälligkeit der Datenerweiterung zu erhöhen und zu beweisen, dass unabhängig davon, wie stark die Daten erweitert werden, dieselbe Datenkategorie mit guter Konsistenz beibehalten werden kann. Wie in Abbildung 1 dargestellt, wurde jedes Video gleichmäßig in Einzelbilder aufgeteilt, um sicherzustellen, dass der Gesichtsbereich gleichmäßig über die gesamte Videodauer verteilt ist . Unser Datenvorverarbeitungsblock enthält auch mehrere Datenerweiterungen und wird zufällig auf jedes aufgenommene Einzelbild angewendet, um verschiedene Eingaben anzureichern. Anschließend extrahiert ein gemeinsamer Transformer-basierter Encoder Darstellungen und berechnet die Ähnlichkeit von Paaren, indem er Kombinationen vornimmt, um einen Konsistenzverlust zu erzielen. Schließlich weist ein überwachter Klassifizierer jedem extrahierten Einzelbildmerkmal die Bezeichnung zu.**

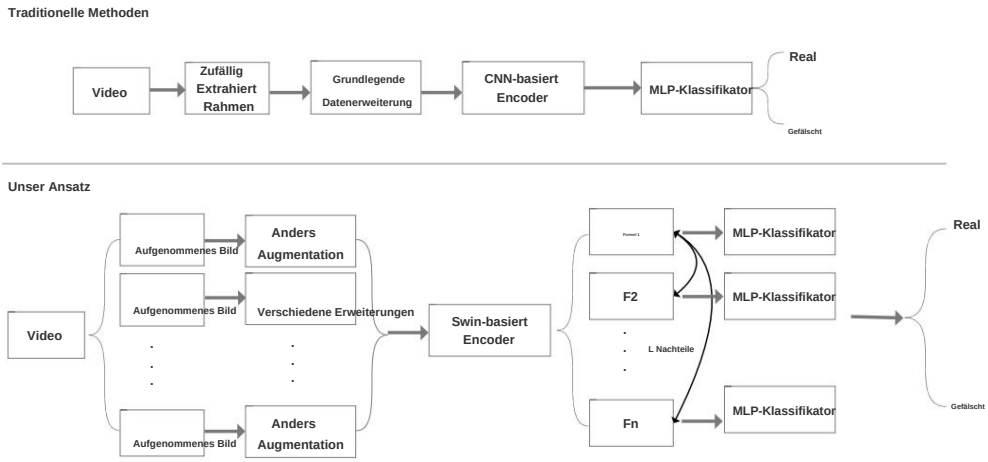


Abbildung 1. Darstellung des Swin-Fake-Frameworks im Vergleich zum herkömmlichen Deepfake-Video-Detektor-Framework.

Unsere Methode weist im Vergleich zu herkömmlichen Deepfake- Detektoren drei wesentliche Verbesserungen auf: (1) Wir verwenden fünf Datenerweiterungsmethoden, darunter Basistransformation, Random Erasing [12], Random Augmentation, Random Crop und DFDC Selim [13], die zufällig angewendet werden und eine Überanpassung in der Eingabephase weitgehend vermeiden. (2) Als Encoder wird Swin Transformer gewählt, der die Korrelationen zwischen verschiedenen aufgeteilten Bildausschnitten kombiniert. (3) Der endgültige Gesamtverlust wird durch zwei Komponenten bestimmt, nämlich Konsistenzverlust und Kreuzentropieverlust, und die Entscheidung wird durch Konsistenzlernen gewährleistet.

2. Verwandte Arbeiten

2.1. Deepfake-Erkennung

Die Deepfake-Erkennung wird immer als binäre Klassifizierung oder Mehrklassen-Klassifizierungsarbeit definiert, da das Modell grundsätzlich in der Lage sein muss, die Medien als echt oder gefäl Das beliebteste Framework ist die Verwendung von Vanilla Xception [14] oder ResNet [15] als Backbones

Machine Translated by Google

um kontrastives Lernen zu kodieren und dann zu kombinieren oder die Konsistenz von Darstellungen zu untersuchen. Capsule Network [4] ist ein effektives, leichtgewichtiges Netzwerk, das „Inverse Graphic“ -Probleme lösen kann und einen dynamischen Routing-Algorithmus zur Neugruppierung der extrahierten Features vorsc CORE [16] ist eine auf Xception Network basierende Methode des Konsistenzlernens, die die Anzahl der Datenerweiterungen verbessert, um die Konsistenz des Modells zu beweisen. Youtu-Lab [17] schlug eine dual-kontrastive Lernarchitektur vor, um die Authentizität von Gesichtern ebenfalls anhand verschiedener Datenerweiterungen zu unterscheiden. Ensemble-Learning-Deepfake-Detektoren könnten die Klassifizierungsleistung ebenfalls erheblich verbessern und wurden in der jüngsten Forschung häufig angewendet; beispielsweise schlug Zhang [18] eine heterogene Ensemble-Lernmethode für Merkmale vor, um zuerst Graustufenmerkmale, Spektrummerkmale und Texturmerkmale aus Bildern zu extrahieren und diese dann durch einen Abflachungsprozess in einen Ensemble-Merkmalsvektor zu integrieren . Die Autoren von [19] entschieden sich , mithilfe von DeepfakeStack mehrere hochmoderne Deepfake-Klassifizierungsmodelle zu kombinieren, um einen verbesserten zusammengesetzten Klassifikator zu erstellen, der in einigen Deepfake-Datensätzen eine Genauigkeit von 99,65 % erreichen könnte. Unsere vorgeschlagene Methode, Swin-Fake, zielt jedoch darauf ab, die Fähigkeit von Swin Transformer zu untersuchen, gefälschte Merkmale wie PRNU- Rauschen und Randfehlanpassungen zu extrahieren und dann verschiedene Frames eines Videos zu ex

2.2. Swin Transformer

Swin Transformer [11] ist ein hierarchischer Vision Transformer, der Multi-Head Attention Blocks (MSAs) zu Window Multi-Head Attention Blocks (W-MSAs) und Shifted Window Multi-Head Attention Blocks (SW-MSAs) verbessert. Er schichtet Features unterschiedlicher Größe übereinander, indem er Fenster unterschiedlicher Größe einrichtet , was sich vom herkömmlichen Vision Transformer [7] unterscheidet und mehr Rechenkosten einsparen könnte als der herkömmliche Vision Transformer. Normalerweise hat das Eingabebild eine Größe von 3 × 3 × 224 und die Ausgabe-Feature-Größe beträgt 7 × 7 × 768. Dieses Framework enthält eine Patch-Portion-Schicht, eine lineare Einbettungsschicht und zwölf Swin Transformer-Blöcke zum Extrahieren von Features, und die Features werden mit verschiedenen Patch-Einbettungen korreliert. Diese Netzwerkstruktur kann das Gefühl von Konsistenz und Inkonsistenz verschiedener Patches effektiver simulieren, wenn das menschliche Auge verschiedene Deepfake-Gesichtsbereiche beobachtet und scannt. Darüber hinaus haben wir vorhandene Arbeiten zu Transformer-basierten Deepfake-Detektoren untersucht und überraschenderweise festgestellt, dass es nicht viele Arbeiten gibt, die sich auf räumliches Konsistenzlernen mit einem Transformer-Backbone konzentrieren. Die meisten von ihnen investieren mehr Aufwand in die Entwicklung von Blöcken zur Extraktion zeitlicher Merkmale, um mehr Informationen zur zeitlichen Konsistenz zu erhalten. Kürzlich schlugen Ilyas et al. [20] ein AVNet zur Erkennung audiovisueller Medien mit Swin Trans

2.3. Datenerweiterung Im

Bereich der Computervision ist die Anwendung einer Datenerweiterung wichtig. Sie ist der erste Schritt zur Vermeidung von Überanpassung und eine effektive Methode zur Erhöhung der Datenmenge. In diesem Dokument verwenden wir nur fünf Bildbearbeitungsmethoden, um Pixelwerte und -positionen zu ändern. Die Basistransformation wird verwendet, um Pixelnormalisierung und Größenänderung Random Erasing [12] ist eine Datentransformation zum Ausschneiden, die zufällig einen rechteckigen Block des Gesichtsbereichs auswählt und seine Farbe in völlig Schwarz ändert. Random Crop wird verwendet, um den Gesichtsbereich in einem zufälligen Verhältnis zuzuschneiden und ihn dann auf eine bestimmte Höhe und Breite zu skalieren. DFDC Selim ist eine komplexe Datenerweiterung, die Gaußsches Rauschen, Gaußsche Unschärfe, zufällige Verschiebung und Skalierung enthält und sich bereits als die nützlichste Datenerweiterung bei der Klassifizierung von Deepfakes erwiesen hat [16]. Random Augmentation wählt zufällig entweder Random Erasing, Random Horizontal Flip oder Random Resize Crop aus. Im Vergleich zum Design herkömmlicher Deepfake-Detektoren haben wir die Vielfalt und Zufälligkeit der Eingabedaten drastisch verbessert und können zur anschließenden Überprüfung der Ähnlichkeit zwischen verschiedenen Videobildern eines Videoclips verwendet werden.

Machine Translated by Google

3. Vorgeschlagene

Methoden In dieser Arbeit skizzieren wir die Pipeline zur Erkennung von Deepfake-Bildern und unterteilen sie in vier Teile: Datenvorverarbeitung (einschließlich Gesichtserkennung, Zuschneiden und Datenerweiterungen), Merkmalsextraktion, Konsistenzberechnung und Vorhersage durch Klassifikatoren. Wie in Abbildung 2 dargestellt, erfassen wir zunächst die Frames eines erkannten Videos gleichmäßig und wählen zufällig Datenerweiterungen aus und wenden sie an, um unterschiedliche Datenansichten zu erstellen. Nachdem wir die Merkmale durch Swin Transformer extrahiert haben, berechnen wir separat die Kosinusdistanzen zwischen Darstellungen in Kombinationen und su Um das Problem der unausgewogenen Klassen zu vermeiden, kombinieren wir echte Videos aus anderen Datenbanken mit unserem Trainingsset, was für die weitere Auswertung sehr wichtig ist. Darüber hinaus entscheiden wir uns für eine Hochskalierung der echten Videobilder. Wenn beispielsweise jedes gefälschte Video gleichmäßig in acht Bilder aufgeteilt ist, wird das echte Videobild entspre

Dies könnte weitgehend verhindern, dass die vorhergesagte Ausgabe des trainierten Modells in **derselben Klasse verbleibt .**

Algorithmus 1 Der Arbeitsablauf von Swin-Fake Input: Der

vortrainierte Swin-Transformator: f(x); Videosequenzen: X. **Teile jedes Video gleichmäßig in** M **Frames auf.**

Gesichtszuschneiden und Breite und Höhe erweitern, um Gesichtsbilder zu erhalten.

Wählen Sie Datenerweiterungsmethoden nach dem Zufallsprinzip aus, um eine Überanpassung zu vermeiden. Für i = 1 bis

M verwenden Sie die Eingabesequenzen X, **um den Satz zu gruppieren.**

Berechnen Sie den Aufmerksamkeitswert zwischen verschiedenen geteilten Patch-Einbettungen.

Berechnen Sie f(x) **durch Vorwärtsausbreitung, um 2D-Merkmale** Fx zu erhalten. **2 Berechnen Sie den Cosinus-Abstand** LCos **zwischen** C- **Paaren mithilfe von Gleichung (3).**

$$LCons = \ddot{y} \frac{1}{M} \sum_{1}^{M} \text{Mittelwert (LCos)}$$

Verwenden Sie die lineare Ebene, um Fx **auf 2 Klassen zu projizieren.**

$$LCE = \ddot{y}\ \ddot{y}\ Ltot = \sum_{1}^{N} [y \log(p) + (1 \ddot{y}\ y)\log(1 \ddot{y}\ p)]$$

ÿLCons + **(1 ÿ** ÿ)LCE **Führen Sie eine**

Rückwärtsausbreitung durch und speichern Sie die Gradienten für jede Schicht. Ende für Ausgabe: ein

trainiertes Swin-Fake-Modell f(x).

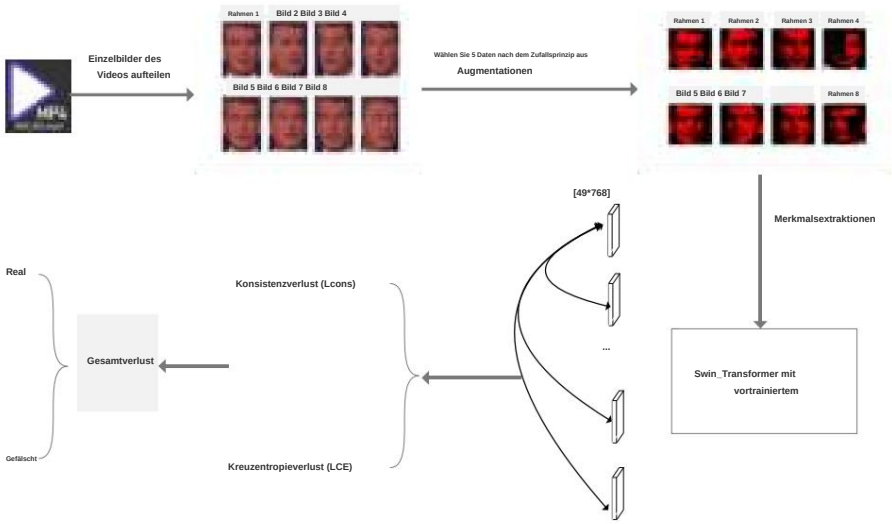


Abbildung 2. Erklärung der Swin-Fake-Architektur. Wir teilen zunächst M **Frames eines Videos auf und wenden für jedes Frame unterschiedliche Datenerweiterungen an. Dann werden sie an einen gemeinsam genutzten Swin Transformer-Encoder gesendet. Nachdem** M Darstellungen erhalten wurden, werden sie gesendet, um Klassifizierungsverlust **und Konsistenzverlust separat zu berechnen .**

Machine Translated by Google

3.1. Vorverarbeitungsarbeit

Zuerst geben wir einen Eingabestapel von B- **Videos ein und teilen jedes Video gleichmäßig in M Videobilder auf, indem wir die Gesamtbildzahl jedes Videos berechnen. In die Vorverarbeitungsphase ist auch ein Gesichtsdetektor, „SCRFD“ [22] ,** eingebettet , **um Gesichtsbereiche zu lokalisieren, die Gesichter auszurichten und sie zuzuschneiden, um spezifische Datensätze zu erstellen. Genauer gesagt basiert dieser Gesichtsdetektor auf Insightface und verwendet spezifische Sample Redistribution (SR)- und Computation Redistribution (CR)-Methoden zur Gesichtserkennung. „scrfd 10g kps.onnx“ ist eine der Modellgewichtsdateien, die diese Gesichtserkennungsmethode [22] verwenden , und kps stellt die 5 in diesem Modell erkannten Gesichtsschlüsselpunkte dar. Um einige Hintergrundinformationen zu erhalten, erweitern wir die Breite und Höhe der zugeschnittenen Bilder um 10 %. Anschließend werden die gesamten** M **×** B- **Frames ordentlich angeordnet, dienen als Eingaben für den Swin Transformer und werden alle auf 3 × 224 × 224 skaliert. Da wir fünf Datenerweiterungen in einem Transformationssatz** T **auf** M **×** B **-Frames anwenden und zufällig auswählen,** haben **die Eingabevektorsätze** I **eine Größe von** [B, **M, 3, 224, 224], was bedeutet, dass jede Gruppe** I

$$I = \{\{x1, \mathbf{x2. . . xM}\}\mathbf{1} . . . \{\mathbf{x1}, x2. . . xM\}\mathbf{B}\}, \quad (1)$$

wobei xn **jede Einzelbildansicht darstellt,** B **einen Videostapel darstellt, {x1, x2 . . . xM}B die B-te Gruppe von Datenansichten darstellt** und M **angibt, wie viele Einzelbilder gleichmäßig aufgeteilt sind.**

Sobald die Eingabedatenansichten vorliegen, werden fünf Datenerweiterungsmethoden zufällig auf diese Datenansichten angewendet. Diese erweiterten Tensoren sind die endgültigen Eingaben des Swin-Fake-Modells, wie in Abbildung 3 dargestellt.



Abbildung 3. Beispiele für zufällig erweiterte Daten. Für das zufällige Löschen legen wir den Wert für das zufällige Löschen auf 0,8 fest, der Maßstab liegt zwischen 0,02 und 0,2 und das Verhältnis der gelöschten Fläche zwischen

Für den zufälligen Zuschnitt haben wir den Maßstab für den zufälligen Zuschnitt von 0,77 bis 1 und das Verhältnis für den zufälligen Zuschnitt von 0,9 bis 1,1 festgelegt.

3.2. Encoder-Netzwerk

Wir verwenden Swin Transformer [11] aufgrund seiner einzigartigen hierarchischen Struktur **als unseren Merkmalsextraktor** f . **Er enthält mehrere gestapelte Transformer-Blöcke, von denen die Blöcke auf niedrigerer Ebene einige Farb- und Kanteninformationen erfassen können und die Blöcke auf höherer Ebene zur Analyse der semantischen Informationen verwendet werden können. Nach Anwendung des Ansatzes der verschobenen Fensterpartitionierung und der linearen Einbettung können wir die Dimension des Eingabemerkmals ändern und die Vektoren für Abfrage (Q), Schlüssel** (K) **und Wert** (V) **separat erhalten. Die Abfragen** (Q) **und Schlüssel** (K) **haben dieselbe Dimension d. Indem wir die Skalarprodukte der Abfrage** (Q) **mit allen Schlüsseln** (K) berechnen und jedes durch **ÿ** d **dividieren , ist der erste Schritt der Aufmerksamkeitskorrelation abgeschlossen. Dann fügen wir den Bias hinzu und wenden eine Softmax-Funktion an, um die Ausgabe in den Bereich von 0 bis 1 zu quetschen und einen Wertwert zu erhalten. Diese Aufmerksamkeitsberechnungsmethode zielt darauf ab, die ge**

Die Berechnung der Selbstaufmerksamkeit wird wie folgt geschrieben:

$$\text{Achtung (Q, } \mathbf{K,} \text{ V) = So f tMax (QKT/ } \mathbf{\ddot{y}} \text{ d + B)V, } \mathbf{wobei\ 1/\ddot{y}} \quad (2)$$

d **der Skalierungsfaktor von Q,** K **und** B **die Verzerrung von Q, K, V** ist.

Wenn die Eingabemenge I durch den Encoder geht und in B **×** M **Darstellungen abgebildet wird , ist die Darstellung** f(xn) **ein zweidimensionaler Vektor mit einer Größe von [49, 768].**

Machine Translated by Google

In der Zwischenzeit wird jede M- **Darstellung als eine Gruppe festgelegt und in die Berechnung des Konsistenzverlusts eingespeist. Nachdem wir eine lineare Schicht auf den letzten grundlegenden Swin-Block angewendet haben, senden wir die Darstellungen zur Berechnung des Cross-Entropy-Verlusts zur Klassifizierung. Daher muss jeder Frame-Stapel mehrere datenerweiterte Operationen durchlaufen, und dann können wir die extrahierten niedrigdimensionalen Merkmale erhalten (siehe Abbildung 4). Schließlich werden zwei Ausgabevektoren berechnet, einer für die Berechnung des Konsistenzverlusts und ein anderer, der von der linearen Schicht für den Klassifizierungsverlust betrieben wird. Eine detaillierte Erklärung dies siehe Abschnitt 3.3.**

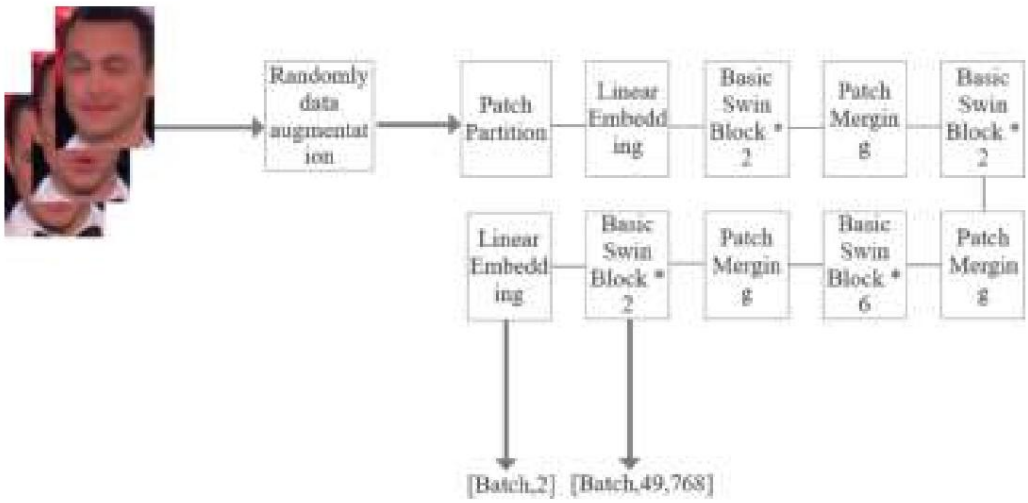


Abbildung 4. Die Architektur des Swin Transformer-Backbones. ÿ stellt dar, wie viele grundlegende Swin-Blöcke im Backbone angewendet werden. Für jeden Eingabedatenansichtssatz in unserem Modell gibt es zwei Ausgabevektoren, nämlich die extrahierten niedrigdimensionalen Merkmale mit einer Größe von [M **× 49 × 768] und der vorhergesagten Wahrscheinlichkeit** [M **×** 2].

3.3. Verlustfunktionen

Die Verlustfunktion dieses Modells besteht aus zwei Komponenten: Klassifizierungsverlust und Konsistenzverlust. Der Konsistenzverlust wird auf jede Gruppe von M **extrahierten Darstellungen angewendet und wir subtrahieren ihre Distanzen. Darüber hinaus übernehmen wir die Kosinusdistanz, um das Konsistenzverlustdesign zu vervollständigen, weil wir nicht jedes Darstellungspaar dazu zwingen, genau gleich zu sein, sondern sich immer ähnlicher zu werden. Die Berechnung der Kosinusdistanz jedes Darstellungspaars kann wie folgt geschrieben werden:**

$$LCos = (1 \text{ ÿ } fxi \cdot fxt) \quad (3)$$

wobei xi **und** xt das **i-te und t-te Frame der Datenansichten darstellen** und f (xi) **und** f (xt) **die von einem gemeinsam genutzten Encoder extrahierten normalisierten Darstellungen darstellen .**

Da wir erweiterte Videoclips mit der Menge der B- **Gruppen als Eingabe für unser Modell erhalten haben, müssen wir auch die Ähnlichkeit zwischen diesen Merkmalen der aufgenommenen Videoclips berechnen . Ziel ist es, einige falsch unterschiedene Ergebnisse von harten Proben durch Verkürzung ihrer Kosinusdistanzen zu beheben, und wir betrachten dies als Konsistenzverlust.**
Ein weiterer Faktor, der den Gesamtverlust beeinflusst, ist der Konsistenzverlust, der dazu dient, die mittleren Kosinusabstände von B **Batch-Videos zu berechnen. Insbesondere können die extrahierten** M **Darstellungen eines Videos** durch $\frac{2}{M}$ die obige Gleichung C **Ähnlichkeitspaare erzeugen. Die einzelnen Die Konsistenz eines Videos kann als Mittelwert der** C- **Paarähnlichkeit** $\frac{2}{M}$ **ausgedrückt werden, und der Gesamtkonsistenzverlust von** B **-Batch-Videos erfordert die Berechnung des Durchschnittswerts aller berechneten Einzelkonsistenzen. Daher kann der Konsistenzverlust wie folgt ausgedrückt werden:**

$$LCons = \frac{1}{B}\sum_{1}^{B} \text{ÿ Mittelwert}(LCos) = \frac{1}{B}\sum_{1}^{B} \text{ÿ Mittelwert}(1 \text{ ÿ } fxi \cdot fxt) \quad (4)$$

Machine Translated by Google

wobei B **die Batchgröße des Videos darstellt.**

Der Hauptzweck der Einrichtung eines Konsistenzverlusts besteht darin, zu beweisen, dass das Modell unabhängig davon, wie realistisch ein bestimmtes Deepfake-Videobild ist, in der Lage ist, seinen räumlichen Abstand zu anderen gefälschten Bildern zu verkürzen. Andererseits validiert der Konsistenzverlust die Klassifizierungskonsistenz verschiedener Bilder eines Videos. Da die Klassifizierungsarbeit untrennbar mit überwachten Beschriftungen verbunden ist, wenden wir den standardmäßigen Kreuzentropieverlust als Klassifizierungsverlust an und formulieren ihn wie folgt:

$$LCE = \ddot{y} \overset{M\times B}{\underset{1}{\ddot{y}}} [y \log(p) + \mathbf{(1 \ddot{y} y)log(1 \ddot{y}} p)] \quad \textbf{(5)}$$

wobei y **die Grundwahrheitsbezeichnung,** p **die vorhergesagte Wahrscheinlichkeit und** M × B die **Gesamtzahl der Videobilder darstellt .**

Da es sich bei unserem Modell um ein Konsistenzlernmodell handelt, verwenden wir zwei Verlustfunktionen (Konsistenzverlust und Kreuzentropieverlust) in der linearen Kombination, um unseren Gesamtv Wir legen jedoch den Hyperparameter *ÿ* **als Gewichtungsausgleich fest, um zu beobachten, welcher Anteil dieser beiden Verlustfunktionen den minimalen Gesamtverlust und die höchste Genauigkeit erreichen In Abschnitt 4.4 führen wir einen Ablationstest durch, um zu beobachten, welcher Wert dieses Hyperparameters die meisten Stichproben vorhersagen kann, indem wir den Gewichtsparameter durchlaufen und in jedem Experiment 0,1 hinzufügen. Wir hoffen, das am besten passende Gewicht zu finden, um diese beiden Verlustfunktionen auszugleichen, und wir hoffen nicht, dass der Gesamtverlust stark zu einem der Ergebnisse der Verlustfunktionen tendiert.**

$$Ltot = \ddot{y}LCons + \mathbf{(1 \ddot{y}} \ddot{y})LCE \quad \textbf{(6)}$$

wobei *ÿ* **der Hyperparameter zum Ausgleich von Kreuzentropieverlust und Konsistenzverlust ist und zwischen 0 und 1 liegt.**

4. Experimente

4.1. Datensätze

Um unsere Methode zu evaluieren, haben wir vier gängige Datensätze in diesem Bereich ausgewählt: FaceForensics ++ [23], DFDC [24], Celeb-DF [25] und FaceShifter [26]. Insbesondere wurden die ersten beiden Datensätze für die Trainingsphase verwendet, und Celeb-DF V2 und FaceShifter wurden für datensatzübergreifende Experimente verwendet, um die Generalisierungsfähigkeit unseres Modells zu beweisen. FaceForensics++ ist ein bahnbrechender groß angelegter Datensatz zur Gesichtsmanipulation, der 1000 echte Videos und 4000 gefälschte Videos enthält, die von Face2Face, FaceSwap, Neural Textures und Dee DFDC ist der größte dieser vier Datensätze und enthält 119.197 von echten Schauspielern gefilmte Videos. Im Vergleich zu den beiden vorherigen Datensätzen besteht Celeb-DF aus 590 Original-YouTube-Videos und 5639 manipulierten Videos mit erweiterter Farbmischung, die für das menschliche Auge schwer zu unterscheiden sind. FaceShifter ist eine Unterklasse der FaceForensics+ +-Datensätze und enthält 1000 gefälschte Videos mit unterschiedlichen Aufnahmeszenarien. Angesichts der unausgewogenen Trainingsdaten von FaceForensics++ haben wir 2000 echte Videos aus der DFDC- Datenbank zufällig mit echten FF++-Videos kombiniert, um dieses Problem zu lösen. Was das DFDC-Training betrifft, haben wir uns entschieden, einige gefälschte Videobilder herunterzurechnen,

4.2. Implementierungsdetails

Für die Vorverarbeitung der Daten extrahierten wir M- **Frames aus jedem Video, indem wir für „scrfd 10g kps.onnx“ einen Konfidenzschwellenwert von 0,5 und einen NMS-Schwellenwert von 0,5 festlegten. Außerdem wurden die erkannten zugeschnittenen Gesichter mit Begrenzungsrahmen in Breite und Höhe um 1,2 vergrößert, um sicherzustellen, dass zu den Eingaben des Modells Hintergrundinformationen vorhanden waren. Was die Einstellungen zur Datenerweiterung betrifft, legten wir eine Wahrscheinlichkeit von 80 % für die Löschung des Gesichtsbereichs fest, den Bereich des Proportions des gelöschten Bereichs im Vergleich zum Eingabebild legten wir auf 0,02 bis 0,2 fest und den Bereich des Seitenverhältnisses des gelöschten Bereichs legten wir im Vorgang „Zufälliges Löschen“ auf 0,5 bis 2,0 fest . Bei der Erweiterung „Zufälliges Zuschneiden“ legten wir die Unter- und Obergrenzen der Skala auf 0,77 und 1 fest, was bedeutet, dass der zufällige Zuschneidebereich vor der Größenänderung entsprechend zwische**

Machine Translated by Google

Es wurden keine menschlichen Gesichter aus dem zuletzt erkannten Index-Frame kopiert, anstatt sie direkt an den Eingabedaten-Ansichtssatz zu senden.



Abbildung 5. Einige von „scrfd 10g kps.onnx“ verarbeitete Deepfake-Beispiele, die den ursprünglichen Input für unser Modelltraining darstellen.

Für das Modelltraining haben wir die Video-Mini-Batch-Größe auf 4 mit einer Bildgröße von 224 × 224 eingestellt und dann einen AdamW-Optimierer [27] mit einem Gewichtsabfall von 0,05 und einer Lernrate von 0,0001 verwendet. Um dieses Modell effektiv zu trainieren, haben wir das Training vorzeitig abgebrochen, wenn sich die Validierungsgenauigkeit in den ersten 5 Epochen nicht verbesserte, und die Gesamtepoche wurde auf 30 eingestellt. Darüber hinaus wurden alle Abbruchraten auf 0,2 eingestellt, um ein Überanpassen in unserem Experiment zu vermeiden. Alle unsere Modelle verwendeten swin-tiny-patch4-window7-224.pth und wurden auf ImageNet-1K als Backbone vorab trainiert . Die Ausgabefunktionen wurden von der linearen Ebene mit einer Größe von [M, **49, 768] und den Vektoren mit einer Größe von** [M, 2] extrahiert. **Wir haben zunächst ein ausgewogenes Gewicht von 0,5 gewählt, da wir den Konsistenzverlust zu Beginn als genauso wichtig erachteten wie den Klassifizierungsverlust.**

4.3. In-Dataset-Studie **In**

diesem Abschnitt bewerten wir die Validierungsleistung unseres vorgeschlagenen Modells in denselben Quellen wie die Trainingsdatasets. Für die In-Dataset-Experimente bewerteten wir DFDC und FaceForensics++, zwei relativ große Datasets, da wir festgestellt haben, dass Swin Transformer immer eine große Menge an Trainingsdaten benötigt, um bei Trainingsexperimenten ein gutes Validierungsniveau zu erreichen; andernfalls weist es in einigen kleinen Trainingsdatasets eine relativ niedrige anfängliche Validierungsgenauigkeit auf. Und der anfängliche Hyperparameter wurde standardmäßig auf 0,5 eingestellt. Wir verglichen unsere Methode des Konsistenzlernens mit drei Basislinien: Capsule Network, STIL [28] und ISTVT [10]. Wie in Tabelle 1 gezeigt , weist Capsule Network die niedrigste Genauigkeit auf und ISTVT schneidet besser ab als STIL. Unser Ansatz schneidet viel besser ab als Capsule Network und die In-Dataset-Genauigkeit ist etwas höher als die von STIL und ISTVT.

Machine Translated by Google

Tabelle 1. Genauigkeit der Deepfake-Erkennung. Die besten Genauigkeitsergebnisse in Datensätzen sind fett gedruckt. Außer Capsule Network, das reproduziert wird, die Ergebnisse der anderen Methoden stammen aus [10,28].

Methoden	FF++				DFDC
	Verteidigung	F2F	FS	NT	
Kapselnetz	98,5 %	96,7 %	94,5 %	91,8 %	-
STIL	99,6 %	99,3 %	97,1 %	95,4 %	89,8 %
ISTVT	99,6 %	99,6 %	100%	96,8 %	92,1 %
Unser Ansatz DF:	99,9 %	99,6 %	100%	98,8 %	93,7 %

Deepfakes; F2F: Face2Face; NT: Neural Textures; FS: FaceShifter.

4.4. Ergebnisse des Ablationstests

Um die Bedeutung des Konsistenzzweigs unseres Modells zu testen, führten wir auch eine Ablationsexperiment (siehe Tabelle 2) zur Swin Transformer Deepfake-Bildklassifizierung mit unser Ansatz (mit überwachtem Konsistenzlernen). Wir fanden heraus, dass Konsistenzlernen kann die Genauigkeit effektiv erhöhen, da unser Ansatz Konsistenzverlust verwendet um die Distanzen zwischen den einzelnen Frames des Videos zu verkürzen (die Ähnlichkeit zu erhöhen), dass, wenn ein Frame durch den Klassifizierungszweig als gefälscht mit einem wirklich hohen Konfidenzwert, besteht eine hohe Wahrscheinlichkeit, dass andere harte Stichprobenrahmen berücksichtigt werden auch als Fälschung.

Tabelle 2. Vergleich der Genauigkeit von Ablationsexperimenten. Wir verwendeten Swin Transformer, um eine Führen Sie einen Binärtest für Deepfake-Bilder durch und vergleichen Sie seine Genauigkeit mit der unseres Ansatzes.

Methoden	FF++ (Binärtest)	DFDC
Swin-Transformator	89,4 %	91,4 %
Unser Vorgehen	94,3 %	93,7 %

Darüber hinaus ist das Bilanzgewicht $\ddot{y}$ eine weitere Variable, die über den endgültigen Gesamtverlust entscheidet; Daher führten wir ein Ablationsexperiment durch, um festzustellen, wie viel Gewicht erforderlich ist für den Konsistenzverlust die beste Leistung in diesem diskriminierenden Modell zu zeigen durch Setzen dieses Hyperparameters auf 0,4, 0,5, 0,6, 0,7 und 0,8. In diesem Ablationstest verwendeten wir immer noch FaceForensics++ als Datensatz zur Untersuchung des Ausgleichsgewichts, das die höchste vorhergesagte Genauigkeit. Darüber hinaus haben wir alle gefälschten Kategorien in einer Klasse zusammengefasst, um zu testen binäre Leistung anstatt mehrere Klassifizierungen zu testen. Da wir dies nicht eingestellt haben Gewichtsbalance als lernbarer Parameter in unserem Modell haben wir uns entschieden, die Gewichtsbalance zu durchlaufen. durch Hinzufügen von 0,1 pro Experiment ausgeglichen. Sobald wir die höchste Genauigkeit erreicht hatten, betrachteten wir Dieses Modell ist das modernste Modell und wir haben die trainierten Parameter dieses Modells verwendet um datensatzübergreifende Experimente durchzuführen. Wie in Tabelle 3 gezeigt, stellen wir fest, dass das Gewicht Die Balance-Einstellung von 0,7 zeigt die beste Leistung und wurde auf dem FaceForensics++ trainiert. Binärtest.

Tabelle 3. Genauigkeit des Ablationsexperiments auf verschiedenen Gewichtswaagen. Damit wird getestet, welches Gewicht Das Gleichgewicht unserer Verlustfunktion erzielt das beste Ergebnis.

$\ddot{y}$	0,4	0,5	0,6	0,7	0,8
Genauigkeit	92,6 %	93,9 %	94,1 %	94,3 %	93,7 %

4.5. Datensatzübergreifende Studie

Aufgrund der zahlreichen Fälschungsmethoden ist es wichtig, die Leistung des Modells über mehrere Datensätze hinweg zu bewerten. Um die Generalisierungsfähigkeit unseres Modells zu testen, haben wir unser Modell trainiert auf FaceForensics++ und testete es auf Celeb-DF, FaceShifter und DFDC. In Anbetracht dessen, dass Die Stichprobengrößen der meisten Deepfake-Datensätze für die reale und die gefälschte Klasse sind unausgewogen.

Machine Translated by Google

Wir haben uns auch für die Nutzung der Area Under the Curve (AUC) entschieden, um die Robustheit unserer Modell, wie in Tabelle 4 dargestellt.

Tabelle 4. Die AUC des datensatzübergreifenden Experiments. Wir haben unser Modell mit FF++ trainiert und getestet die ungesehenen Daten. Die besten AUC-Ergebnisse über mehrere Datensätze hinweg sind fett gedruckt. Die Ergebnisse der anderen Methoden sind aus [10,28–31].

Modelle	Promi-DF	Gesichtsschieber	DFDC	Durchschnitt
Kapselnetz	34,5 %	42,4 %	-	38,5 %
STIL	75,6 %	-	-	75,6 %
ISTVT	84,1 %	99,3 %	74,2 %	85,9 %
Gesichtsröntgen	79,5 %	92,8 %	65,6 %	79,3 %
FTCN	86,9 %	98,8 %	74,0 %	86,6 %
LipForensics	82,4 %	97,1 %	73,5 %	84,3 %
Unser Vorgehen	88,8 %	98,3 %	71,8 %	86,3 %

Wir fanden heraus, dass unser Ansatz nur bei Celeb-DF die besten Ergebnisse erzielen konnte. liegt daran, dass Celeb-DF erweiterte Nachbearbeitungsfunktionen wie Farbmischung enthält und es ist wirklich schwer herauszufinden, wann es in generierten Medien zu Grenzfehlanpassungen kommt. Um diesen Standpunkt zu beweisen, haben wir uns entschieden, Heatmaps unseres Modells und anderer CNN-basierter Erkennungsmodelle zu erstellen, um die Visualisierung der Fokusbereiche von Deepfakes durch das Modell zu erl Medien. Durch die Beobachtung der Heatmap unseres Modells haben wir festgestellt, dass sich unser Modell mehr konzentriert auf generierten menschlichen Gesichtsinhalten wie Lippen und Augen statt auf Randtexturen Informationen (siehe Abbildung 6). Durch die Überprüfung einiger Beispielvideos von FaceShifter stellten wir fest, dass Fälschungshinweise kommen immer im Gesichtsgrenzbereich vor, also in der Gesichtsregion, die Unser Modell achtet weniger darauf, aber die Fälschungshinweise von Celeb-DF finden sich in den Gesichtszügen, wie unnatürlich manipulierte menschliche Augen und Lippen. Darüber hinaus verglichen wir die Daten in Celeb-DF mit den Daten von DFDC, und wir fanden heraus, dass DFDC mehr Daten enthält mit verschiedene Aufnahmeszenarien, wie dunkle Innenhintergründe und Gesichter mit extremen Posen. Dies macht es für unser Modell wirklich schwierig, Konsistenz zu erreichen und nützliche Informationen zur Unterscheidung von echten und gefälschten Medien, da wir das von trainierte Modell verwendet haben FaceForensics++, und es enthielt mehr menschliche Gesichter mit normalen Posen und hellen Aufnahmen Szenarien, die der Celeb-DF-Datenverteilung ähnlich sind.

Darüber hinaus verglichen wir auch die Transformer-basierte Methode ISTVT als Benchmark Modell und erzielte bei FaceShifter und DFDC etwas bessere AUC-Ergebnisse (siehe Abbildung 7). Dies liegt daran, dass die Autoren, die ISTVT verwendeten, einen Selbstsubtraktionsmechanismus anwendeten und zeitliche Selbstaufmerksamkeit in ihrer Arbeit, was unserem Ansatz fehlt. Die Selbstsubtraktion Mechanismus kann sich mehr auf die Verzerrung zwischen den Bildern konzentrieren, was eine weitere wichtige Technik ist zur Deepfake-Erkennung. Es erfordert die Anwendung eines Eingabetokens vor der Projektion zu Abfrage- und Schlüsselvektoren. Es generiert Merkmalsresiduen, um die ursprünglichen Merkmale zu ersetzen. Dennoch erzielt unser Modell im Vergleich zu anderen bisherigen hochmoderne Modelle wie STIL und Face X-ray.

Machine Translated by Google



Abbildung 6. Vergleichs-Heatmap zwischen anderen Deepfake-Erkennungsmethoden und unserem Ansatz. Aus dieser Abbildung geht hervor, dass sich die meisten Deepfake-Detektoren auf Randinformationen der generierten gefälschten Medien konzentrieren.

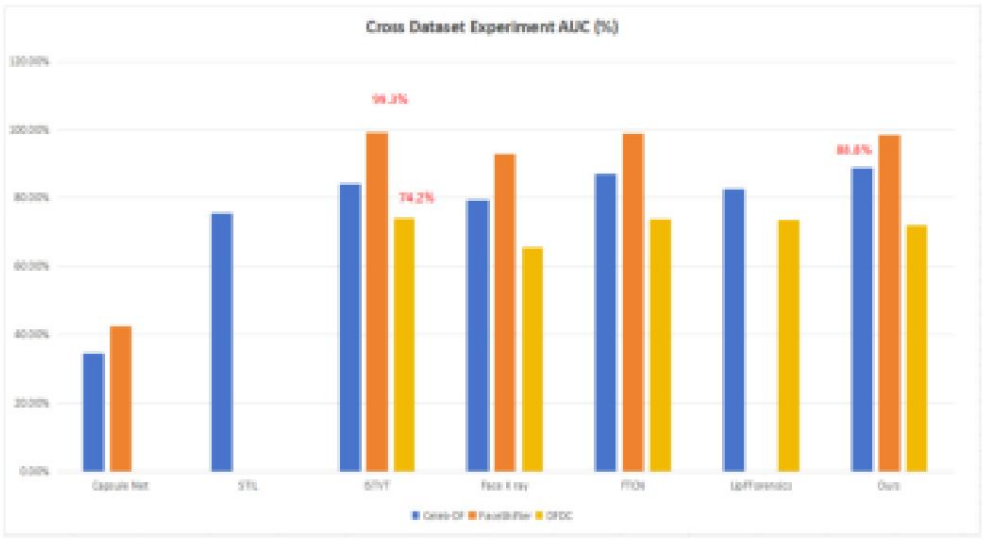


Abbildung 7. Datensatzübergreifende AUC-Ergebnisse basierend auf Capsule Net, STIL, ISTVT, Face X-ray, FTCN, LipForensics und unserem Ansatz. Die beste Leistung für jeden Datensatz ist in der Abbildung gekennzeichnet.

5. Schlussfolgerungen und zukünftige Arbeiten

In diesem Artikel schlagen wir Swin-Fake als Konsistenzlernmethode zur Deepfake-Erkennung vor, die erstmals Swin Transformer mit Konsistenzverlust kombiniert. Wir verwenden Swin Transformer als Merkmalsextraktor und nutzen die durchschnittliche Kosinusdistanz als Maß für den Konsistenzverlust, was die Generalisierungsfähigkeit in einigen Deepfake- Datenbanken verbessert. Wir haben bewiesen, dass ein Transformer-basierter Klassifikator im Bereich der Deepfake-Erkennung ein höheres Niveau erreichen kann als andere frühere CNN-basierte Methoden. Während der Trainings- und Validierungsphasen kamen wir zu dem Schluss, dass Swin Transformer eine bessere Fähigkeit zur Unterscheidung von Deepfakes hat, während große Datenbanken wie FF++ und DFDC bei einigen kleinen Datensätzen wie Celeb DF keine vielversprechende Leistung zeigen. Darüber hinaus könnte die Berechnung der Konsistenz zwischen den Merkmalen der Frames den Verlust erheblich verringern und die Bewertungsmetriken verbessern. Daher glauben wir, dass die Anwendung der Konsistenzlernmethode mit Transformer zur Deepfake-Erkennung ein neuer Trend in der Erkennung gefälschter Medien sein könnte .

Die Erkennung von Deepfake-Videos kann ohne zeitliche Informationen nicht abgeschlossen werden, da in den benachbarten Bildern immer Frame-Jitter auftritt. In der jüngsten Forschung wird die zeitliche Aufmerksamkeit bei der Videoerkennung immer mit der räumlichen Aufmerksamkeit kombiniert. Als eine weitere Technik im Bereich der Videoerkennung kann die zeitliche Aufmerksamkeit als dynamischer zeitlicher Auswahlmechanismus betrachtet werden, der bestimmt, wann auf ein bestimmtes Merkmal eines Videos geachtet werden soll, und wird häufig bei der Videoverarbeitung angewendet. Es ist auch eine effektive Meth

Machine Translated by Google

Metriken in diesem Bereich, wenn wir Informationen zur zeitlichen Konsistenz mit extrahierten räumlichen Informationen kombinieren können. Die gängigste Methode besteht darin, spezielle Module zu entwickeln, um durch Anwenden horizontaler und vertikaler Faltungsschichten die zeitliche Konsistenz oder Inkonsistenz zu untersuchen. Daher werden wir in künftigen Arbeiten die zeitliche Selbstaufmerksamkeit des auf dem Swin Transformer basierenden Backbones weiter untersuchen. Obwohl sich unsere derzeitige Arbeit auf videobasierte Detektoren konzentriert, legen wir dennoch unser Augenmerk auf die räumliche Merkmalsextraktion, und die Merkmalskonsistenz wird für den Intraklassenvergleich angewendet. Die Kombination von zeitlicher Konsistenz mit räumlicher Konsistenz würde wahrscheinlich mehr Informationen beinhalten, was der Klassifizierung von Videos zugute käme. Was die Gewichtungsbalance ÿ betrifft, haben wir in jedem Experiment 0,1 hinzugefügt, um die E Um einen genaueren Wert dieses Hyperparameters zu erhalten, werden wir in zukünftigen Untersuchungen die Gewichtsbalance als lernbaren Parameter festlegen. Darüber hinaus werden wir weitere Visualisierungsmethoden für auf Swin Transformer basierende Deepfake-Detektoren untersuchen, um zu interpretieren, welche räumlichen Merkmale erfasst werden und zum Deepfake-Klassifikato

Beiträge der Autoren: Konzeptualisierung, LYG und XJL; Methodik, LYG und XJL; Software, LYG; Validierung, LYG; formale Analyse, LYG; Untersuchung, LYG; Ressourcen, LYG und XJL; Datenkuratierung, LYG und XJL; Schreiben – Vorbereitung des Originalentwurfs, LYG, XJL und PHJC; Schreiben – Überprüfung und Bearbeitung, LYG, XJL und PHJC; Visualisierung, LYG und XJL; Überwachung, XJL und PHJC; Projektverwaltung, XJL und PHJC. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Forschung erhielt keine externe Finanzierung.

Datenverfügbarkeitserklärung: Die entsprechenden Datensätze wurden von [24] heruntergeladen, die reproduzierten DFDC- Codes stammen von [13].

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Verweise

1. **Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Ware-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Network.** Commun. Acm **2018,** 63, **139–144. [CrossRef]**
2. **Karras, T.; Laine, S.; Aila, T. Eine stilbasierte Generatorarchitektur für generative kontradiktorische Netzwerke. In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung, Long Beach, CA, USA, 15.–20. Juni 2019.**
3. **Ho, J.; Jain, A.; Abbeel, P. Rauschreduzierung bei probabilistischen Diffusionsmodellen. In Proceedings der 34. Internationalen Konferenz übe Neural Information Processing System, Red Hook, NY, USA, 6.–12. Dezember 2020; S. 6840–6851.**
4. **Nguyen, HH; Yamagishi, J.; Echizen, I. Einsatz eines Kapselnetzwerks zur Erkennung gefälschter Bilder und Videos. In Proceedings of the Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16.–20. Juni 2019.**
5. **Das, S.; Seferbekov, S.; Datta, A.; Islam, MS; Amin, MR Zur Lösung des Deepfake-Problems: Eine Analyse zur Verbesserung der Deepfake-Erkennung durch dynamische Gesichtserweiterung. In Proceedings der IEEE/CVF International Conference on Computer Vision, Montreal, BC, Kanada, 11.–17. Oktober 2021.**
6. **Zhao, T.; Xu, X.; Xu, M.; Ding, H.; Xiong, Y.; Xia, W. Lernen der Selbstkonsistenz zur Erkennung von Deepfakes. In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung, Nashville, TN, USA, 20.–25. Juni 2021.**
7. **Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. Ein Bild sagt mehr als 16 × 16 Worte: Transformatoren für die Bilderkennung im großen Maßstab. In Proceedings der International Conference on Learning Representations, virtuell, 26. April–1. Mai 2020.**
8. **Khan, SA; Dai, H. Video Transformer zur Deepfake-Erkennung mit inkrementellem Lernen. In Proceedings der 29. ACM Internationale Konferenz für Multimedia, virtuelle Veranstaltung, 20.–24. Oktober 2021; S. 1821–1828.**
9. **Khormali, A.; Yuan, J. DFDT: Ein End-to-End-DeepFake-Erkennungsframework mit Vision Transformer.** Appl. Sci. **2022,** 12, **2953. [Querverweis]**
10. **Zhao, C.; Wang, C.; Hu, G.; Chen, H.; Liu, C.; Tang, J. ISTVT: Interpretierbarer räumlich-zeitlicher Video-Transformator für Deepfake Erkennung.** IEEE Trans. Inf. Forensics Secur. **2023,** 18, **1335–1348. [CrossRef]**
11. **Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin Transformer: Hierarchischer Vision Transformer mit verschobenen Fenstern. In Proceedings der IEEE/CVF International Conference on Computer Vision, Montreal, BC, Kanada, 11.–17. Oktober 2021.**
12. **Zhong, Z.; Zheng, L.; Kang, G.; Li, S.; Yang, Y. Datenerweiterung durch zufälliges Löschen. In Proceedings of the AAAI Conference on Künstliche Intelligenz, New York, NY, USA, 7.–12. Februar 2020.**
13. **DFDC Selim. Online verfügbar: https://github.com/selimsef/dfdc_deepfake_challenge (abgerufen am 15. Mai 2024).**

Machine Translated by Google

14. **Chollet, F. Xception: Deep Learning mit tiefenweise trennbaren Faltungen. In Proceedings of the Computer Vision and Pattern Anerkennung, Honolulu, HI, USA, 21.–26. Juli 2017.**
15. **He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning für die Bilderkennung. In Proceedings der 2016 IEEE-Konferenz zu Computer Vision und Mustererkennung, Las Vegas, NV, USA, 27.–30. Juni 2016.**
16. **Ni, Y.; Meng, D.; Yu, C.; Quan, CB; Ren, D.; Zhao, Y. CORE: Konsistentes Darstellungslernen zur Erkennung von Gesichtsfälschungen. In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung, New Orleans, LA, USA, 18.–24. Juni 2022.**
17. **Sun, K.; Yao, T.; Chen, S.; Ding, S.; Li, J.; Ji, R. Duales kontrastives Lernen zur allgemeinen Erkennung von Gesichtsfälschungen.** AAAI Conf. Artif. Intell. **2022,** 36, **2316–2324. [CrossRef]**
18. **Zhang, J.; Cheng, K.; Sovernigo, G.; Lin, X. Eine auf heterogenem Feature-Ensemble-Learning basierende Deepfake-Erkennungsmethode. In Proceedings of the ICC 2022—IEEE International Conference on Communications, Seoul, Republik Korea, 16.–20. Mai 2022; S. 2084–2089.**
19. **Rana, MS; Sung, AH DeepfakeStack: Eine auf Deep Ensemble basierende Lerntechnik zur Erkennung von Deepfakes. In Proceedings der 7. IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/6. IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom) 2020 , New York, NY, USA, 1.–3. August 2020; S. 70–75.**
20. **Ilyas, H.; Javed, A.; Malik, KM AVFakeNet: Ein einheitliches End-to-End Dense Swin Transformer Deep Learning-Modell für audiovisuelle Deepfakes-Erkennung.** Appl. Soft Comput. **2023,** 136, **110124. [CrossRef]**
21. **Khalid, F.; Akbar, MH; Gul, S. SWYNT: Swin Y-Net Transformers zur Deepfake-Erkennung. In Proceedings der 2023 International Conference on Robotics and Automation in Industry (ICRAI), Peshawar, Pakistan, 3.–5. März 2023; S. 1–6.**
22. **Guo, J.; Deng, J.; Lattas, A.; Zafeiriou, S. Proben- und Berechnungsumverteilung für eine effiziente Gesichtserkennung. In Proceedings of die Internationale Konferenz zur Lernrepräsentation, virtuelle Veranstaltung, 3.–7. Mai 2021.**
23. **Zhou, TF; Wang, WG; Liang, ZY; Shen, JB Gesichtsforensik in freier Wildbahn. In Proceedings of the Computer Vision and Pattern Anerkennung (CVPR), virtuell, 19.–25. Juni 2021.**
24. **Kaggle. Online verfügbar: https://www.kaggle.com/c/deepfake-detection-challenge/overview (abgerufen am 12. Dezember 2023).**
25. **Li, YZ; Yang, X.; Sun, P.; Qi, HG; Lyu, S. Celeb-DF: Ein umfangreicher, anspruchsvoller Datensatz für die DeepFake-Forensik. In Proceedings der Computer Vision und Mustererkennung, Seattle, WA, USA, 14.–19. Juni 2020.**
26. **Li, L.; Bao, J.; Yang, H.; Chen, D.; Wen, F. FaceShifter: Auf dem Weg zu hochpräzise und okklusionsbewusstem Gesichtsaustausch. In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung, Seattle, WA, USA, 13.–19. Juni 2020.**
27. **Loshchilov, I.; Hutter, F. Entkoppelte Gewichtsabfall-Regularisierung. In Proceedings der International Conference on Learning Representations, New Orleans, LA, USA, 6.–9. Mai 2019.**
28. **Gu, Z.; Chen, Y.; Yao, T.; Ding, S.; Li, J.; Huang, F.; Ma, L. Räumlich-zeitliches Inkonsistenzlernen zur Erkennung von Deepfake-Videos. In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung, virtuell, 19.–25. Juni 2021.**
29. **Li, L.; Bao, J.; Zhang, T.; Yang, H.; Chen, D.; Wen, F.; Guo, B. Gesichtsröntgen für allgemeinere Gesichtsfälschungserkennung. In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung, Seattle, WA, USA, 13.–19. Juni 2020.**
30. **Zheng, Y.; Bao, J.; Chen, D.; Zeng, M.; Wen, F. Untersuchung der zeitlichen Kohärenz zur allgemeineren Erkennung von Gesichtsfälschungen in Videos. In Proceedings der IEEE/CVF International Conference on Computer Vision, Montreal, QC, Kanada, 10.–17. Oktober 2021; S. 15044–15054.**
31. **Haliassos, A.; Vougioukas, K.; Petridis, S.; Pantic, M. Lippen lügen nicht: Ein verallgemeinerbarer und robuster Ansatz zur Erkennung von Gesichtsfälschungen. In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung, virtuell, 19.–25. Juni 2021; S. 5039–5049.**

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Ein Prosumer-Wasserkraftwerksnetz als nachhaltiges verteiltes Energiedepot

Michaÿ Morawski und Przemysÿaw Ignaciuk *

Fakultät für Technische Physik, Informationstechnologie und Angewandte Mathematik, Technische Universität Lodz, 93-005 Lodz, Polen; michal.morawski@p.lodz.pl *
Korrespondenz: przemyslaw.ignaciuk@p.lodz.pl

Zusammenfassung: Der Mangel an effizienten, kostengünstigen Speicherdepots behindert die großflächige Einführung von von Natur aus volatilen, erneuerbaren Energiequellen (RSEs). In diesem Artikel beschreiben wir, wie man Wasserkraftwerke im Besitz von Prosumern mit einigen bis mehreren kW als verteilte, kurzfristige Energiespeicherlösung nutzen kann, die mit geringen Investitionen und geringen Betriebskosten eingesetzt werden kann. Die vorgeschlagene Lösung ist ein System aus miteinander verbundenen Wasserkraftwerken mit einem aktiven Wasserfluss-Kontrollalgorithmus, der die Lastvariabilität des Netzes reduziert und den Prosumern Vorteile bringt. Den durchgeführten Tests zufolge steigt der Prosumer-Umsatz je nach Wetterbedingungen im Vergleich zum Fall des freien Flusses um einige Prozent auf über 30 Prozent. Der Puffereffekt des verteilten Energiepuffers wiederum gleicht die Schwankungen aus, die durch andere RSEs verursacht werden, z. B. solche auf Photovoltaik oder Windkraft. Dies kommt nicht nur den beteiligten Parteien zugute, sondern erleichtert auch di

Schlüsselwörter: Energiedepot; Wasserkraftwerke; verteilte Systeme

1. Einleitung

Elektrische Energie gilt als Lebensader der heutigen Gesellschaft. Der Stromverbrauch in Haushalten, Industrie, Verkehr sowie bei der Kühlung und Heizung von Gebäuden führt täglich zu einem erhöhten Bedarf. Gleichzeitig ist die Energieerzeugung sowohl aus ökologischer als auch aus wirtschaftlicher Sicht weiterhin teuer. Globale Umweltverschmutzung, die steigenden Kosten für fossile Brennstoffe und politischer Druck sind nur einige Faktoren, die die Energieerzeugung aus erneuerbaren Quellen (RSEs) fördern, z. B. Sonne, Wind, Wellen, Geothermie und Wasserbewegung [1]. Leider hängt die Menge der aus RSEs gewonnenen Energie stark von den zeitlichen Wetterbedingungen ab. RSEs liefern kostengünstige Energie entweder zu zufälligen Zeiten (Wind) oder um die Mittagszeit (Solar, meist Photovoltaik (PV)). Im Gegensatz dazu bleibt der Bedarf bis zum Abend auf einem ähnlichen Niveau, wie in Abbildung 1 dargestellt. Daher müssen teure, auf fossilen Brennstoffen basierende Kraftwerke abends in Betrieb sein, ihre Aktivität sollte jedoch zugunsten von RSEs eingeschränkt werden, wenn Sonne oder Wind verfügbar sind. Darüber hinaus kann die Energieerzeugung aus herkömmlichen Wärmekraftwerken aufgrund technologischer Faktoren nicht zu stark eingeschränkt werden. Aufgrund der Marktdynamik werden die Energiepreise zu bestimmten Zeiten niedrig (oder sogar negativ) angesetzt, um dann einige Stunden später eine höhere Größenordnung zu erreichen, was der in Abbildung 2 dargestellten „Duck Curve“ ähnelt, die die ersten Tage des Juli 2023 auf dem polnischen Energiemarkt zeigt. Im Interesse aller Beteiligten sollte diese Diskrepanz verringert werden.

Eine selbst auferlegte Lösung für das Ungleichgewicht zwischen Erzeugern und Verbrauchern besteht darin, überschüssige Energie zu einem späteren Zeitpunkt zu speichern, wenn sie billig ist, um sie dann zu einem späteren Zeitpunkt zu nutzen, wenn kostspieligere Quellen genutzt werden. Leider erlaubt die heutige Technologie es nicht, dies in größerem Maßstab umzusetzen [2]. Beispielsweise können Pumpspeicherkraftwerke in Polen höchstens 5 % des gesamten Strombedarfs decken, und die Batteriedepots liefern höchstens 1 %. Beide dieser traditionellen Methoden sowie zukünftige wasserstoffbasierte Systeme sind teuer in der Installation und Wartung. Die geringe Lebensdauer elektromechan

Daher **ist die Suche nach einer neuen Option zur Energiespeicherung mit geringen Investitions- und Betriebskosten** wünschenswert.

check for updates

Zitat: Morawski, M.; Ignaciuk, P. Ein Prosumer-Wasserkraftwerksnetzwerk als nachhaltiges dezentrales Energiedepot. Electronics **2024,** 13, **3043. https://doi.org/10.3390/ Elektronik13153043**

Wissenschaftliche Herausgeber: Temitayo Olowu und Raghav Khanna

Empfangen: 24. Juni 2024
Überarbeitet: 22. Juli 2024
Akzeptiert: 30. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

Machine Translated by Google

In jüngster Zeit ist ein grundlegender Wandel auf dem Strommarkt zu beobachten. Traditionelle Großkraftwerke werden durch Kleinanlagen ergänzt, die von Bürgern auf ihrem eigenen Grundstück installiert werden. Diese Bürger sind zu bestimmten Tageszeiten Stromverbraucher und zu anderen Zeiten Erzeuger und werden daher als Prosumer bezeichnet. Einerseits fördert die Popularität von Prosumer-Anlagen die Einführung von RSEs, andererseits bedroht sie die Stabilität des Energieversorgungssystems, indem sie traditionelle Kraftwerke zu extremen Lastwechseln zwin

ff

Diese Studie trägt auf folgende Weise zu den oben genannten Herausforderungen bei:

- **Wir schlagen den Einsatz von Prosumer-Wasserkraftwerken mit einer Spitzenleistung im Bereich von einig kW, um ein dezentrales Energiedepot zu bilden;**
- **Wir entwickeln ein aktives Steuerungsschema, das sowohl eine Erhöhung der Erträge der Anlagenbesitzer als auch eine Reduzierung der Lastschwankungen bei bestehenden thermischen Anlagen** ermöglicht **(„Köpfen der Ente“ [3]) und so eine Win-Win-Lösung erreicht;**
- **Im Gegensatz zum typischen Ansatz, der sich auf einen einzelnen Hybridgenerator konzentriert, wie z. B. [4],** zeigen wir, wie ein Netzwerk von Anlagen für einen sicheren und effizienten verteilten **Systembetrieb koordiniert werden kann .**

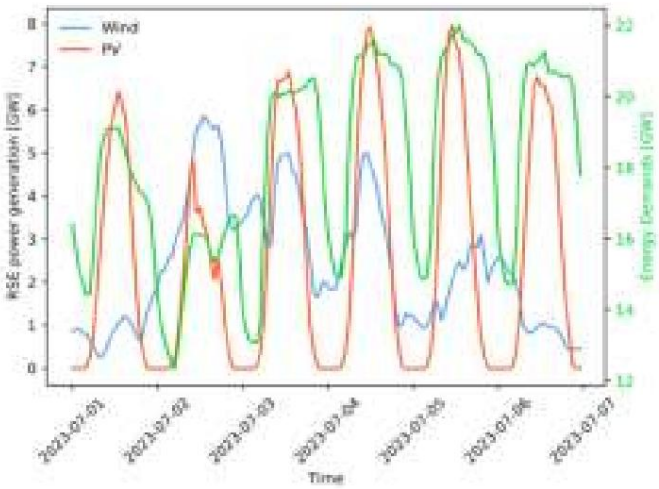


Abbildung 1. Energieerzeugung aus den wichtigsten RSEs und prognostizierter Energiebedarf in Polen für den Zeitraum vom 1. bis 6. Juli 2023 (der 1. Juli war ein Samstag [5]).

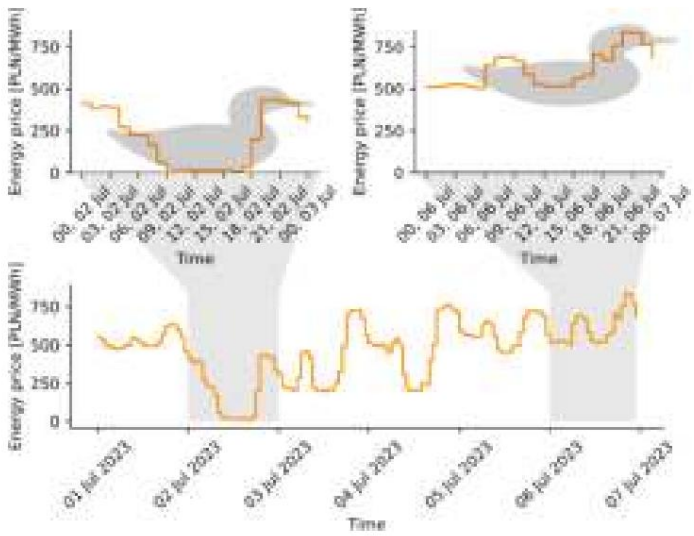


Abbildung 2. Schwankung der Energiepreise für den nächsten Tag [PLN/MWh] auf dem polnischen Markt vom 1. bis 6. Juli 2023 gemäß der „Duck Curve“ [5].

Machine Translated by Google

Der Rest des Dokuments ist wie folgt gegliedert. In Abschnitt 2 werden die traditionellen Methoden der Wasserkraftwerkssteuerung erörtert. In Abschnitt 3 gehen wir näher auf die Unterschiede zwischen konventionellen und Prosumer-Kraftwerken ein. In Abschnitt 4 wird ein zeitdiskretes dynamisches Modell des Wasserflusses in einem System mit mehreren Wasserkraftwerken erstellt und ein Verfahren zur Optimierung seines Betriebs beschrieben. In Abschnitt 5 bewerten wir die Leistung des Algorithmus für ein breites Spektrum von Testdaten. Anschließend diskutieren wir in Abschnitt 6 die Überlegungen zur Implementier Abschließend werden in Abschnitt 7 Schlussfolgerungen gezogen.

2. Steuerung von

Wasserkraftwerken In der Literatur werden Wasserkraftwerke (WKW) als normale (>100 MW), mittlere (<100 MW), kleine (<15 MW), Mini- (<1 MW), Mikro- (<100 kW) oder Pico- (<5 kW) Wasserkraftwerke klassifiziert, wobei die tatsächlichen Werte je nach Region variieren [6,7]. Der Großteil der Forschung konzentriert sich auf kleine, mittlere und normale Wasserkraftwerke, die eine ausgeprägte Rolle bei der Stromversorgung spielen. In diesem Artikel werden aufgrund ihrer Bedeutung für Prosumer Mikro- und Pico-Wasserk Sie umfassen zahlreiche ehemals verlassene Mühlen mit kleinen künstlichen Teichen, Sümpfen und anderen natürlichen Formationen, die gemeinhin als Laufwasser-Wasserkraftwerke bekannt sind [8,9], deren Einsatz **somit nur unbedeutende Kapitalausgaben erfordert . Derzeit werden beispielsweise in Polen weniger als 5 % der möglichen Anlagen zur Energieerzeugung genutzt [10]. Darüber hinaus** zeigen zahlreiche Veröffentlichungen von Online-Diensten, dass Pico-Wasserkraftwerke problemlos ohne **Spezialausrüstung eingesetzt werden können [11]. Daher ist der Zielbereich der in dieser Arbeit durchgeführten Studie wichtig. Natürlich ist ein einzelnes Prosumer-Depot kein bahnbrechender Schritt. Ein umsatzorientierter, optimal geplanter Satz von Prosumer-Wasserkraftwerken kann jedoch ein effizientes und störungsresistentes Energiespeichersystem mit geringen Betriebskosten darstellen. Tatsächlich führten ähnliche** Umstände zur Popularität von PV-Anlagen, die zu einem gemeinsamen **Bestandteil der heutigen Landschaft geworden sind .**

Es gibt jedoch einen grundlegenden Unterschied zwischen PV- und HP-Anlagen. Die PV-Anlagen sind voneinander unabhängig, während der ordnungsgemäße Betrieb einer HP davon abhängt, dass alle Stationen stromaufwärts in einer Kaskade verbunden sind. In der Vergangenheit wurde die Planung des Betriebs einer Kaskade von HPs als Fallstudie [12] untersucht **und nicht als vernetztes System, das einer Steuerungsmaßnahme unterworfen ist, d. h. jede Anlage wurde manuell angepasst, ohne dass ein gemeinsamer Rahmen durchgesetzt wurde. Obwohl auf diese Weise ein geeigneter Zeitplan definiert werden kann , berücksichtigt er möglicherweise nicht die optimalen Betriebsbedingungen oder stellt ein Risiko in Bezug auf den Dürre-/ Hochwasserschutz dar. Andere Ansätze berücksichtigen vereinfachte Modelle, die die Kaskade auf einen Einzelanlagenfall** reduzieren **[4], bei dem die Dämme nah beieinander liegen [8,13] oder an einem Fluss, aber nicht an seinen Nebenflüssen** liegen **[9,14]. Eine Lösung, die der hier diskutierten konzeptionell am nächsten kommt,** wurde in **[13] vorgeschlagen , obwohl sie auf große, konzentrierte Anlagen und nicht auf die verteilten HPs unter Prosumer-Management abzielt.**

Technisch gesehen ist die Steuerung der Energieproduktion in einem Wasserkraftwerk kein komplizierter Prozess. Durch Anheben der Wehre wird der Wasserfluss eingeschränkt, und durch Absenken der Wehre wird die Energieproduktion erhöht. Die richtige Reihenfolge des Anhebens und Absenkens der Wehre festzulegen, ist jedoch alles andere als offensichtlich. Normalerweise wird die Planungsstrategie in langfristig (1–5 Jahre), mittelfristig (3–18 Monate) und kurzfristig (2–14 Tage) eingeteilt [15]. Im Fall von Prosumer-Wasserkraftwerken sind solche Strategien jedoch aufgrund des geringen Volumens der Teiche nur von begrenztem Nutzen. Im betrachteten Fall sollte eine sinnvolle Planungsstrategie den Bereich von wenigen Stunden oder höchstens einem Tag abdecken. Darüber hinaus sollte sie sich darauf konzentrieren, das Ungleichgewicht zwischen Erzeugung und Verbrauch abends und morgens zu verringern (Abbildung 2) oder, gleichwertig, die schwankenden Marktpreise auszunutzen, die sich aus der Diskrepanz zwischen prognostiziertem und tatsächlichem Bedarf ergeben **[16] (Abbildung 1).**

Die optimale Planung des Problems der Wasserfreisetzung durch Hochdruckwasserkraftwerke wurde lange Zeit untersucht ([15,17] und die darin enthaltenen Referenzen). In diesen Studien wurden Hochdruckwasserkraftwerke jedoch nur im Zusammenhang mit Hochwasser-/Dürreschutz oder Stromerzeugung untersucht. Hier wird eine dritte Rolle von Hochdruckwasserkraftwerken untersucht, nämlich das Potenzial, ein Energiedepot bereitzustellen [2]. Anstatt sich nur auf die Menge des erzeugten Stroms zu konzentrieren, können Hochdruckwasserkraftwerke den Flussfluss einschränken, wenn die Energie billig

Machine Translated by Google

Energiepuffer. Die Gesamtmenge der erzeugten Energie bleibt gleich, aber der Umsatz des Wasserkraftwerkbesitzers steigt, da er die kurzfristigen Änderungen im Preisprofil berücksichtigt. Gleichzeitig verringert sich die Fluktuation der Energieerzeugung aus herkömmlichen Wärmekraftwerken . Eine ähnliche Idee wurde in [18] untersucht, sie ist jedoch für große Anlagen bestimmt und nicht für die in dieser Arbeit betrachteten Prosumer-Anlagen.

3. Prosumer-Wasserkraftwerke

Nicht alle RSEs sind empfindlich gegenüber dem Einfluss von Wetterschwankungen, wie z. B. geothermische, gezeitenbasierte oder große, teure Wasserkraftwerke [19]. Hier liegt der Schwerpunkt auf kleinen, flexiblen Anlagen, d. h. Mikro- [6] und Pico **- Wasserkraftwerken [20]. Sie werden typischerweise als Flusslaufanlagen [8,21] eingesetzt, d. h. ohne große Dämme und Reservoirs, und verursachen daher keine erheblichen Kosten oder Eingriffe in die Umwelt [21]. Im Gegensatz zu größeren Wasserkraftwerken, die nicht von lokalen atmosphärischen Bedingungen beeinflusst werden, da die Wassermenge über den Vorhersagen im Vergleich zum Reservoirvolumen vernachlässigbar ist, macht der begrenzte Speicherraum von Mikro- und Pico-Reservoirs sie anfällig für Wetterschwankungen. Beispielsweise kann ein lokaler Sturm den Teich in der Nähe eines Mikro-Wasserkraftwerks sofort füllen, während ein nahe gelegenes gleichzeitig sonnige und trockene Bedingungen aufweist. Heutzutage sind Niederschlagsvorhersagen nicht präzise genug, um mit solchen Situationen umzugehen. Auch die Wasserabflussmodelle sind nicht genau [22]. Daher muss bei instabilem Wetter der Kraftwerksplan häufig neu berechnet werden, um das System an die neuesten Umstände oder die Preisvariabilität in einem Echtzeit-Ausgleichsmarkt anzupassen [16]. Im Gegensatz zu den üblicherweise verwendeten numerischen Optimierungsansätzen dauert die Ausführung des in diesem Manuskript vorgestellten Algorithmus selbst auf einem Low-End-Gerät nur den Bruchteil einer Sekunde .**

Eine Reihe von Wasserkraftwerken nutzt das Wasser aus vorgelagerten Anlagen und Nebenflüssen. Eine unsachgemäße Steuerung kann daher zu Überschwemmungen führen. Daher ist es zwingend erforderlich, den Betrieb der Wasserkraftwerke zu synchronisieren und dabei verschiedene

Faktoren explizit zu berücksichtigen: (1) Geografischer Faktor – die Entfernung zwischen den Anlagen und die daraus resultierende Verz

ter-Reflow;

(2) Wetterbezogene Faktoren – die vorhergesagte Niederschlagsmenge und der Unsicherheitsgrad;
(3) Anwendungsorientierter Faktor – Nutzung der HD-Kaskade als dezentrales Energiedepot

kombiniert mit der gemeinsamen Funktion der Stromerzeugung.

Normalerweise wird der Steuerungsplan mithilfe numerischer Optimierungsverfahren erstellt, wobei zur Reduzierung des Rechenaufwands Techniken der künstlichen Intelligenz einbezogen werden [15,17,21–29]. Diese Verfahren weisen jedoch im **betrachteten Anwendungsbereich selbst in relativ einfachen Fällen Robustheitsprobleme auf [29]. Darüber hinaus widerspricht die lange Zeit, die zur Erzielung einer Lösung benötigt wird (unter Verwendung angemessener Rechenressourcen), der Notwendigkeit der Plananpassung, d. h. der Einführung von Änderungen am Steuerungsplan, wenn das Wetter instabil ist oder eine Störung auftritt.**

Um die Faktoren (1)–(3) zu erfassen, werden Informationen über unterschiedliche Verzögerungen des Wasserflusses auf den Verbindungen zwischen den Stauseen explizit in das in dieser Arbeit erstellte dynamische Modell integriert. Unter Verwendung der dynamischen Darstellung des Systems wird ein Optimierungsproblem formuliert. Es kann analytisch gelöst werden, wenn keine Sättigung vorliegt [30]. Im Gegensatz dazu sind hier die natürlichen Beschränkungen in die algorithmische Lösung eingebettet. Dieser Vorschlag bringt nicht nur den Prosumenten Vorteile, indem er ihren wirtschaftlichen Gewinn erhöht, sondern auch den Stromnetzbetreibern, indem er die Lastschwankungen für

4. Systemmodell

Normalerweise konzentriert sich die HP-Modellierung auf die Optimierung der Arbeit von Generatoren und ignoriert das Versorgungssystem [23]. Dabei wird der Generator als Blackbox betrachtet und der Fokus auf die angeschlossenen HPs als Wasserversorgungssystem gelegt. Ein wichtiger Punkt, der in einem Wasserrückflusssystem berücksichtigt werden muss, ist die nicht zu vernachlässigende Zeit zwischen der Ausgabe der Steueraktion in einem Reservoir, bevor sie den Wasserstand in einem nachgelagerten beeinflusst. Im Gegensatz zu den früheren Modellen von Speichernetzwerken, z. B. [31], werden daher **in dem hier vertretenen Ansatz die Steuerprinzipien von zeitverzögerten Speichersystemen [32,33] ange**

Machine Translated by Google

Die in **[32,33] vorgeschlagenen Modelle gehen von einer kontinuierlichen Regelung aus, die in einem Wasserregelungssystem aufgrund der Besonderheiten der mechanischen Komponenten, die die Stauwehre steuern, schwierig zu realisieren ist. Das Modell in dieser Arbeit deckt explizit die Auswirkungen einer endlichen Abtastzeit ab und wird direkt im diskreten Zeitbereich konstr**

4.1. Einzelanlagendepot

Betrachten wir das Modell einer einzelnen Wasserkraftanlage als Bestandteil des in Abbildung 3 dargestellten Systems. Die Dynamik des Wasserhaushalts in der Anlage kann durch die rekursive Beziehung beschrieben werden

$$sj(k + 1) = sj(k) ÿ fj(k) + ÿ i ÿ Anlagen \quad fi(k ÿ Tij) + rj(k), \qquad (1)$$

stromaufwärts

Wo

- k ÿ **[1, m] ist ein Zeitpunkt und** m **ist ein Planungshorizont;**
- sj (k) **ist die Wassermenge, die im Reservoir in der Nähe von Anlage** j gespeichert ist **(in Abbildung 3 ist Anlage** j **= 3 vergrößert dargestellt);** sj (k) **umfasst Wasserbauwerke sowie nahegelegene Sümpfe und ähnliche Gebiete;**
- fj (k) **ist die Wassermenge, die zum Antrieb der am Damm** j installierten Stromgeneratoren **zwischen den Zeitpunkten** k **und** k **+ 1**
- **verwendet wurde;** rj (k) **ist die Versorgung aus externen hydrologischen Quellen wie Regen (und dessen Abfluss), schmelzendem Schnee, unkontrollierten Zuflüssen und Verdunstung. Die** rj (k) -Werte können **aus der Wettervorhersage und hydrologischen Modellen innerhalb des Planungshorizonts von** m **Perioden gewonnen werden .** rj (k) **kann positiv (Häufigkeit von Niederschlägen) oder negativ (Häufigkeit von Verdunstung) sein.** rj (k) **wird im Modell als bekannt vorausgesetzt. Wenn der wahre Wert von** rj (k) **von der Prognose abweicht, wird der Algorithmus erneut ausgeführt, was keine komplexen Berechnungen Die Zuflüsse versorgen das Reservoir mit dem Wasser, das zuvor von den flussaufwärts gelegenen Anlagen verbraucht wurde. Das Wasser aus der flussaufwärts gelegenen Anlage** i **erreicht die Anlage** j **mit** einer Verzögerung von Tij **> 0. Die Periodenlänge** ÿk, **d. h. die Zeit zwischen den Zeitpunkten** k **und** k **+ 1, kann beliebig gewählt werden, aber je nach Geschwindigkeit der Preisänderungen ist es sinnvoll, je nach örtlichen Vorschriften 1 h bis 15 min zu wählen. Ebenso kann der Planungshorizont** m **ein 24-Stunden-Fenster bekannter Energiepreise (den Markt des nächsten Tages) abdecken . Der Anfangsdurchfluss** fj (k ÿ **0) und der Anfangswasserstand** sj **(0) werden als bekannt vorausgesetzt. Die Endbedingung** sj (m) **kann beliebig gewählt werden.**

Der Ertrag aus der Anlage, der in einem Zeitraum erzielt wird, kann wie folgt berechnet werden:

$$Jj(k) = ÿj\ pj(k)fj(k)ÿk, \qquad (2)$$

Wo

- pj (k) **ist der Energiepreis zum Zeitpunkt k. Die Tatsache, dass alle WPs an dasselbe System angeschlossen sind, führt nicht dazu, dass jedes Kraftwerk j dasselbe Preisprofil hat . Die Eigentümer können Verträge mit verschiedenen Unternehmen haben, unterschiedliche Tarife verwenden oder selbst unterschiedliche Verbrauchsmuster aufweisen ;**
- ÿj **ist der Wirkungsgrad von Stromgeneratoren unter Berücksichtigung des Einflusses der Staudammhöhe. Bei den Prosumer-Generatoren im Tiefland entspricht ein Durchfluss von 1 m3/s einer Stromerzeugung von 5–7 kW.**

Machine Translated by Google

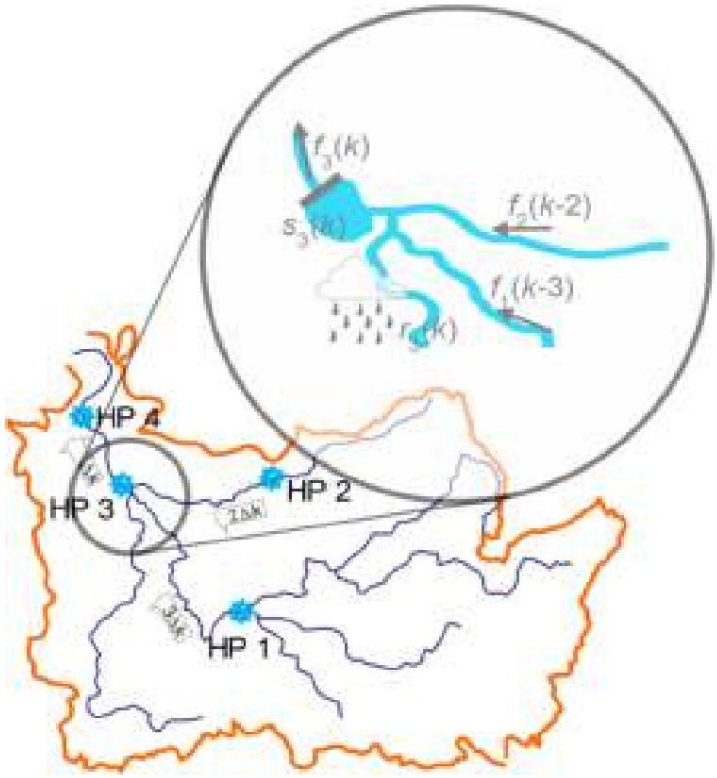


Abbildung 3. Modell verbundener Wasserkraftwerke in einem Flussgebiet. Anlage 3 (vergrößert) wird versorgt durch die Zuflüsse aus den Reservoiren 1 und 2 mit Verzögerungen, die durch dicke Pfeile angezeigt werden (ÿk – Probenahmeperiode) und exogene Quellen, z. B. Regen oder Zuflüsse im Süden. Die Wassermenge, die an das Werk geliefert wird aus diesen Quellen innerhalb von ÿk zum Zeitpunkt k ist f 1 (k ÿ 3), f 2 (k ÿ 2) bzw. r3 (k). Der Abfluss von den Staudammgeneratoren in Anlage 3 ist f 3 (k). Die Differenz zwischen Eingangs- und Ausgangsfluss wird im Teich, dessen aktuelles Wasservolumen mit s3 (k) bezeichnet wird.

4.2. Mehrwerksdepot

Ähnlich wie bei Photovoltaikanlagen erhöht sich mit der Anzahl der HPs auch deren Einfluss auf das Netz. Eine Ansammlung von Pico- und Mikro-HPs kann ein großes, verteiltes Energiedepot bilden.

Im Falle von n **Kraftwerken im gleichen Flussgebiet werden die Modellvariablen gruppiert in Vektoren:**

$$s(k) = \begin{bmatrix} s1(k) \\ s2(k) \\ \vdots \\ sn(k) \end{bmatrix}, f(k) = \begin{bmatrix} f1(k) \\ f2(k) \\ \vdots \\ fn(k) \end{bmatrix}, r(k) = \begin{bmatrix} r1(k) \\ r2(k) \\ \vdots \\ rn(k) \end{bmatrix}, \quad (3)$$

wobei s(k) **der Vektor des Stauseewasserspiegels,** f(k) **der Vektor der Wassermengen in den Flüssen zwischen den Stauseen und** r(k) **der Vektor der Wassermengen aus exogenen Quelle**

Abgeleitet aus (1) wird die Dynamik vernetzter Systeme beschrieben durch

$$s(k+1) = s(k) + \sum_{t=0}^{T} \ddot{y}t\, f(k-t) + r(k), \quad (4)$$

wobei T **die maximale Verzögerung ist und die Matrizen** ÿt **die Informationen über die grafische Topologie und Flussverzögerungen.** ÿt = ÿij $_{n\times n}$ **wird** ÿij = 1, **wenn der Durchfluss vom Reservoir** j **erreicht Reservoir** i **mit Verzögerung t, andernfalls ist es** 0. ÿ0 **beschreibt den Abfluss aus das System. Im Gegensatz zu [30] ist hier der Abstand zwischen den Anlagen nicht vernachlässigbar. Für die Beispiel aus Abbildung 3, die längste Verzögerung** T = 3 **(der Durchfluss zwischen den Reservoirs 1 und 3) und**

Machine Translated by Google

$$\ddot{y}0 = \begin{bmatrix} \ddot{y}1\,0 & 0\,0\,0 \\ \ddot{y}\ 0\ \ddot{y}1\ 0 & \ddot{y} \\ 0\ 0\ \ddot{y}1\ 0 & \ddot{y}\ ,\ \ddot{y}1 = \ddot{y} \\ \ddot{y}\ 0\ 0 & 0\ \ddot{y}1 \end{bmatrix} \begin{bmatrix} 0\,0\,0\,0 \\ \ddot{y} \\ \ddot{y}0\,0\,0\,0 \\ 0\,0\,0\,0 \\ \ddot{y}0\,0\,1\,0 \end{bmatrix} \begin{bmatrix} \ddot{y} \\ \ddot{y}\ ,\ \ddot{y}2 = \ddot{y} \\ \ddot{y} \end{bmatrix} \begin{bmatrix} 0\,0\,0\,0 \\ \ddot{y} \\ \ddot{y}0\,0\,0\,0 \\ 0\,1\,0\,0 \\ \ddot{y}0\,0\,0\,0 \end{bmatrix} \begin{bmatrix} \ddot{y} \\ \ddot{y}\ ,\ \ddot{y}3 = \ddot{y} \\ \ddot{y} \end{bmatrix} \begin{bmatrix} 0\,0\,0\,0 \\ \ddot{y} \\ \ddot{y}0\,0\,0\,0 \\ 1\,0\,0\,0 \\ \ddot{y}0\,0\,0\,0 \end{bmatrix} \begin{bmatrix} \ddot{y} \\ \ddot{y} \\ \ddot{y} \end{bmatrix} . \quad (5)$$

Für das Mehrpflanzensystem wird der Qualitätsindex (2)

$$J = \ddot{y}_{k=1}^{N} Jj = \ddot{y}_{k=1}^{N} \ddot{y}_{j=1}^{M} Jj(k) = \ddot{y}_{k=1}^{N} \ddot{y}_{j=1}^{M} \ddot{y}\,\ddot{y}j\,pj(k)fj(k)\ddot{y}k. \quad (6)$$

Die Interessengruppen haben oft gegensätzliche Ziele. Prosumenten versuchen, J zu maximieren, während Netzbetreiber J maximieren wollen. Ein bequemer Weg, ihre Jj-Interessen in Einklang zu bringen , **ist ein Preissystem. Der im letzten Teil dieses Dokuments vorgestellte Algorithmus erfüllt die Ziele sowohl der Anlagen- als auch der Netzbetreiber.**

4.3. Steuerungsalgorithmus

Die größte Herausforderung bei der Lösung des Problems (6) liegt in der Komplexität der Systemdynamik, die sich aus der Anwesenheit einer nicht zu vernachlässigenden Verzögerung in (4) ergibt. Eine weitere Schwierigkeit ergibt sich aus den physikalischen Beschränkungen, die den Variablen f(k) **und** s(k) **auferlegt werden.**

$$\begin{aligned} &\ddot{y}_{j,k}\ \ddot{y}_{f\,j}^{\text{Mindest}}\ \ddot{y}\ fj(k)\ \ddot{y}\ f_j^{\max}, \\ &\min_{j,k}\ \ddot{y}_{J}^{s}\ \ddot{y}\ sj(k)\ \ddot{y}\ s_j^{\max}, \\ &\ddot{y}_{j}\ sj(m) = Sj, \end{aligned} \quad (7)$$

wobei Sj **ein willkürlich gewählter Endwert ist.**

Wie oben erwähnt, sind die Standardlöser in der hier betrachteten Problemklasse nicht robust **[15,17,21–29]. Um die rechnerischen Aspekte zu berücksichtigen, wird daher ein dreistufiges Min-Max- Optimierungsverfahren angewendet, das in Abbildung 4** skizziert **ist. Jeder Schritt des Algorithmus wird unten erläutert. Die Kommunikation zwischen den HPs entspricht dem bekannten Industriestandard Message Queuing Telemetry Transport (MQTT) [34], der sowohl Sicherheit als auch Synchronisierung zwischen den Anlagensteuerungen bietet.**

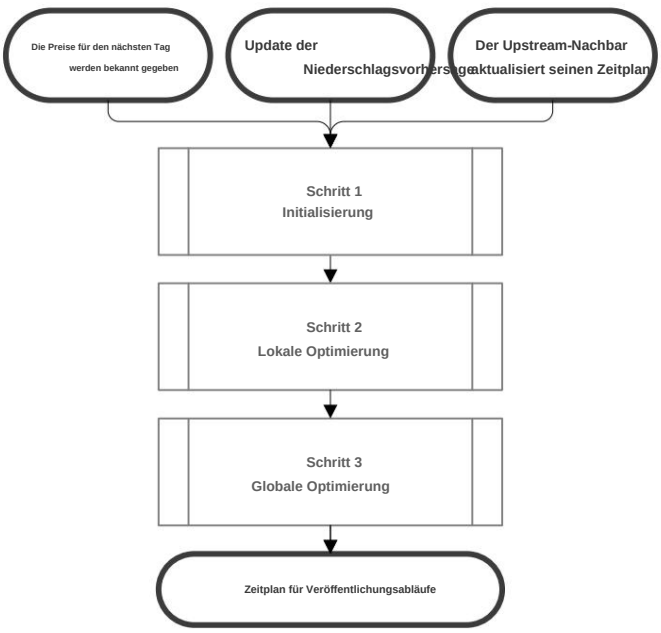


Abbildung 4. Flussdiagramm des Algorithmus.

Machine Translated by Google

Schritt 1, Initialisierung:

Eingabe: Preisprofil.

Ausgabe: vorläufiger Wasserflussplan.

Ziel: Bestimmung der Akkumulations- und Ertragsintervalle innerhalb des betrachteten Zeitrahmens, wie in Abbildung 5 dargestellt. Wenn die Energiepreise über dem Durchschnitt liegen, ist es Zeit, Geld zu verdienen (Ertragsintervall). Andernfalls sollte man das Wasser speichern (Akkumulationsintervall). Im ersten Fall ist f(k) **= f max und im zweiten Fall** ist f(k) **= f für jedes** k **im jeweiligen Intervall. Als Ergebnis erhält man** für Anlage **j** hj **Perioden mit der Länge** mjp . **Da die HPs aufgrund unterschiedlicher Tarife oder Nachfragen unabhängig voneinander verwaltet werden, kann das Segmentierungsmuster für jede HP unterschiedlich sein. In beiden Fällen erfordert dieser Schritt keine komplexen Berechnungen. Die in Schritt 1 erhaltene Näherung kann dazu führen, dass die in (7) definierten Grenzen von** s(k) überschritten werden .

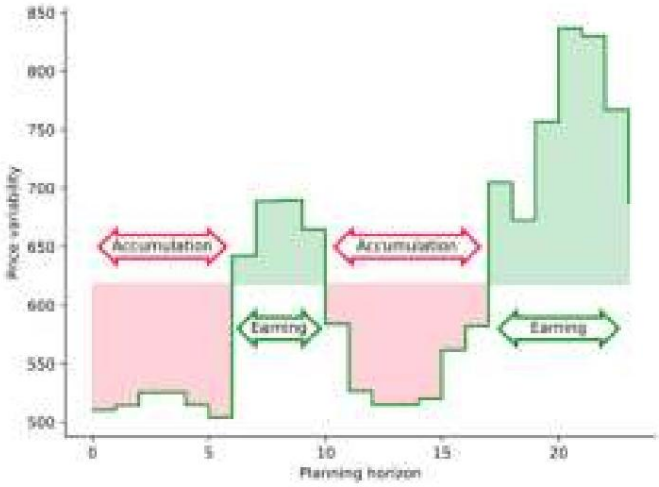


Abbildung 5. Darstellung von Schritt 1: Identifizierung von Akkumulations- und Ertragsintervallen für das Preisprofil vom 6. Juli. Markierte Zeiträume wirken sich auf die Version von Schritt 2 des vorgestellten Algorithmus aus.

Schritt 2, lokale Optimierung:

Eingabe: Preisprofil und vorläufiger Wasserdurchflussplan, berechnet in Schritt 1, sowie der Zufluss (d. h. vorhergesagter Niederschlag, Zufluss der Nebenflüsse, der Plan der vorgelagerten Anlagen) zum System.

Ausgabe: lokal optimierter Flussplan. Das Ergebnis überschreitet nicht die in (7) definierten Grenzen.

Ziel: Suche nach dem optimalen Fluss innerhalb jedes Zeitraums, der durch Ausführung von Schritt 1 ermittelt wurde.

Da es zwei Arten von Zeiträumen gibt, ist es sinnvoll, Schritt 2 entsprechend aufzuteilen.

Schritt 2a (Akkumulationsintervall): Wenn für jedes k gilt, dass sj (k) **>** sj **max, bedeutet dies, dass der Wasserfluss zu stark eingeschränkt ist. Der Wasserzufluss überschwemmt daher den Teich. Daher ist es wünschenswert, den Zeitpunkt k* mit dem höchsten Preis innerhalb des Intervalls** nj zu finden **und den Fluss um** sj **(k*)** ÿ sj **max zu erhöhen, jedoch nicht über** fj **max. Dieser Schritt wird wiederholt, bis die Bedingung** sj (k) ÿ sj **für alle** k **im untersuchten Intervall erfüllt ist.**

Schritt 2b (Ertragsintervall): Wenn für jedes k gilt, dass sj (k) **<** sj **min ist, bedeutet dies, dass der Fluss übermäßig ist. Man sollte also einen Zeitpunkt k* mit dem niedrigsten Preis in diesem Intervall finden** und den Fluss **min so einschränken, dass** sj **für alle** k **im untersuchten Intervall** erfüllt **ist .** ÿ sj **(k*), aber nicht unter** fj **min. Dieser Schritt wird wiederholt, bis die Bedingung** sj (k) ÿ sj

Schritt 2 wird für jeden HP ausgeführt, beginnend mit den vorgelagerten. Die maximale Anzahl m2 von Iterationen beträgt ÿ für jeden HP j, was niedrig genug ist, um in Echtzeit ausgeführt zu werden, selbst ein groß angelegtes System. Schritt 2 ist in Abbildung 6 dargestellt. Nach den Berechnungen wird innerhalb jedes Intervalls der maximal mögliche Gewinn erzielt.

Machine Translated by Google

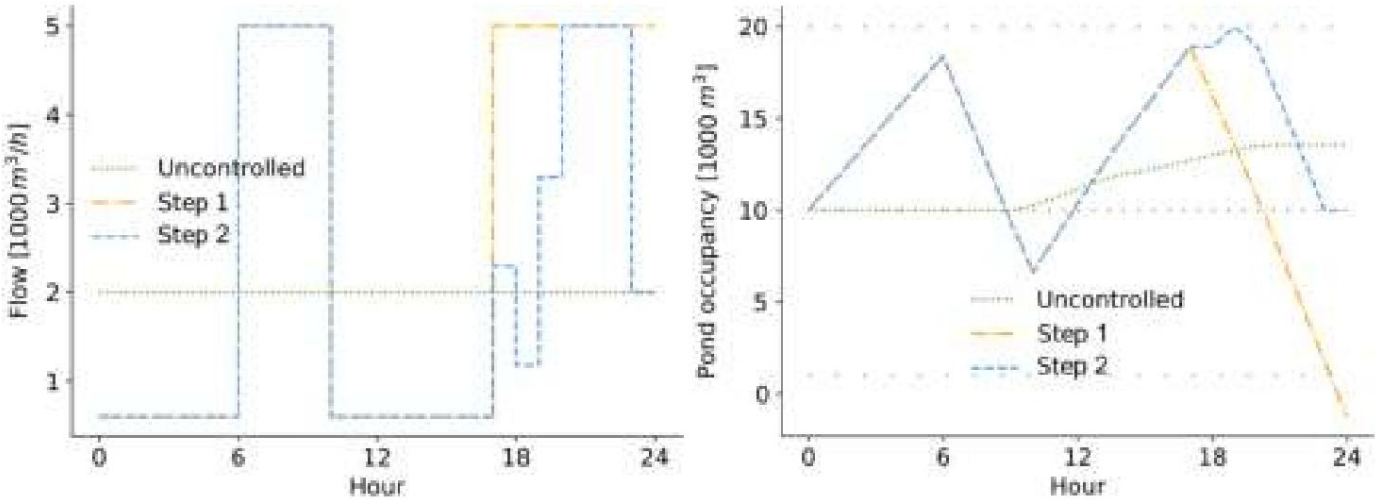


Abbildung 6. In Schritt 2 ermittelte Flussintensität für das Muster aus Abbildung 5. Für die ersten drei Perioden **ist keine Aktion erforderlich . der letzten Periode wird die Endbedingung** s(m) **verletzt, sodass der Fluss eingeschränkt werden muss.**

Schritt 3, globale Optimierung:

Eingabe: Preisprofil und lokal optimierter Wasserflussplan, berechnet in Schritt 2.
Ausgabe: global optimierter Flussplan.
Ziel: Neuplanung des Wasserflusses innerhalb des gesamten Optimierungshorizonts.
Schritt 2 führt zum optimalen Durchfluss innerhalb jedes Ertrags-/Akkumulationsintervalls und garantiert, dass Bedingung (7) für jedes Intervall nj erfüllt ist . **Möglicherweise besteht Spielraum für eine allgemeine Leistungssteigerung des Systems. Bei der Beobachtung der in Abbildung 6** dargestellten Ereignisse **kann man feststellen, dass in den Stunden 0–19 die Teichbelegung nicht monoton wächst, bis der Behälter voll ist. Die Füll- und Entleerungsintervalle werden durch die Zeitpunkte abgegrenzt, an denen der Teich gesättigt ist. Die Neuplanung innerhalb dieser Intervalle beeinflusst das endgültige Volumen der Behälter nicht, sodass es möglich ist, den Durchfluss zu erhöhen, wenn die Energie teuer ist, und ihn um denselben Wert zu verring**

Die maximale Anzahl der Iterationen in Schritt 3 ist ähnlich wie in Schritt 2. Das Ergebnis von Ein Beispiel für Schritt 3 ist in Abbildung 7 dargestellt.

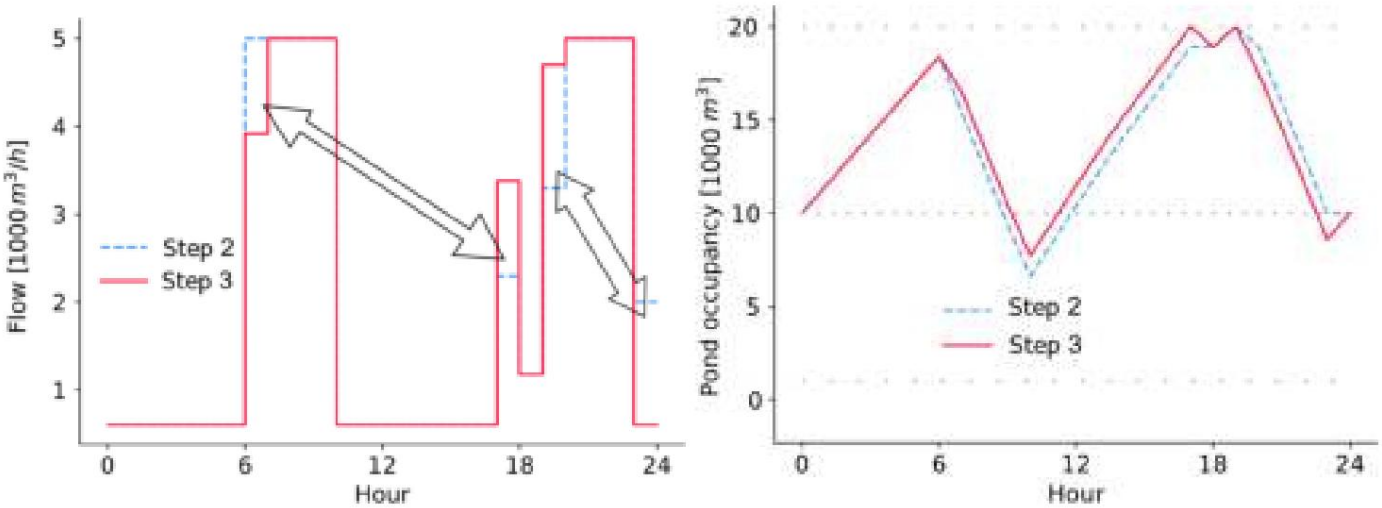


Abbildung 7. Ergebnis von Schritt 3 im Vergleich zum Ergebnis von Schritt 2 (Abbildung 6). Pfeile markieren die neu geplante Wassermenge.

Machine Translated by Google

Schlussbetrachtungen

Der Algorithmus wird immer dann ausgeführt, wenn sich die Eingangswerte ändern, z. B. wenn neue Energiepreise bekannt gegeben werden oder eine aktualisierte Wettervorhersage rj (k) **vorliegt. Die Ergebnisse müssen an alle betroffenen Eigentümer im HP-Netzwerk gesendet werden, da Schritt 2 des Algorithmus die Ausgangsdaten der vorgelagerten Anlagen benötigt. Der Algorithmus kann lokal mit kostengünstiger Hardware, z. B. einem Mikrocontroller, ausgeführt werden.**

5. Auswertung

Die Vier-Knoten-Topologie in Abbildung 3 wurde für die Bewertung des Algorithmus verwendet. Sie ist einfach genug, um die Eigenschaften des Algorithmus zu veranschaulichen und gleichzeitig die wichtigsten Herausforderungen eines Prosumer-HP-Netzwerks zu berücksichtigen, das sich aus der baumartigen Topologie und der nicht zu vernachlässigenden Verzögerung zwischen den Anlagen ergibt. Es wird kein gemeinsames Management angenommen – die Prosumer treffen die Steuerungsentscheidungen in ihren Anlagen unabhängig voneinander. Es ist nur die Übertragung der Steuerungsentscheidungen erforderlich. Formale Vorschriften können dies lei

Der Algorithmus aus Abschnitt 4 wurde anhand verschiedener Preisprofile aus [5], einschließlich der in Abbildung 2 dargestellten, bei Vorhandensein einer sich bewegenden Welle intensiven Regens und des entsprechenden Abflusses, der in Abbildung 8 dargestellt ist, getestet. Die Berechnungen wurden für jedes Profil und unterschiedliche Teichvolumina durchgeführt. In allen Fällen versorgt ein exogener Fluss konstanter Intensität f(k) **= [2, 3, 5, 5] × 103 [m3/h] das System. Andere Systemparameter** f_{max} **angenommen als = [5, 6, 10, 10] × 103 [m3/h],** f_{min} **sind = [0,5, 0,6, 1, 1] × 103 [m3/h], s(0) =** s(m), s_{max} **= 0,1 × s(0),** s_{max} **= 2 × s(0), wobei der Anfangswasserstand s(0) wie folgt festgelegt ist:**

Szenario 1 (kleine Staubecken): s(0) = s(m) **= [0,5, 0,75, 1,25, 1,25] × 104 [m3]; Szenario 2 (mittelgroße Staubecken): s(0) =** s(m) **= [1, 1,5, 2,5, 2,5] × 104 [m3]; Szenario 3 (große Staubecken): WPs sind an größere künstliche Seen angeschlossen, deren Sättigungswahrscheinlichkeit vernachlässigbar ist, s(0) =** s(m) **= [2, 3, 5, 5] × 104 [m3].**

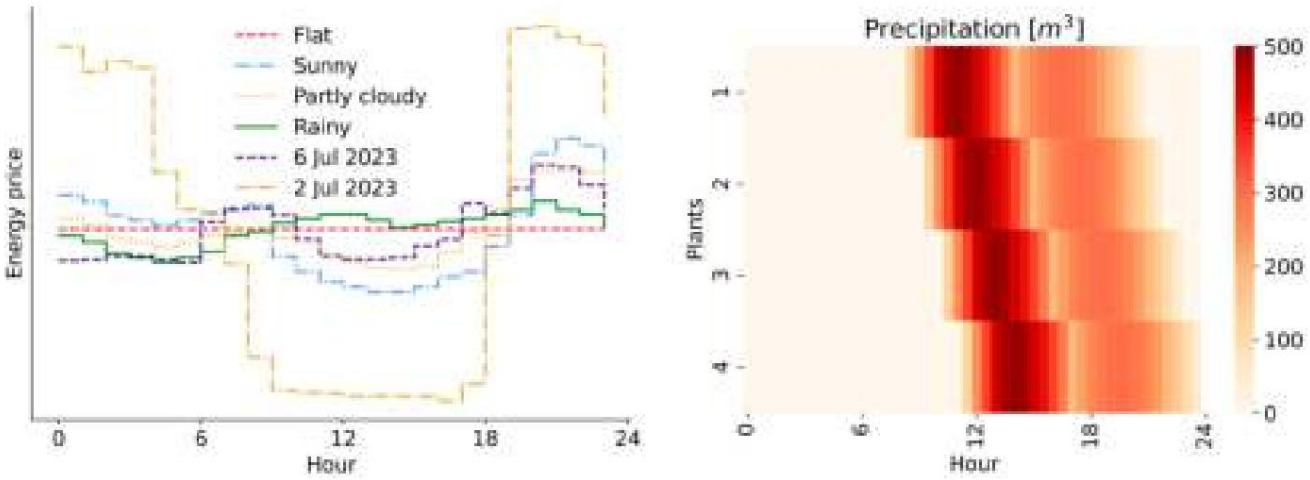


Abbildung 8. Beispielhafte Preisprofile, skaliert auf das gleiche durchschnittliche und angenommene Niederschlagsmuster, wobei der Abfluss zur Validierung der Ergebnisse f_{tt} **verwendet wurde.** f_{ff}

Obwohl der Algorithmus keine bestimmte Beziehung zwischen den Preisprofilen erwartet, verwenden alle HPs der Lesbarkeit halber dieselben Daten. Aus diesem Grund wird die Turbinenleistung auch als lineare Funktion der Strömungsintensität f_{ff} **(k)) =** const betrachtet.

Tabelle 1 fasst den Gewinn des Prosumers und den Gesamtgewinn des Netzbetreibers zusammen. Der Gewinn wird im Hinblick auf den Fall ermittelt, in dem die HPs nicht kontrolliert werden. Die Diskrepanz im Endzustand unter freiem Fluss wird zum Durchschnittspreis wieder aufgefüllt oder freigegeben.

Machine Translated by Google

Tabelle 1. Gewinnzuwachs [%] in Bezug auf den freien Durchfluss [m3] für die Preisprofile aus Abbildung 7 und verschiedene Reservoirgrößen.

Profil	Werk 1	Werk 2	Anlage 3	Anlage 4	Alle Anlagen	Anmerkung
Kleinteiche (Szenario 1)						
Sonnig	10,85	11,80	10,98	6,74	9,71	
Teilweise bewölkt	4,52	4,80	6,95	1,69	4,44	
Regnerisch	2,54	2.41	6,62	198,52	34,55	Wenn Anlage 4 unkontrolliert bleibt, dann Die Produktion fällt für 5 Stunden unter das Minimum.
6. Juli 2023	10,13	10,35	13,39	28,16	17.30	
2. Juli 2023	29,25	33,81	59,53	46,64	47,30	
Teiche mit mittlerem Volumen (Szenario 2)						
Sonnig	20.13	19,38	15,91	11,44	15,65	
Teilweise bewölkt	11,98	12.07	13,86	38,17	21.29	Wenn Anlage 4 unkontrolliert bleibt, dann Die Produktion fällt für 3 Stunden unter das Minimum.
Regnerisch	9.31	11.37	17.44	65,61	33,05	Wenn die Pflanzen 3 und 4 unkontrolliert bleiben, dann Produktion fällt für 5 h unter das Minimum (beide), und Anlage 3 erlebt während 2 Stunden eine Überschwemmung.
6. Juli 2023	14.47	13,77	14.25	49,92	25.15	Wenn Anlage 4 unkontrolliert bleibt, dann Die Produktion fällt für 4 Stunden unter das Minimum.
2. Juli 2023	68,50	68,34	69,79	82,83	74,03	Wenn Anlage 4 unkontrolliert bleibt, dann Die Produktion fällt für 2 Stunden unter das Minimum.
Großvolumige Teiche (Szenario 3)						
Sonnig	21.04	20,82	19,52	16,15	18,86	
Teilweise bewölkt	14,73	14,72	14.43	19.43	16,76	Wenn Anlage 4 unkontrolliert bleibt, dann Die Produktion fällt für 1 Stunde unter das Minimum.
Regnerisch	6,93	6,79	6,61	5,78	6.41	
6. Juli 2023	15,39	15,18	13,83	12,36	13,82	
2. Juli 2023	59,52	56,83	80,13	65,47	67,83	

Die in Tabelle 1 **dargestellten Daten zeigen, dass je größer das Teichvolumen, desto mehr Gewinn erhalten. Mit dem vorgeschlagenen Algorithmus jedoch, selbst für kleine Reservoirs, die Der relative monetäre Nutzen erreicht mehrere Dutzend und wächst mit der Preisvariabilität, wenn auch nicht linear. Die Daten zeigen auch, dass kleine Änderungen in** rj (k) **oder** pj (k) **zu einer erheblichen Diskrepanz im Ergebnis. Der Gewinn wächst, wenn der Abstand zwischen den HPs zunimmt, was auf einen größeren Unterschied zwischen** pj (k) **und** pj (k ÿ Tij) zurückzuführen ist. **Es ist bemerkenswert, dass Der Algorithmus neigt dazu, die Teiche gegen Mitternacht zu niedrigen Preisen zu füllen, erlaubt jedoch nicht für Überschwemmungen. Dieses Verhalten ist vorteilhaft, da die Teiche zur Bodenfeuchtigkeit beitragen und somit Linderung des Problems anhaltender Dürre in verschiedenen Regionen.**

Der Algorithmus wählt die maximale oder minimale Flussintensität f(k) **anstelle von Zwischenprodukte. Diese Eigenschaft erleichtert den Bau und die Wartung der Turbine durch Steigerung der Effizienz unter Nennbedingungen.**

Die mit dem vorgeschlagenen Algorithmus erzielten Ergebnisse, zusammengefasst in Tabelle 1, sind viel besser als die von Standardoptimierungsalgorithmen [29]. Darüber hinaus sind alle

Die flussabwärts gelegenen Wasserkraftwerke sind vor den sonst üblichen Überschwemmungen und Dürren geschützt (Szenario 2).

Der Ausgleichsmarkt kann sowohl den Prosumern als auch dem Netzbetreiber zusätzliche Gewinne bescheren Betreiber [16]. Wenn der Netzbetreiber, an den das Prosumer-Netz angeschlossen ist, feststellt, ein Mangel an Energie, kann es eine sofortige Wasserfreisetzung auslösen oder umgekehrt die im Überschussfall fließen.

6. Implementierungsaspekte

Vor dem Einsatz sollten noch einige weitere Aspekte beachtet werden:

1. Rechtliche Aspekte. Der Systemeinsatz erfordert die Zustimmung der Wasserwirtschaftsbehörde und entsprechende Vereinbarungen mit den Netzbetreibern.

2. Informationsaustausch. Der entworfene Algorithmus erfordert eine zuverlässige Kommunikation Plattform für Prosumer-Anlagen. Wir schlagen hierfür MQTT **[35]** vor.

Machine Translated by Google

3. **Wettervorhersage. Die Vorhersagen für den nächsten Tag sind derzeit nicht genau genug, um lokale Stürme mehrere Stunden im Voraus vorherzusagen. Radargestützte Bilder liefern jedoch innerhalb weniger Stunden genaue Niederschlagsschätzungen.**
4. **Depots als Echtzeitsystem. Jeder HP-Controller arbeitet unabhängig von den anderen. Daher ist das System potenziell anfällig für ungünstige Phänomene von verteilten Systemen wie Rushes oder Deadlocks. Das MQTT-Protokoll löst dieses Problem, ohne die Rechenlast zu erhöhen, da die Zeitplanberechnung nur den Bruchteil einer Sekunde dauert und die Prognoseaktualisierungen selten erfolgen.**
5. **Depots als Kontrollsystem. Das System arbeitet in einem offenen Kreislauf. Die Aktualisierungen des Zeitplans wirken sich nur auf die nachgelagerten HPs aus. Das System ist stabil bei begrenzter Eingabe und begrenzter Ausgabe (BIBO).**
6. **Sicherheit. Die vorgeschlagene Kommunikationsplattform entspricht den hohen Industriesicherheitsstandards Normen [35].**
7. **Rechenressourcen. Der vorgeschlagene Algorithmus erfordert keine High-End-Rechenressourcen zur Ausführung. Er läuft und kommuniziert effizient auf Low-End- und kostengünstigen Geräten.**

7. Schlussfolgerung

In diesem Artikel wird eine optimale Strategie zur Steuerung eines Systems verbundener HPs vorgestellt, die ein nachhaltiges, verteiltes Energiedepot bilden. Ein solches System kommt sowohl den Stromnetzbetreibern zugute, indem es Schwankungen in der Netzlast reduziert, als auch den Anlagenbesitzern, indem es ihre Einnahmen erhöht. Es kann mit geringen Kapital- und Betriebskosten eingesetzt werden. Darüber hinaus erhöht das System durch die Verlangsamung des Niederschlagsabflusses die Widerstandsfähigkeit der Region gegen Überschwemmungen und Dürren, was angesichts des Klimawandels von größter Bedeutung ist. Der Steuerungsalgorithmus geht nicht vom gleichen Preismuster oder dem gleichen Zeitplanhorizont aus, den verschiedene Interessengruppen verwenden. Die HP-Besitzer sind nicht zur Zusammenarbeit gezwungen – sie müssen ihre Steuerungsentscheidungen lediglich an die Nachbarn weitergeben. Im Gegensatz zum ungünstigen „ Wechselrichterkampf“ der Photovoltaik **[35] steigert die Zusammenarbeit der Prosumer im vorgeschlagenen Schema ihre Gewinne. Im Gegensatz zu Pumpspeicherkraftwerken verursacht die Energiespeicherung außerdem keine zusätzlichen Kosten, sodass selbst**

Der vorgestellte Algorithmus ist in Bezug auf numerische Stabilität und Robustheit besser als klassische Optimierungsverfahren [29]. Er kann mit vernachlässigbarem Rechenaufwand implementiert und asynchron an den HPs ausgelöst werden. Während das endgültige Einkommen vom Wetter, der Prognosegenauigkeit, der Variabilität der Energiepreise und anderen externen Bedingungen abhängt, liegt der Anstieg im Vergleich zu einem unkontrollierten Fall im Bereich von 20–70 % und sollte mit dem Einfluss von RSEs, insbesondere PV- und Wind-RSEs, wachsen.

Beiträge der Autoren: Konzeptualisierung, MM und PI; Methodik, MM; Software, MM; Validierung, MM und PI; formale Analyse, MM und PI; Untersuchung, MM; Ressourcen, MM und PI; Datenkuratierung, MM und PI; Schreiben – Vorbereitung des Originalentwurfs, MM; Schreiben – Überprüfung und Bearbeitung, PI; Visualisierung, MM und PI; Überwachung, PI; Projektverwaltung, PI; Mittelbeschaffung, PI Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Arbeit wurde im Rahmen des Projekts Nr. 2021/41/B/ST7/00108: „Robuste Steuerungslösungen für mehrkanalige vernetzte Ströme“ durchgeführt, finanziert vom Nationalen Wissenschaftszentrum Polen.

Erklärung zur Datenverfügbarkeit: Alle Daten sind öffentlich verfügbar [5].

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen. Die Geldgeber hatten keinen Einfluss auf die Gestaltung der Studie, auf die Erhebung, Analyse oder Interpretation der Daten, auf das Verfassen des Manuskripts oder auf die Entscheidung, die Ergebnisse zu veröffentlichen.

Verweise

1. **Kolosok, S.; Bilan; Vasylieva, T.; Wojciechowski, A.; Morawski, M. Eine Scoping-Überprüfung von erneuerbaren Energien, Nachhaltigkeit und der Umwelt.** Energien **2021,** 14, **4490. [CrossRef]**

Machine Translated by Google

2. **Rahman, MM; Oni, AO; Gemechu, E.; Kumar, A. Bewertung von Energiespeichertechnologien: Eine Übersicht.** Energy Convers. Manag. **2020,** 223, **113295.** [CrossRef]
3. **Wirfs-Brock, J. IE Questions: Why Is California Trying to Behead the Duck? 2014. Online verfügbar: https://insideenergy.org/2014/10/02/IE-Fragen-Warum-versucht-Kalifornien-die-Ente-zu-enthaupten/ (abgerufen am 12. Mai 2023).**
4. **Borkowski, D.; Cholewa, D.; Korze´n, A. Ein Laufwasserkraftwerk-Hydro-PV-Batterie-Hybridsystem als Energielieferant für lokale Lasten.** Energien **2021,** 14, **5160.** [CrossRef]
5. **PSE. Marktenergiepreise. Online verfügbar : https://www.pse.pl/dane-systemowe/funkcjonowanie-rb/raporty-dobowe-z-funkcjonowania-rb/podstawowe-wskazniki-cenowe-i-kosztowe/rynkowa-cena-energii-elektrycznej-rce (abgerufen am 1. August 2023).**
6. **Anaza, SO; Abdulazeez, MS; Yisah, YA; Yusuf, YO; Salawu, BU; Momoh, SU Mikrowasserkraft-Energieerzeugung – An Überblick.** Am. J. Eng. Res. **2017, 6, 5–12.**
7. **Sachdev, H.; Akella, AK; Kumar, N. Analyse und Bewertung kleiner Wasserkraftwerke: Eine bibliographische Übersicht.** Erneuern. Sustain. Energy Rev. **2015,** 51, **1013–1022.** [CrossRef]
8. **Yiwei, Q.; Jin, L.; Liu, F.; Dai, N.; Song, Y.; Chen, G.; Ding, L. Tertiäre Regulierung von kaskadierter Laufwasserkraft im Inselsystem für erneuerbare Energien unter Berücksichtigung der Dynamik mehrerer Zeitskalen.** IET Renew. Power Gener. **2021,** 15, **1778–1795.**
9. **Jamali, S.; Jamali, B. Optimaler Betrieb von Kaskaden-Wasserkraftsystemen: Auswirkungen auf die Wasserkraftsysteme des Großen Karun im Iran.** Appl. Water Sci. **2019, 9, 66.** [CrossRef]
10. **Wyszkowski, K.; Piwowarek, Z.; Paÿejko, Z. Maÿe Elektrowne Wodne w Polsce. Technischer Bericht, UN Global. 2022. Online verfügbar: https://ungc.org.pl/wp-content/uploads/2022/03/Raport_Male_elektrownie_wodne_w_Polsce.pdf (abgerufen am 1. Juni 2024). (Auf Polnisch).**
11. **YouTube. Constriction General Channel. 2024. Online verfügbar: https://www.youtube.com/watch?v=MBDV_Xj6Dj8 (Zugriff am 1. Juni 2024).**
12. **IEA Hydropower. Managementmodelle für Wasserkraft-Kaskadenreservoirs, Fallsammlung. Technischer Bericht. 2021. Online verfügbar: https://www.ieahydro.org/media/02e6ec8e/EAHydro_AnnexXIV_Management%20Models%20for%20Hydropower%20Kaskadenreservoirs_FALLZUSAMMENSTELLUNG.pdf (abgerufen am 1. Juni 2024).**
13. **Yu, Y.; Wu, Y.; Sheng, Q. Optimale Planungsstrategie von Kaskaden-Wasserkraftwerken im gemeinsamen Markt für Day-Ahead-Energie und Frequenzregulierung.** IEEE Access **2021, 9, 87749–87762.** [CrossRef]
14. **Zhang, H.; Chang, J.; Gao, C.; Wu, H.; Wang, Y.; Lei, K.; Long, R.; Zhang, L. Betrieb von Kaskadenwasserkraftwerken unter Berücksichtigung umfassende ökologische Wasseranforderungen.** Energy Convers. Manag. **2019,** 180, **119–133.** [CrossRef]
15. **Thaeer Hammid, A.; Awad, OI; Sulaiman, MH; Gunasekaran, SS; Mostafa, SA; Manoj Kumar, N.; Khalaf, BA; Al-Jawhar, YA; Abdulhasan, RA Eine Überprüfung der Optimierungsalgorithmen bei der Lösung von Planungsproblemen bei der Stromerzeugung aus Wasserkraft.** Energies **2020,** 13, **2787.** [CrossRef]
16. **Mielczarski, W.** Energiesysteme und -märkte; **Verband polnischer Elektroingenieure: Warszawa, Polen, 2018.**
17. **Singh, VK; Singal, S. Betrieb von Wasserkraftwerken – Eine Übersicht.** Renew. Sustain. Energy Rev. **2017,** 69, **610–619.** [CrossRef]
18. **Li, J.; Zhao, Z.; Li, P.; Mahmud, MA; Liu, Y.; Chen, D.; Han, W. Umfassende Nutzenbewertungen für die Integration von Pumpspeicherkraftwerke und schwimmende Photovoltaik.** Energy Convers. Manag. **2023,** 296, **117651.** [CrossRef]
19. **Wijesinghe, A.; Lai, L. Analyse und Entwicklung kleiner Wasserkraftwerke. In Proceedings der 4. Internationalen Konferenz zur Deregulierung und Umstrukturierung von Elektrizitätsversorgungsunternehmen und Energietechnologien (DRPT) 2011, Weihai, China, 6.–9. Juli 2011; Seiten 25–30.**
20. **Khomsah, A.; Laksono, AS Pico-Hydro als erneuerbare Energie: Lokale natürliche Ressourcen und Verfügbarkeit von Geräten bei der Stromerzeugung.** IOP Conf. Ser. Mater. Sci. Eng. **2018,** 462, **012047.** [CrossRef]
21. **Anderson, D.; Moggridge, H.; Warren, P.; Shucksmith, J. Die Auswirkungen der Laufwasserkraft auf die physikalischen und ökologischen Zustand der Flüsse.** Water Environ. J. **2015,** 29, **268–276.** [CrossRef]
22. **Nguyen, A.; Vu, V.; Hoang, D.; Nguyen, T.; Nguyen, K.; Nguyen, P.; Ji, Y. Aufmerksamkeitsorientiertes Ensemblemodell zur genauen Vorhersage von Abfluss und Wasserstand mit Trainingsdatenverbesserung.** Eng. Appl. Artif. Intell. **2023,** 126, **107073.** [CrossRef]
23. **Alvarez, G. Ein Optimierungsmodell für den Betrieb von Großwasserkraftwerken.** IEEE Lat. Am. Trans. **2020,** 18, **1631–1638.** [Querverweis]
24. **Shaw, AR; Sawyer, H.; LeBoeuf, E.; McDonald, M.; Hadjerioua, B. Wasserkraftoptimierung mithilfe künstlicher neuronaler Netzersatzmodelle eines hochpräzisen Hydrodynamik- und Wasserqualitätsmodells.** Water Resour. Res. **2017,** 53, **9444–9461.** [CrossRef]
25. **Ahmadianfar, I.; Samadi-Koucheksaraee, A.; Asadzadeh, M. Extrahieren nichtlinearer Betriebsregeln von Mehrbehältersystemen unter Verwendung eine effiziente Optimierungsmethode.** Sci. Rep. **2022,** 12, **18880.** [CrossRef]
26. **Bernardes, J., Jr.; Santos, M.; Abreu, T.; Prado, L., Jr.; Miranda, D.; Julio, R.; Viana, P.; Fonseca, M.; Bortoni, E.; Bastos, GS Optimierung des Wasserkraftbetriebs mittels maschinellem Lernen: Eine systematische Übersicht.** AI **2022, 3, 78–99.** [CrossRef]
27. **Marcelino, CG; Leite, GMC; Delgado, CADM; de Oliveira, LB; Wanner, EF; Jiménez-Fernández, S.; Salcedo-Sanz, S. Ein effizienter multiobjektiver evolutionärer Ansatz zur Lösung des Betriebs der Mehrspeichersystemplanung in Wasserkraftwerken .** Expert Syst. Appl. **2021,** 185, **115638.** [CrossRef]
28. **Linke, H. Ein modellprädiktiver Regler zur optimalen Wasserkraftnutzung von Flussreservoirs. In Proceedings der 2010 IEEE International Conference on Control Applications, Yokohama, Japan, 8.–10. September 2010; S. 1868–1873.**
29. **Morawski, M.; Ignaciuk, P. Ausgleich des Energiehaushalts in Prosumer-Wasserwerken unter expliziter Berücksichtigung der Durchflussverzögerung .** IFAC-PapersOnLine **2023,** 56, **11748–11753.** [CrossRef]

Machine Translated by Google

30. Ignaciuk, P.; Morawski, M. Optimale Steuerung von Kaskadenwasserkraftwerken als prosumerorientiertes dezentrales Energiedepot. Energien **2024,** 17, **469. [CrossRef]**

31. Danielson, C.; Borrellia, F.; Oliver, D.; Anderson, D.; Phillips, T. Eingeschränkte Flusskontrolle in Speichernetzwerken: Kapazität Maximierung und Balancieren. Automatica **2013,** 49, **2612–2621. [CrossRef]**

32. Basin, MV; Guerra-Avellaneda, F.; Shtessel, YB Bestandsverwaltungsproblem: Adaptives konvergentes kontinuierliches Reglerentwurf. IEEE Trans. Syst. Man Cybern. Syst. **2020,** 50, **4974–4983. [CrossRef]**

33. Ignaciuk, P. Totzeitkompensation in kontinuierlich überprüften Lagersystemen für verderbliche Waren mit mehreren Lieferalternativen. J. Process Control **2012,** 22, **915–924. [CrossRef]**

34. Sengul, C. Message Queuing Telemetry Transport (MQTT) und Transport Layer Security (TLS) Profile of Authentication and Authorization for Constrained Environments (ACE) Framework; RFC 9431. 2023. Online verfügbar: https://datatracker.ietf.org/ doc/rfc9431/ (abgerufen am 1. Juni 2024).

35. Gu, Y.; Green, TC-Stromversorgungssystemstabilität mit einer hohen Durchdringung von Wechselrichter-basierten Ressourcen. Proc. IEEE **2023,** 111, **832–853. [Querverweis]**

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Energieeffiziente neuronale Netzwerkbeschleunigung mit den meisten Signifikanter bitgeführter Näherungsmultiplikator

Pengfei Huang, Bin Gong, Ke Chen * und Chenghua Wang

Hochschule für Elektro- und Informationstechnik, Universität für Luft- und Raumfahrt Nanjing, Nr. 29 Jiangjun Avenue, Bezirk Jiangning, Nanjing 210016, China; pfhuang@nuaa.edu.cn (PH); bingong@nuaa.edu.cn (BG); chwang@nuaa.edu.cn (CW)
*** Korrespondenz: chen.ke@nuaa.edu.cn**

Zusammenfassung: Die steigenden Rechenleistungsanforderungen von Deep Learning und groß angelegten Modellen haben zu einem deutlichen Anstieg des Energieverbrauchs geführt und den dringenden Bedarf an energieeffizienteren Hardwaredesigns hervorgehoben. Diese Studie präsentiert eine neuartige Gewichtsapproximationsstrategie, die speziell für quantisierte neuronale Netzwerke (NNs) entwickelt wurde. Das Ergebnis ist die Entwicklung eines effizienten Näherungsmultiplikators, der die Verschiebung der signifikantesten Eins (MSO) nutzt. Im Vergleich zu energieeffizienten logarithmischen Näherungsmultiplikatoren und genauigkeitspriorisierten nicht-logarithmischen Näherungsmultiplikatoren erreicht unser vorgeschlagenes logarithmisches Design ein beispielloses Gleichgewicht zwischen Genauigkeit und Hardwarekosten . Im Vergleich zum Basis-Exaktmultiplikator weist unser innovatives Design bemerkenswerte Einsparungen auf, darunter eine Flächenreduzierung von bis zu 28,31 %, eine bemerkenswerte Reduzierung des Stromverbrauchs um 57,84 % und eine Verringerung der Verzögerung um 11,86 %. Experimentelle Ergebnisse zeigen, dass der vorgeschlagene Multiplikator bei Anwendung in neuronalen Netzwerken etwa 60 % Energie einsparen kann, ohne die Aufgabengenauigkeit zu beeinträchtigen. Gleichzeitig veranschaulichen Experimente mit dem Transformatorbeschleuniger und der Bildverarbeitung die erheblichen Energieeinsparungen, die **durch die Implementierung unseres vorgeschlagenen Designs für Large Language Models (LLMs) und Bildverarbeitungsaufgaben**

Schlüsselwörter: Näherungsberechnungen; großes Sprachmodell; Koeffizientenmultiplikator; logarithmischer Multiplikator

check for updates

Zitat: Huang, P.; Gong, B.; Chen, K.; Wang, C. Energieeffiziente Beschleunigung neuronaler Netze durch Verwendung des höchstwertigen bitgeführten Näherungsmultiplikators. Electronics **2024,** 13, **3034. https://doi.org/10.3390/Elektronik13153034**

Wissenschaftliche Herausgeber: Ke Xie, Shoue Chen und Xiaolong Wang

Erhalten: 5. Juli 2024
Überarbeitet: 30. Juli 2024
Akzeptiert: 31. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

Im Bereich der künstlichen Intelligenz wurden erhebliche Fortschritte erzielt, insbesondere durch die Entstehung tiefer neuronaler Netzwerke (DNNs) und großer Sprachmod
Allerdings bringen diese komplexen Modelle auch erhebliche Herausforderungen hinsichtlich Rechenleistung und Speicherbedarf mit sich, die auf ihre immer größer werdende Größe zurückzuführen sind, wie in [1] ausführlich analysi
Diese Herausforderungen sind für Edge-Geräte noch akuter, da diese durch inhärente Einschränkungen ihrer Rechenleistung eingeschränkt sind. Um den praktischen Einsatz neuronaler Netzwerke (NNs) auf Edge-Plattformen zu ermöglichen, wurden verschiedene softwarebasierte Modellkomprimierungstechniken wie Quantisierung und Bereinigung als praktika
Darüber hinaus haben jüngste Bemühungen, das Paradigma des Approximativen Rechnens (AxC) auf Schaltungsebene auszunutzen, überzeugend dessen Potenzial zur Verbesserung der Systemleistung bewiesen [2]. Diese Verbesserung ist auf die Verringerung der Rechenkomplexität und des Speicherbedarfs zurückzuführen, die durch den Ersatz präziser Rechenelemente durch ihre weniger genauen Äquivalente erreicht wird. Die wichtigsten Operationen von Rechenprozessoren, nämlich Multiplikation und Addition, wirken sich stark auf ihren Gesamtenergieverbrauch aus. Insbesondere verbraucht die Multiplikation mehr Energie als die Addition. Die wichtigste arithmetische Berechnung, die in neuronalen Netzwerken durchgeführt wird, ist die Multiplikations-Akkumulationsoperation (MAC). Da ein neuronaler Netzwerkbeschleuniger Millionen von Multiplikationsoperationen durchführt, haben ungefähre Multiplikatoren (AMs) aufgrund ihrer erheblichen Bedeutung für die Verbesserung der Effizienz und Genau

Machine Translated by Google

Frühere Forschungen zu für DNNs angepassten Näherungsmultiplikatoren haben ermutigende Ergebnisse hinsichtlich der Effizienz geliefert [6]. Beim Einsatz von NNs auf ressourcenbeschränkten Edge-Geräten bietet der Einsatz von Näherungsmultiplikatoren eine attraktive Möglichkeit, ein Gleichgewicht zwischen Genauigkeit und Energiekosten zu erreichen. Diese Leistung beruht auf der Abschwächung und Lockerung der strengen Anforderungen an perfekt präzise Multiplikationsergebnisse, wodurch die Erstellung von ungefähren Hardwaredesigns erleichtert wird, die ressourcen- und energieeffizienter sind. Die Studie in [7] demonstrierte die Effizienz von AxC in DNNs durch die Präsentation einer neuen AxC-DNN-Schicht, die Näherungsmultiplikatoren in den DNN-Designprozess einbezieht. In [8] wurden drei hybride Näherungsmultiplikatoren vorgeschlagen, indem NAND-Gatter anstelle von AND-Gattern verwendet und verschiedene Kombinationen der Kompressoren angewendet wurden, um unterschiedliche Kompromisse zwischen Genauigkeit und Hardwareeffizienz zu erzielen . In [9] wurde ein Open-Source-Design eines anpassbaren Näherungsmultiplikators vorgeschlagen, das von der Eingangsverteilung und -polarität Ziel ist es, das durchschnittliche Quadrat des absoluten Fehlers eines Näherungsmultiplikators auf der Grundlage der Wahrscheinlichkeitsverteilungen der aus der Zielanwendung extrahierten Operanden unter Berücksichtigung der Eingangspolarität zu minimieren. In [10] wurde eine Methodik zum Entwurf von Näherungsarraymultiplikatoren vorgestellt, die das Konzept der Übertragsvernachlässigung nutzt. Durch die strategische Auswahl der Positionen für die Übertragsvernachlässigung wird es möglich, eine kleine Anzahl von Überträgen zu vernachlässigen, wodurch eine höhere Geschwindigkeit und eine reduzierte Fläche erreicht werden, während gleichzeitig eine überlegene Gesamtgenauigkeit beibehalten und die Eignung für die beabsichtigte Anwend

Näherungsweise logarithmische Multiplikatoren bieten ein einfacheres Design, führen aber zu einem deutlich höheren Genauigkeitsverlust, während näherungsweise nicht-logarithmische Multiplikatoren einen geringeren Rechenfehler aufweisen, allerdings auf Kosten einer höheren Designkomplexität [11]. Um dem Bedürfnis nach reduziertem Energieverbrauch bei gleichzeitig akzeptablem Genauigkeitsverlust gerecht zu werden, wird in dieser Studie ein energieeffizienter Most Significan

Das MSAM kann eine Reihe hocheffizienter Multiplikatordesigns erhalten, indem der Approximationsfaktor (k) und der Präzisionsfaktor (m) dynamisch konfiguriert werden. Durch Nutzung des Approximationsfaktors (k) wird der Multiplikator strategisch in exakte und ungefähre Untermultiplikatoren unterteilt. Innerhalb des Approximationsbereichs wird das Prinzip des rekonfigurierbaren Multiplikators mit konstantem Koeffizienten (RCCM) [12] angewendet, wobei herkömmliche Multiplikationsoperationen durch Verschiebungs-, Additions- und Subtraktionsoperationen ersetzt werden. Dieser Zeitschriftenartikel baut auf unserem vorherigen Konferenzpapier [13] auf und erweitert es, wobei er mehrere wichtige Unterschiede und neue Beiträge bietet. Eine Zusammenfassung dieser Fortschritte finden Sie

unten: • Wir haben eine alternative und aggressivere Approximationsstrategie für den Multiplikator-Approximationsprozess entwickelt und damit die Grenzen der Kompromisse zwischen Genauigkeit und Effizienz erweitert. • Wir haben außerdem eine umfassende Fehleranalyse durchgeführt und die Kompromisse zwischen Genauigkeit und Energieverbrauch.

• Neben der Analyse gängiger Datensätze untersuchen wir auch die Bildverarbeitung mithilfe verschiedener AxC-Schemata. • Die

Vorteile unserer vorgeschlagenen ungefähren Multiplikatordesigns werden durch eine strenge theoretische Analyse und umfangreiche experimentelle Demonstrationen untermauert.

Der Rest dieses Dokuments ist wie folgt strukturiert: Abschnitt 2 bietet einen kurzen Überblick über den Hintergrund und verwandte Arbeiten. Abschnitt 3 beschreibt unsere innovative Methodik zur energieeffizienten Beschleunigung neuronaler Netzwerke im Detail, die einen höchstwertigen bitgesteuerten dynamischen Näherungsmultiplikator zur deutlichen Leistungssteigerung enthält. Abschnitt 4 bietet eine Fehleranalyse und Simulationsergebnisse. Abschnitt 5 schließt das Dokument ab.

2. Vorbemerkungen

Ein logarithmischer Multiplikator (LM) wandelt Zahlen im Binärformat in ein logarithmisches Format um und erleichtert so die Transformation von der Multiplikation zur Addition. Mitchells Algorithmus wurde erstmals in [14] vorgeschlagen. Gegeben seien zwei Binärzahlen A und B, wobei A =

Machine Translated by Google

und B = **2 k2 (1** + x2), **wandelt der logarithmische Multiplikator sie zunächst in ihre logarithmischen Formen wie folgt:**

$$\log2\ (A \cdot B) = \log2\ (2^{k1+k2}\ (1 + x1)(1 + x2)) = (k1 + k2) + \log2\ (1 + x1) + \log2\ (1 + x2), \quad (1)$$

wobei k1 und k2 die Positionen des höchstwertigen Bits (mit einem Wert von 1) in den Operanden A **bzw. B bezeichnen und x1 und x2 die Bruchteile von** A **und** B darstellen **(x1, x2 ÿ [0, 1]). In Mitchells Algorithmus** ist **der Ausdruck log2 (1** + x) **ungefähr gleich** x, wenn 0 **ÿ** x < 1 **ist. Daher kann der Logarithmus des Produkts ausgedrückt werden als wie folgt:**

$$\log2\ (A \cdot B) \approx k1 + k2 + x1 + x2. \quad (2)$$

Der logarithmische Multiplikator umfasst mehrere wichtige Komponenten: führende Einsendetektoren (LODs), logarithmische Konverter, einen Addierer und einen antilogarithmischen Konverter. Die LODs haben die Aufgabe, das höchstwertige 1-Bit innerhalb der eingegebenen Binärzahlen Anschließend erzeugen die Logarithmuskonverter die logarithmischen Darstellungen dieser Binärzahlen. Diese Logarithmuswerte werden addiert, um das Produkt zu erhalten, das dann mit dem Antilogarithmuskonverter wieder in eine Binärzahl dekodiert wird.

3. Der vorgeschlagene signifikanteste einsgetriebene Verschiebungs-Approximationsmultiplikator

In dieser Studie stellen wir einen äußerst bedeutenden einsgetriebenen, verschiebenden Näherungsmultiplikator vor , der einen logarithmischen Ansatz verwendet. Dieser neuartige Multiplikator erreicht eine verbesserte Genauigkeit bei gleichzeitiger Beibehaltung der Hardwarekosten, die mit denen herkömmlicher logarithmischer Näherungsmultiplikatoren vergleichbar oder sogar niedriger sind.

3.1. Die Architektur des Entwurfs

Die grundlegenden Operationen in neuronalen Netzwerken (NNs) umfassen Multiplikationen, insbesondere die Multiplikation von Gewichten w **mit Aktivierungen x. Die Architektur eines NN, das den vorgeschlagenen MSAM (Most Significant One-Driven Shifting Approximative Multiplier) verwendet, ist in Abbildung 1 dargestellt. Der MSAM besteht im Wesentlichen aus zwei Kernkomponenten: einem ungefähren Submultiplikator (ASM) und einem exakten Submultiplikator (ESM), erweitert mit zusätzlichen Shiftern und Addierern für ein einfaches und kohärentes Strukturdesign. Insbesondere wird das Gewicht** w **basierend auf der Signifikanz seiner Bits in zwei Teile aufgeteilt: Die höchstwertigen Bits bilden** wH, **das vom ESM verarbeitet wird, während die weniger signifikanten Bits manipuliert werden, um** w zu bilden $\hat{w}_L$, **das vom ASM verarbeitet wird. Anschließend wird das Ergebnis des ESM verschoben und eine Additionsoperation ausgeführt, um das Endergebnis r zu erzeugen.**

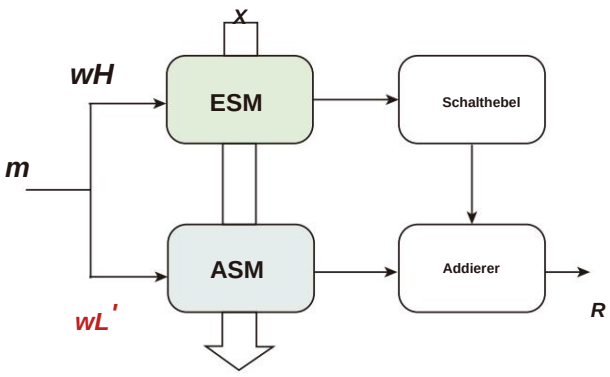


Abbildung 1. Die Architektur des bedeutendsten einmotorigen Shifting Approximative Multiplier (MSAM) -Blockdesigns.

Ein Vorzeichenkonverter ähnlich dem in [15] wird verwendet, um den MSAM so zu erweitern, dass er vorzeichenbehaftete Zahlen unterstützt , wie im linken Teil von Abbildung 2 dargestellt. Die internen Eigenschaften des MSAM-Blocks sind ebenfalls in Abbildung 2 dargestellt. Wir führen zwei zentrale Parameterfaktoren ein , die dem Multiplikator das Attribut der dynamischen Konfigurierbarkeit verleih

Machine Translated by Google

Bereich von Präzisionsanforderungen, die je nach Anwendungskontext schwanken können. Der erste ist der Näherungsfaktor, bezeichnet als k, der für die Partitionierung des Multiplikators verwendet wird. Diese Partitionierungsoperation trennt den n-Bit-Multiplikator in zwei unterschiedliche Segmente: k **Bits niedrigerer Ordnung und** (n ÿ k) **Bits höherer Ordnung. Dieses Partitionierungsschema zerlegt eine einzelne Multiplikationsoperation effektiv in zwei diskrete Multiplikationsprozesse, die den Multiplikanden und zwei Multiplikatoren mit kleineren Bitbreiten umfassen. Zweitens** ist **der Präzisionsfaktor, bezeichnet als m, in die Approximationsstrategie des ASM integriert und spielt eine entscheidende Rolle bei der Ermöglichung konfigurierbarer Präzisionseinstellungen im Rahmen des Untermultiplikators.**

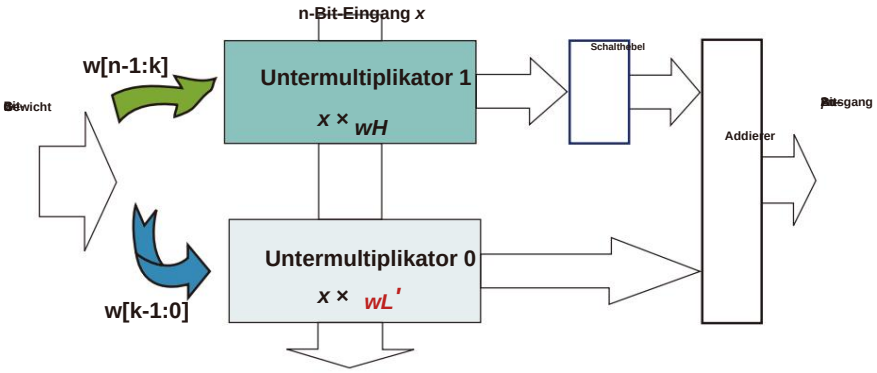


Abbildung 2. Erweiterung des vorgeschlagenen Ansatzes zur Unterstützung vorzeichenbehafteter Nummern, einschließlich der internen Funktion des MSAM-Blocks.

3.2. Die vorgeschlagenen Strategien zur dynamischen Gewichtsapproximation

Um im MSAM- Design ein Gleichgewicht zwischen Rechenleistung und Genauigkeit zu erreichen, führen wir mehrere auf die ASM-Komponente zugeschnittene Gewichtsapproximationsstrategien ein , die auf den Grundprinzipien des logarithmischen Multiplikators (LM) und des rekonfigurierbaren Multiplikators mit konstantem Koeffizienten (RCCM) basieren. Diese Strategien zielen darauf ab, verschiedene Anwendungsanforderungen zu erfüllen und die Leistung unter verschiedenen Bedingungen zu optimieren. Dazu gehört die Umwandlung von Multiplikationsoperationen in einfachere Rechenoperationen wie Verschiebungen, Addition oder Subtraktion. Die ESM-Komponente verwendet unterdessen konventionelle, präzise Multiplikationstechniken, um genaue Berechnungen sicherzustellen.
Um eine genaue kumulative Addition zu gewährleisten, wird die vom ESM generierte Ausgabe einer Reihe von Verschiebungsoperationen unterzogen, um sie mit der Ausgabe vom ASM abzugleichen. Diese Ausrichtung ist entscheidend für den kumulativen Additionsprozess, der letztendlich die endgültige Ausgabe des MSAM-Blocks bestimmt. Durch die Kombination dieser unterschiedlichen Strategien in den ASM- und ESM-Komponenten möchten wir die Gesamtleistung des MSAM-Designs optimier
Der Aktivierungswert x **dient als Multiplikand, während das Gewicht** w **die Rolle des Multiplikators bei der Multiplikation innerhalb eines neuronalen Netzwerks (NN) übernimmt. Um den ungefähren Submultiplikator (ASM) für neuronale Netzwerke zu optimieren, schlagen wir eine Methode vor , die die Gewichte des Modells in unterschiedliche Datenbereiche aufteilt und die Gewichte innerhalb dieser Bereiche als konstante Werte approximiert. Diese Sammlung ungefährer konstanter Werte, bezeichnet als** $W^{\ddot{y}}$ **entspricht effektiv dem konstanten Koeffizientensatz** C, **der im RCCM-Rahmenwerk [12] verwendet wird. Ein grundlegender Unterschied besteht darin, dass der vorgeschlagene Satz** $W^{\ddot{y}}$ **nur die Bestimmung einer einzigen Anfangskonstante,** w, **erfordert , während die übrigen Konstanten im Satz in Bezug auf diese Anfangskonstante abgeleitet werden. Darüber hinaus ermöglicht diese innovative Gewichtsapproximationsstrategie die dynamische Auswahl geeigneter approximierter Konstantenwerte basierend auf dem signifikantesten Wert (MSO) in den Eingangsgewichtsdaten.**

3.2.1. Eins-dominierende Approximationsstrategie (OAS)

Der Prozess der Näherung des einzelnen Multiplikators (w) **unter Verwendung der vorgeschlagenen Methode ist in Abbildung 3 dargestellt. Die Erkennung der führenden Einsen (LOD) wird verwendet, um die höchstwertige Eins (MSO) des Multiplikators w zu erkennen. Sobald die MSO identifiziert ist, werden die nachfolgend**

Machine Translated by Google

auf 1 gesetzt, während die restlichen Bits auf 0 gesetzt werden. Wenn die Position des MSO beim mso-ten Bit liegt (gezählt von 0, wobei mso **der Index des MSO-Bits ist), werden alle Zahlen im Bereich** w **ÿ [2** mso, 2mso+1 **ÿ** 1] **gemäß Gleichung (3) als konstanter Wert von eins angenähert.**

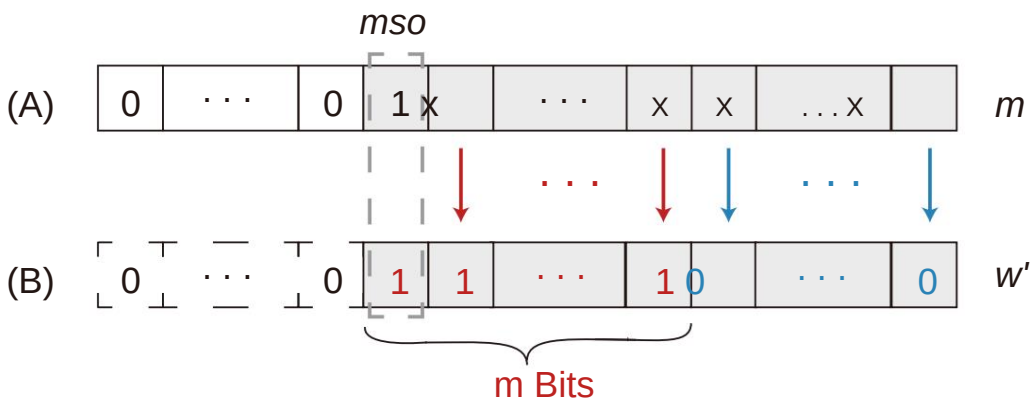


Abbildung 3. Beispiel für den den Multiplikator Eins dominierenden Näherungsprozess: (a) ursprüngliche Zahl; (b) angenäherte Zahl.

$$m^{ÿ} = \begin{cases} ÿ w_0, & mso < m \\ i ÿ 2^{mso} i=msoÿm+1, & mso \ ÿ\ m. \end{cases} \quad (3)$$

Der Wert von w **0** ^{ÿ} **wird durch den Faktor m bestimmt. Wenn beispielsweise** m = **1 ist, wird** $w_0^{ÿ}$ **auf 1b gesetzt ; währenddessen** wird w auf 11b **gesetzt ,** wenn m = **2 ist . Bei dieser Näherungsmethode wird der Multiplikator dynamisch in separate Datenbereichsintervalle aufgeteilt, basierend auf der Position des MSO in den Eingabedaten. Diese Intervalle werden dann mithilfe eines Satzes konstanter Koeffizienten angenähert.**

Wir bezeichnen m **als Präzisionsfaktor, da er die Anzahl der Näherungsintervalle steuert.**

Die Auswahl von m **aufeinanderfolgenden Bits, beginnend beim MSO, und deren Setzen auf 1 basiert auf zwei Hauptgründen. Erstens ermöglicht dies die Herstellung einer Verbindung zwischen dieser bestimmten Gruppe von Näherungswerten und dem anfänglichen Näherungskonstantenkoeffizienten w** durch **eine Abfolge von Verschiebungsoperationen. Infolgedessen können wir** durch Anbieten einer **anfänglichen Näherungskonstante** $w_0^{ÿ}$ von **0 die Multiplikationsausgaben der verbleibenden Koeffizienten in Bezug auf den Eingang x effektiv wiederherstellen. Diese Methode rationalisiert die Hardwareimplementierung, indem sie redundante Berechnungen reduziert, und basierend auf dem Produkt von** $x \cdot w_0^{ÿ}$, **eine sinnvolle Wahl des Wertes von** m die strategische Positionierung **der Näherungskonstanten in der Nähe des Durchschnittswerts innerhalb jedes Gewichtsintervalls, wodurch** die Präzision des Multiplikators maximiert wird. Im Wesentlichen ermöglicht es den Näherungskonstanten, sich zum Mittelwert des einzelnen Gewichtsbereichs hin zu bewegen, wodurch die **Gesamtgenauigkeit der Multiplikatorarchitektur verbessert wird .**

3.2.2. Null-dominierende Approximationsstrategie (ZAS)

Bei der oben beschriebenen einsdominierenden Näherungsstrategie werden alle m **Bits ab der MSO- Position aus einem wichtigen Grund auf 1 gesetzt: Diese angenäherte m-Bit-Zahl sollte dem Durchschnittswert aller k-Bit-Zahlen sehr nahe kommen. Im Extremfall, bei dem** m **= k ist, wird die angenäherte Zahl jedoch die größte, anstatt dem Mittelwert zu entsprechen. Um dieses Dilemma zu lösen, schlagen wir eine alternative nulldominierende Näherungsstrategie vor, wie in Abbildung 4 dargestellt. Unmittelbar nach der Position des MSO werden nicht alle verbleibenden Bits auf 1 gesetzt. Stattdessen wird das erste Bit unmittelbar nach der MSO-Position auf 0 gesetzt, während der Rest dieser Bits auf 1 gesetzt wird. Dadurch wird das Problem übermäßig großer Näherungszahlen effektiv vermieden und gleichzeitig die Nähe zum Mittelwert sichergestellt . Mit einer 0 als Trennzeichen ermöglicht dies auch die sinnvolle Anwendung von Carry- Disc-Techniken [10], wodurch Latenz und Stromverbrauch reduziert werden.**

Machine Translated by Google

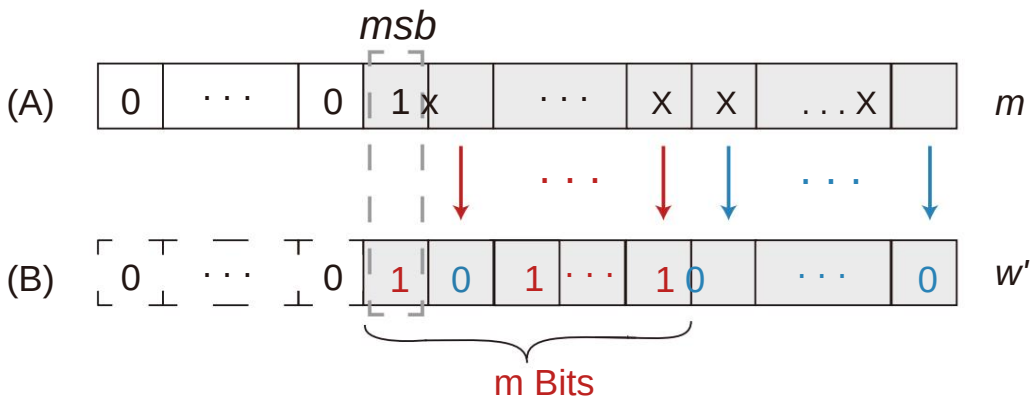


Abbildung 4. Beispiel für den Multiplikator-Null-dominierenden Näherungsprozess: (a) ursprüngliche Zahl und b) ungefähre Zahl.

3.2.3. Kompensation für die Null-dominierende Approximationsstrategie

Um die Vorteile der Näherung besser nutzen zu können, wählen wir normalerweise nicht m = **k. Empirisch gesehen führt die Einstellung** m ÿ **log2** (k) + **1 zu einem günstigen Gleichgewicht zwischen Energie Verbrauchs- und Genauigkeitsvorteile.**

Wie aus Tabelle 1 hervorgeht, ist im Hinblick auf den MSAM-42-Ansatz (d. h. k = **4 und** m = **2) die Das durchschnittliche Ergebnis mit dem OAS ist etwas größer als der Medianwert, während das durchschnittliche Das Ergebnis mit der ZAS ist deutlich kleiner als der Median.**

1 Tabelle 1. Mit MSAM-42 erzielte Ergebnisse unter Verwendung verschiedener Näherungsstrategien.

Die k Bits (Eingabe w[3 : 0])	Eins-Dominant Annäherung	Null-Dominanz Annäherung	Entschädigung für Null-Dominanz
0000	0000	0000	0000
0001 0010 0011	0011	0010	0101
0100 0101 0110 0111	0110	0100	0101
1001 1010 1011 1100 1101 1110 1111	1100	1000	1010

[1] **MSAM-42 gibt an, dass der Näherungsfaktor** k **4 und der entsprechende Präzisionsfaktor** m **2 ist.**

Obwohl dies den Energieverbrauch besser senken kann, kann es auch zu übermäßigen Fehler. Um diesbezüglich ein besseres Gleichgewicht zu erreichen, führen wir ein Kompensationsschema ein um Fehler zu verringern und sicherzustellen, dass die Gesamtgenauigkeit nicht merklich beeinträchtigt wird. Wie in Abbildung 5 dargestellt, ist die geschätzte Zahl mit ZOS wahrscheinlich kleiner als der tatsächliche Wert, durch Hinzufügen einer zusätzlichen Kompensation von 1 nach dem m-Bit, können wir die Fehlerdistanz effektiv reduzieren und gleichzeitig den zusätzlichen Energieverbrauch im Wesentlichen vernachlässigbar.

Machine Translated by Google

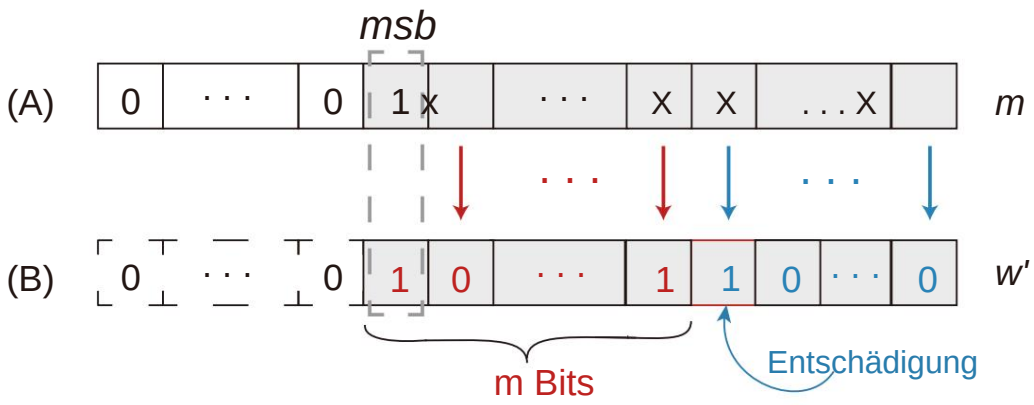


Abbildung 5. Beispiel der Multiplikator-Null-dominierenden Näherung mit Kompensationsprozess: (a) ursprüngliche Zahl und (b) angenäherte Zahl.

3.3. Umsetzung des Hardware-Designs

Der Kern der Exzellenz des MSAM-Designs liegt in der sorgfältigen Verwaltung des ASM, da die Verbesserung der Hardware den Gesamtenergieverbrauch der Multiplikationseinheit direkt und erheblich beeinflusst.

Wie in Abbildung 6 dargestellt, beginnt der Prozess mit einer Reihe von Schiebe- und Additionsoperationen , die gemeinsam das Produkt des Multiplikanden x **und der spezifischen Substitution gemäß den Faktoren** k **und m erzeugen. Die Position** (pos) **des höchstwertigen** w **für** W **0** L **legt das Steuersignal für signifikanten Einsbits innerhalb des Eingangsgewichts** wL, **das vom führenden Einsendetektor (LOD) identifiziert wird, nachfolgenden Multiplexer (MUX) fest. Dieses Signal steuert die Verschiebung in** r0tmp. **Anschließend wird der Multiplexer verwendet, um die Ausgabe r0 für das ungefähre Submultiplikatormodul (ASM) zu erzeugen. In der Zwischenzeit bietet der ESM einen genauen Ansatz zur Multiplikation. Schließlich durchläuft die Ausgabe des ESM eine Reihe von Schiebeoperationen und wird dann mit dem Ergebnis des ASM zusammengeführt, was letztendlich die endgültige Ausgabe des MSAM ergibt.**

Um vorzeichenbehaftete Operationen zu ermöglichen, kann die MSAM-Architektur durch die Integration von Vorzeichenkonvertierungsschaltkreisen vor und nach vorzeichenlosen Operationen erweitert werden, wie in Abbildung 2 dargestellt. Diese Schaltkreise verwenden das Komplement und führen eine ODER-Operation mit einer logischen 1b **im niederwertigsten Bit aus, was als Näherung für die Negation bei der Verarbeitung negativer Eingaben dient.**

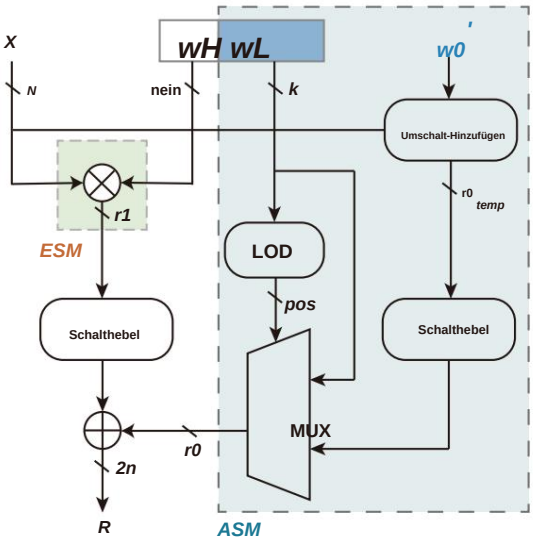


Abbildung 6. Hardware-Designimplementierung von MSAM.

Machine Translated by Google

4. Ergebnisse und Diskussion

Unser vorgeschlagenes MSAM-Design wird mit Verilog HDL implementiert, zusammen mit anderen fortgeschrittene Näherungsmultiplikatoren. Die Designs wurden mit dem Synopsys Design Compiler (DC) mit 28 nm CMOS-Technologie, und die Simulation wurde durchgeführt unter Verwendung von ungefähr 1.000.000 Zufallszahlensätzen, die den gesamten Eingaberaum abdecken von 8 × 8 Multiplikatoren (65.536 Kombinationen) bei einer Taktfrequenz von 200 MHz. Das VCS Für die Simulation wurde ein Tool verwendet, das rückannotierte Schaltaktivitätsdateien generierte um die Leistungsanalyse zu erleichtern. Die NMED-Metrik wurde verwendet, um den Fehler von die vorgeschlagenen Designs, die umfassend über den gesamten Eingaberaum hinweg bewertet werden unter Verwendung gleichmäßig verteilter Daten. Die vorgeschlagenen unsignierten MSAMs wurden verglichen mit Mitch-w [14], DR-ALM-7 [15], DRUM6 [16] und AxRM2 [17], während die signierten MSAMs wurden mit Mitch-w [14], **DR-ALM-7 [15], C-SSM** (m = **4) [18] und Approx-Booth [19] verglichen . Für klare und bequeme Vergleiche verweist MSAM auf die Verwendung der OSA Approximationsstrategie, MSAMZ bezeichnet die Verwendung der ZSA-Approximationsstrategie und MSAMC steht für ZSA mit zusätzlicher Vergütung.**

Die in diesem Manuskript verwendeten Abkürzungen sind in Tabelle 2 aufgeführt.

Tabelle 2. Abkürzungen für ungefähre Multiplikatordesigns.

Multiplikator	Techniken	Beschreibung
Genau	-	**Genauer Multiplikator.**
MSAM-km	**Logarithmisch ähnlich**	**Der vorgeschlagene signifikanteste einsgetriebene Verschiebungs-Approximationsmultiplikator unter Variation der Approximationsfaktoren k** und **Präzisionsfaktoren m.**
Mitch-w [14]	**Logarithmisch**	**Mitchells vorzeichenloser approximativer logarithmischer Multiplikator mit w-Bit-Operanden.**
Mitch-wÿ [14]	**Logarithmisch**	**Mitchells vorzeichenbehafteter ungefährer logarithmischer Multiplikator mit w-Bit-Operanden.**
DR-ALM-7 [15]	**Logarithmisch**	**Ungefährer nicht-iterativer vorzeichenloser logarithmischer Multiplikator mit** Li **-Bit-Operanden** (Li = **7).**
DR-ALM-7ÿ [15]	**Logarithmisch**	**Ungefährer nicht-iterativer vorzeichenbehafteter logarithmischer Multiplikator mit** Li **-Bit-Operanden** (Li = **7).**
TROMMEL6 [16]	**Logarithmisch**	**Unvoreingenommener Multiplikator mit** k(k = 6) **Bits im Betrieb vom Führenden.**
AxRM2 [17]	**Nicht logarithmisch**	**Ungefähre rekursive Multiplikatoren unter Verwendung von Hochleistungsbausteinen.**
C-SSM (m = **4) [18]**	**Nicht-logarithmisch**	**Ungefährer statischer segmentierter Multiplikator mit einem zusammenhängenden Segment von** m **Bits** (m = **4).**
Approx-Booth [19]	**Nicht-logarithmisch**	**Kostengünstiger approximativer Radix-4-Booth-Multiplikator mit vereinfachte Kompressor- und Encoder-Designs.**

4.1. Fehleranalyse

Deep-Learning-Beschleuniger mit approximativen Multiplikatoren und Addierern haben wurde entwickelt, um den Energieverbrauch und die Verzögerung zu verringern, allerdings mit einem akzeptablen Kompromiss bei der Genauigkeit [20]. Als allgemeines Prinzip gilt, dass ein höherer Grad an Beschleunigu führen zu einer größeren Verringerung der Genauigkeit. Die Fehlereigenschaften einer Näherung Multiplikator (AM) werden typischerweise durch umfassende Simulationen ermittelt, die umfassen alle denkbaren Eingangskombinationen oder alternativ durch Monte Carlo Simulationen, die verwenden zufällige Eingaben aus einer vorgegebenen Verteilung, wie etwa der Normalverteilung oder gleichmäßige Verteilung.

Bedenken Sie, dass Mÿ die Ausgabe des Näherungsmultiplikators bezeichnet und M die genaue Ausgabe. Die Fehlerdistanz (ED) **liefert eine genaue Messung der absoluten Differenz zwischen diesen beiden Ausgaben, die als** ED = **|M ÿ Mÿ | formuliert wird. Die Die relative Fehlerdistanz (ROT) drückt den Fehler als Prozentsatz im Verhältnis zum genauen Ergebnis aus. Darüber hinaus sind die mittlere Fehlerdistanz (MED) und die mittlere relative Fehlerdistanz (MRED) sind weit verbreitete Metriken zur Bewertung der Genauigkeit von Näherungsdesigns. Darüber hinaus**

Machine Translated by Google

Im Vergleich zu MED und MRED sind die normalisierte MED (NMED) und die normalisierte RED (NRED) stabile Metriken, die unabhängig von der Implementierungsgröße sind, was sie für die Bewertung von Näherungsdesigns unterschiedlicher Größe wertvoll macht. Wie in [21] gezeigt, hat **die NMED eines Näherungsmultiplikators (AM) einen erheblichen Einfluss auf den Ausgabefehler eines Neurons in einem neuronalen Netzwerk, insbesondere bei einer großen Anzahl vo**

Für das angegebene MSAM, wobei der Wert von k **konstant bleibt und die Konsistenz der Operandenbitbreite für alle Operationen sichergestellt ist, verwenden wir MRED und NMED, zwei effektive und weit verbreitete Metriken, um den durch den ASM-Block eingeführten Fehler zu bewerten. Diese Metriken dienen als unverzichtbares Mittel zur Bewertung der dem vorgeschlagenen Design innewohnenden Fehler, wobei der Schwerpunkt auf verschiedenen Konfigurat Abbildung 7 zeigt die MRED-Kurven für** k- **Werte von 3, 4 und 5, aufgetragen gegen Variationen des Präzisionsfaktors m, wodurch die Identifizierung deutlicher Trends innerhalb der Kurven möglich ist. Für verschiedene k-Werte erreicht die MRED ihr Minimum, wenn** m **1 oder 2 ist. Im Gegensatz zu anderen AMs ist der Präzisionsfaktor** m im Design des MSAM **nicht unbedingt besser, wenn er größer ist. Tatsächlich werden, wenn** m **log2** k + **1 überschreitet, übermäßig viele Bits auf 1 gesetzt, was zu einer erheblichen Abweichung des angenäherten Werts vom tatsächlichen Wert führt. Darüber hinaus führt ein größeres** m **dazu, dass mehr Bits an der Berechnung teilnehmen, wodurch mehr Energie verbraucht wird. Es ist erwähnenswert, dass die aus der NMED-Metrik erhaltenen Ergebnisse eng mit diesen experimentellen Ergebnissen übereinstimmen. Folglich haben wir in nachfolgenden Vergleichsanalysen und detaillierten Untersuchungen** m- **Werten von 1 oder 2 den Vorzug gegeben. Diese strategische Wahl wurde mit dem Ziel getroffen, ein optimales Gleichgewicht im Design zu erreichen und die Präzisionsanforderungen und Überlegungen zu den Hardware-Overheadkosten effektiv in Eink**

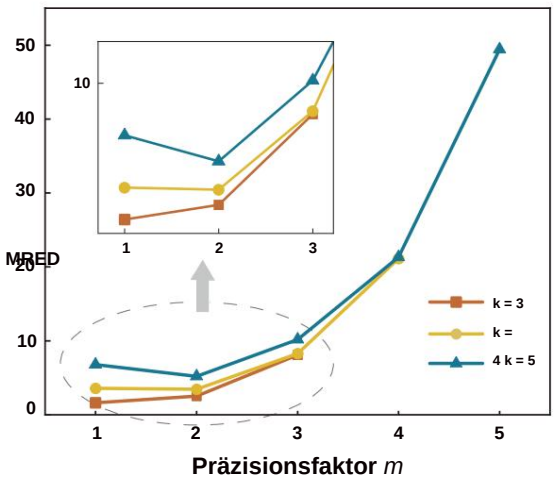


Abbildung 7. Mittlere relative Fehlerdistanzkurven von MSAM unter variierenden Näherungsfaktoren k **und Präzisionsfaktoren m.**

4.2. Ergebnisse der Synthese und

Simulation **Verschiedene MSAM-km- Multiplikatoren, vorhandene logarithmische Multiplikatoren und fortschrittliche nicht- logarithmische Multiplikatoren wurden bewertet, um ihre Hardwareleistung und Genauigkeitsmetriken einzuschätzen .**

In unserem Entwurf des Näherungsmultiplikators wird ein Operand in zwei Komponenten aufgeteilt: ESM und ASM, gesteuert durch den Näherungsfaktor k, der die Anzahl der angenäherten Bits bestimmt. Der Präzisionsfaktor m **stellt den Grad der Annäherung dar. Der Grad der Annäherung wird hauptsächlich durch k bestimmt; empirisch gesehen führt** ein größerer k zu einem höheren **Grad der Annäherung, einem einfacheren Entwurf und einem geringeren Energieverbrauch, während ein kleinerer** k **zu weniger Annäherung und geringeren Energieeinsparungen führt. Der Parameter** m **entscheidet, wie die** k **angenäherten Bits behandelt werden. Ein größerer** m **behält mehr Bits bei, aber es ist wichtig zu beachten, dass dies nicht bedeutet, dass die ursprünglichen präzisen Bits beibehalten werden; stattdessen werden** m **Bits verwendet, die nur aus Einsen bestehen . Daher führt ein größerer** m **nicht unbedingt zu einer höheren Präzision. Statistisch gesehen tendiert ein** m , **das näher an** k liegt, dazu, dass

Machine Translated by Google

Aus experimentellen und statistischen Analysen geht hervor, dass ein sinnvoller Wert für m bei etwa log2(k) **liegt .** m **näher an den Mittelwert aller** k-Bit-Zahlen **zu bringen . Die Einführung der Parameter** k **und** m **ermöglicht die Feinabstimmung des Designs, um das gewünschte Gleichgewicht zwischen Genauigkeit zu erreichen und Energieeffizienz. Diese zusätzlichen Parameter erfordern jedoch möglicherweise eine sorgfältige Berücksichtigung und Optimierung während der Entwurfsphase, um eine optimale Leistung sicherzustellen.**

Logarithmische Multiplikatoren verbrauchen weniger Energie, verlieren aber an Genauigkeit. Das Design mildert dies, indem es eine höhere Genauigkeit bei gleichzeitiger Beibehaltung einer niedrigeren Stromverbrauch als nicht-logarithmische Multiplikatoren. Ein Vergleich von NMED und dem Das Leistungsverzögerungsprodukt (PDP) für vorzeichenlose 8-Bit-Multiplikatoren ist in Abbildung 8 dargestellt. vorgeschlagene MSAMs sind mit einem roten Quadrat hervorgehoben. Während DR-ALM-7 [15] und Mitch- w [14] eine geringere PDP aufwiesen, führten sie zu einer höheren NMED. Umgekehrt zeigten DRUM6 [16] MSAM-21 produzierte weniger NMED, benötigte aber eine höhere PDP. Insbesondere AxRM2 [17], MSAM-52, MSAM-41 und MSAM-42 erreichten einen relativ niedrigeren PDP sowie kleinere NMED. Unter diesen ragen MSAM-41 und MSAM-42, wie durch den roten Kreis angezeigt, als die attraktivsten Designs und erreichen den optimalen Kompromiss zwischen PDP und NMED.

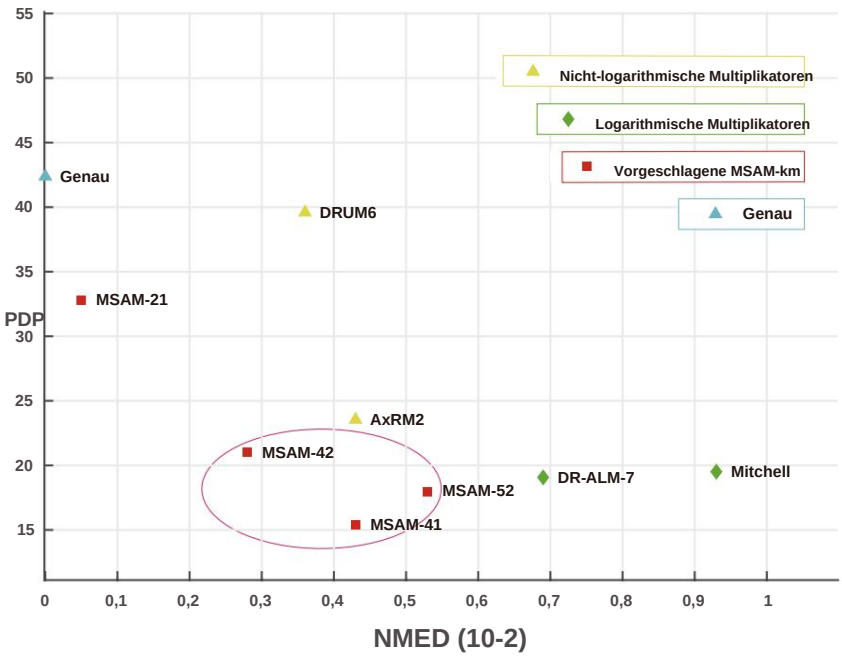


Abbildung 8. NMEDs und PDPs von vorzeichenlosen 8-Bit-Approximationsmultiplikatoren.

Die detaillierten Ergebnisse dieser Bewertung sind in Tabelle 3 zusammengefasst. Durch sorgfältige Durch die Konfiguration der Parameter k **und m erreichten wir eine Hardware-Performance vergleichbar mit (oder in einigen Fällen sogar höher) als die der Mitchell- [14] und DR-ALM-7- [15] Multiplikatoren (für Beispiel mit** km = **41 oder** km = **52). Beim Vergleich von MSAM-41 und MSAM-52 mit DR- ALM-7 ist selbst in Situationen, in denen die Leistung des Power Delay Product (PDP) die Im letzteren Fall wurden immer noch bemerkenswerte Verbesserungen in Bezug auf NMED beobachtet, und zwar 37,68% bzw. 23,19%. Im Vergleich zu AxRM2 [17]** war **MSAM-41 weist erhebliche Verbesserungen auf, darunter eine Reduzierung der Flächennutzung um 32,11 %. eine Verringerung der Verzögerung um 23,33 % und eine Verbesserung des PDP um 27,95 % bei gleichzeitiger Beibehaltung einer äquivalente NMED. Diese Ergebnisse unterstreichen die beeindruckende Hardware-Effizienz, die erreicht wurde mit dem MSAM-Design im Vergleich zu modernsten AM-Designs.**

Obwohl das ZAS ein vereinfachtes Hardware-Design bietet, ist es in der Lage, Energieverbrauch und Flächennutzung wird der NMED im Vergleich zur entsprechenden OAS-Strategie erhöht. Unser Design bietet mehrere Näherungsstrategi und OAS, um eine dynamischere Bandbreite an Kompromissen zu erreichen und damit den Anforderungen einer breiteres Anwendungsspektrum.

Machine Translated by Google

Tabelle 3. Leistungsvergleich von vorzeichenlosen 8-Bit-Approximationsmultiplikatoren.

Design	Fläche (um2)	Leistung (uW)	Verzögerung (ns)	PDP (fJ)	NMED (10ÿ2)
Genau	154,48	76,75	0,59	45,28	0
MSAM-21	147,29	59,55	0,59	35,13 (22 %)	0,05
MSAMZ-21	144,22	56,2	0,58	32,60 (28 %)	0,07
MSAM-32	143.01	52,3	0,52	27,20 (40 %)	0,17
MSAMZ-32	139,05	50,4	0,5	25,20 (44 %)	0,19
MSAM-31	129,53	48,39	0,52	25,16 (44 %)	0,17
MSAMZ-31	126,45	44,66	0,49	21,88 (52 %)	0,19
MSAM-42	125,5	41,78	0,52	21,73 (52 %)	0,28
MSAMZ-42	120,79	38,98	0,49	19,10 (58 %)	0,34
MSAM-41	111,89	36,86	0,46	16,96 (63 %)	0,43
MSAMZ-41	103,76	35,33	0,45	15,90 (65 %)	0,47
MSAM-52	110,75	32,36	0,52	16,83 (63 %)	0,53
MSAMZ-52	104,39	31,33	0,51	15,98 (65 %)	0,59
Mitchell [14]	103,45	30,01	0,65	19,51 (ÿ57 %)	0,93
DR-ALM-7 [15]	98,28	28,46	0,67	19,07 (-58 %)	0,69
TROMMEL6[16]	192,53	56,10	0,77	43,20 (ÿ5 %)	0,36
AxRM2 [17]	164,81	39,23	0,60	23,54 (ÿ48 %)	0,43

Die vorgeschlagenen Designs finden Sie im schattierten Abschnitt.

4.3. Auswertung, Simulation und Analyse

4.3.1. Leistungsbewertung tiefer neuronaler Netzwerke

Für eine umfassendere Bewertung der Leistung des Multiplikators bei Verwendung in In großen Modellen führten wir Experimente mit dem CIFAR-10-Datensatz durch und verwendeten drei repräsentative tiefe neuronale Netzwerke: VGG-19, bekannt für seine tiefen Faltungsschichten; ResNet-50, das das Problem des verschwindenden Gradienten in tiefen Netzwerken effektiv mildert durch die Verwendung von Restverbindungen; und DenseNet-121, das eine effiziente Nutzung durch dichte Konnektivitätsmechanismen. Wir führten Experimente durch mit Diese komplexen Netzwerke mit dem Ziel, die Leistung des Multiplikators gründlich zu bewerten und Anwendbarkeit. Darüber hinaus haben wir unsere experimentelle Analyse auf die anspruchsvolleren ImageNet-Datensatz durch Einbindung von SqueezeNet, einem kompakten und dennoch effizienten Deep Neural Netzwerkarchitektur für optimierte Leistung auf eingeschränkten Geräten.

Um die Leistung der Näherungsmultiplikatoren zu beurteilen, verwendeten wir die Anpassung des Frameworks [22] zur Erleichterung der Simulation und Bewertung verschiedener Multiplikatoren Designs. Wie in Abbildung 9 dargestellt, bestehen die Eingaben aus dem vortrainierten NN-Modell, der Quantisierung Konfigurationen und die Entwurfsdatei des ungefähren Multiplikators (basierend auf Verilog HDL). Zunächst wird der Schichtgraph mit den Parametern der Faltungsschicht aufgebaut und die angrenzende Batch-Normalisierungsschicht. Anschließend wird die Entwurfsdatei des ungefähren Multiplikator wurde analysiert. Basierend auf den Quantisierungskonfigurationen wurden die entsprechenden werden Näherungsebenen erzeugt. Diese Näherungsebenen ersetzen dann die Ebenen in der ursprünglichen Layergraphen. Anschließend wird der rekonstruierte Layergraph durch die AdaPT-Framework für Umschulung und Inferenz unter Einhaltung der angegebenen Umschulungsstrategie. Es ist bemerkenswert, dass für den Trainingsdatensatz nur eine repräsentative Teilmenge benötigt wird, die kann zu Kalibrierungszwecken etwa 10 % des ursprünglichen Trainingssatzes betragen. Schließlich Es werden Leistungsberichte mit detaillierten Angaben zur Genauigkeit und zum Energieverbrauch bereitgestellt.

Um die Genauigkeitsverschlechterung zu mildern , **die durch bestimmte Näherungen wie Pruning und approximative Multiplikation entsteht, können Modellneutraining und Feinabstimmung eingesetzt werden [23]. Diese Ansätze ermöglichen es dem Modell, sich anzupassen und resistenter gegenüber den Fehler, die durch solche Näherungen entstehen. Während der Inferenzphase werden diese Techniken ermöglichen es den AxC-Modellen, geringere Hardwarekosten zu verursachen und gleichzeitig deutlich höhere Genauigkeit. Um die Leistung der Näherungsmultiplikatoren effektiv zu bewerten, wurden Experimente mit dem CIFAR-10-Datensatz durch direkte Inferenz ohne**

Machine Translated by Google

Umschulung. Diese Methode ermöglichte eine schnelle Beurteilung der Leistung der Multiplikatoren. Für den ImageNet-Datensatz wurde inzwischen eine effizientere Umschulungsstrategie übernommen durch Auswahl von nur 10 % der Trainingsteilmenge. Diese Strategie zielte auf die Durchführung von Modell Feinabstimmung während der Trainingsphase, die eine Verbesserung der Genauigkeit und eine erhebliche Energieeinsparung ermöglicht. Reduktion während der eigentlichen Inferenzphase. Es zeigt die Wirksamkeit von Umschulung und Modelloptimierung zur Fehlerminderung und Verbesserung der Genauigkeit ohne die Belastung von teure Hardwareanforderungen.

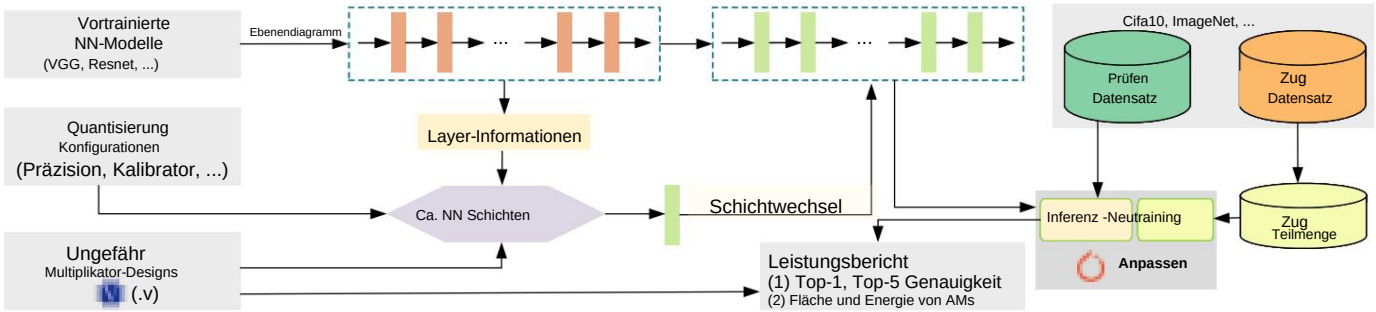


Abbildung 9. Auswertungsablauf basierend auf dem AdaPT-Framework.

Mehrere signierte Versionen des MSAM-Multiplikators wurden implementiert und gründlich verglichen mit dem präzisen Basis-4-Booth-Multiplikator, neben zwei logarithmischen Multiplikatoren— nämlich Mitchell [14] und DR-ALM-7 [15] – sowie zwei nicht-logarithmische Näherungsmodelle Multiplikatoren – C-SSM [18] und Approx-Booth [19]. Um einen fairen Vergleich zu gewährleisten, haben wir die Entwürfe des 8-Bit C-SSM (mit m **= 4) und Approx-Booth (mit** wG, wA **= 6, 6). Die Die Auswertung dieser signierten Multiplikatoren folgte streng dem zuvor beschriebenen Versuchsaufbau. Die Ergebnisse dieses umfassenden Vergleichs sind in Tabelle 4 dargestellt.**

Bei der Betrachtung der Ergebnisse des CIFAR-10-Datensatzes ist die MSAM-Approximation Der Multiplikator zeigte sowohl hinsichtlich der Genauigkeit als auch der Energieeinsparung die beste Leistung. Insbesondere übertraf der MSAM-31-Multiplikator den DR-ALM-7 nicht nur in Bezug auf der Netzwerkauswertungsgenauigkeit, sondern erzielte auch zusätzliche 9% Energieeinsparungen pro Einzelmultiplikator im Vergleich zu DR-ALM-7, was seine Effizienz unterstreicht. DR-ALM-7 hingegen Andererseits zeigte sich eine bessere Netzwerkleistung im Vergleich zu den anderen ungefähre Multiplikatoren, mit Ausnahme der MSAM-Varianten.

Tabelle 4. Vergleich der Hardwarekosten und Genauigkeiten von 8-Bit-Vorzeichenmultiplikatoren auf dem CIFAR10 und ImageNet-Datensätze.

Multiplikator	NMED-Bereich	Leistungsverzögerung (10ÿ2)	Energie (um2)	Energie (uW)	Energie (ns)	Energie (fJ)	CIFAR10 VGG-19	CIFAR10 ResNet-50	CIFAR10 DenseNet-121	ImageNet Top-1 Top-5
Genaue Basis 4	0	276,70	94,27	0,60	56,56	93,69 % 0,62 32,75 (ÿ42 %)	93,57 %	93,84 %	57,37 % 80,19 %	
Vorzeichenbehaftet MSAM-21		173,88	52,82	93,66 % 0,58 27,71 (ÿ51 %) 93,23 % 0,55			92,81 %	93,58 %	54,96 % 78,21 %	
0,34 Vorzeichenbehaftet MSAM-32		169,60	47,78	23,55 (ÿ58 %) 93,60 % 0,60 23,83 (ÿ58			91,34 %	93,30 %	52,02 % 76,00 %	
0,40 Vorzeichenbehaftet MSAM-31		156,11	42,82	%) 93,08 % 0,49 16,71 (ÿ70 % 91,24 %			93,06 %	93,64 %	54,01 % 77,60 %	
0,61 Vorzeichenbehaftet MSAM-42		152,08	39,71	0,71 27,95 (ÿ51 %) 89,64 % 0,73 28,64			91,71 %	92,93 %	50,37 % 74,35 %	
0,60 Vorzeichenbehaftet MSAM-41		138,47	34,10	(ÿ49 %) 93,29 % 0,43 35,42 (ÿ37 %)			91,26 %	91,46 %	43,31 % 68,00 %	
1,20 Mitchellÿ [14]	1,12	132,30	39,37	40,24 % 0,55 40,25 (ÿ29 %) 12,29 %			92,52 %	91,75 %	50,19 % 74,20 %	
DR-ALM-7ÿ [15]	0,78	127,13	39,23				92,24 %	93,37 %	53,92 % 77,38 %	
C-SSM (m = 4) [18]	3.62	276,32	82,37				61,17 %	73,28 %	2,22 % 7,47 %	
Approx-Stand [19]	0,17	218,23	73,18				10,65 %	9,99 %	17,44 % 36,33 %	

Die drei besten ungefähren Multiplikatoren, die die größte Anpassungsfähigkeit zeigten Zum ImageNet-Datensatz gehörten MSAM-21, MSAM-31 und DR-ALM-7. Diese Multiplikatoren zeigte eine außergewöhnliche Leistung in den verschiedenen Netzwerkarchitekturen und betonte ihre Vielseitigkeit und Robustheit. Für Anwendungen, bei denen hohe Netzwerkgenauigkeit im Vordergrund steht,

Machine Translated by Google

MSAM-21 stach als bemerkenswerter Leistungsträger hervor, der im Vergleich zu DR-ALM-7 eine Verbesserung der Top-1-Genauigkeit um 1,04 % und der Top-5-Genauigkeit um 0,83 % erreichte. Diese Verbesserungen sind signifikant und können einen erheblichen Einfluss auf die Gesamtleistung des Netzwerks haben. Die in Tabelle 4 dargestellten Ergebnisse **bieten einen umfassenden Vergleich der Hardwarekosten und Genauigkeiten der 8-Bit-Vorzeichenmultiplikatoren für die Datensätze CIFAR10 und ImageNet. Die MSAM- und DR-ALM-7-Multiplikatoren zeigten durchweg eine starke Leistung, wobei die MSAM-Varianten in Bezug auf Genauigkeit und Energieeffizienz besonders vielversprechend waren. Diese Erkenntnisse haben wichtige Auswirkungen auf das Design und die Optimierung von Multiplikatoren**

4.3.2. Leistungsbewertung des Transformer Matrix Multiplication Accelerator

Die Ressourcen- und Energieeinsparungen, die mit dem vorgeschlagenen Multiplikator beim Einsatz in LLMs auf Edge-Geräten erzielt werden, wurden verifiziert. Diese Verifizierung wurde durch die Anwendung des Multiplikators auf einen Transformer-Matrixmultiplikationsbeschleuniger erreicht, der speziell für die Verarbeitung von Matrizen der Größe 16 × 16 entwickelt wurde. Bei der Bewertung des Hardware-Overheads, wobei exakte Multiplikatoren als Vergleichsbasis dienten, konzentrierten wir uns auf die Beurteilung des Stromverbrauchs des Beschleunigers in Verbindung mit kombinatorischer Logik. Ref. [24] hebt die Möglichkeit hervor, Inferenzen in LLMs mit bis zu 175 Mrd. Parametern ohne Leistungseinbußen durchzuführen, indem LLM.int8() verwendet wird, wodurch die Multiplikation von über 99,9 % der Werte in 8 Bit ermöglicht wird. Ref. [25] demonstriert die Effizienz ungefährer Multiplikatoren im Kontext von Vision-Transformer-Modellen. Wie in Abbildung 10 dargestellt, umfasst unser Matrixmultiplikationsbeschleuniger zwei 8-Bit-Ganzzahlmatrizen, XI8 **und WI8. Um den Multiplikationsprozess zu beschleunigen, haben wir die exakt**

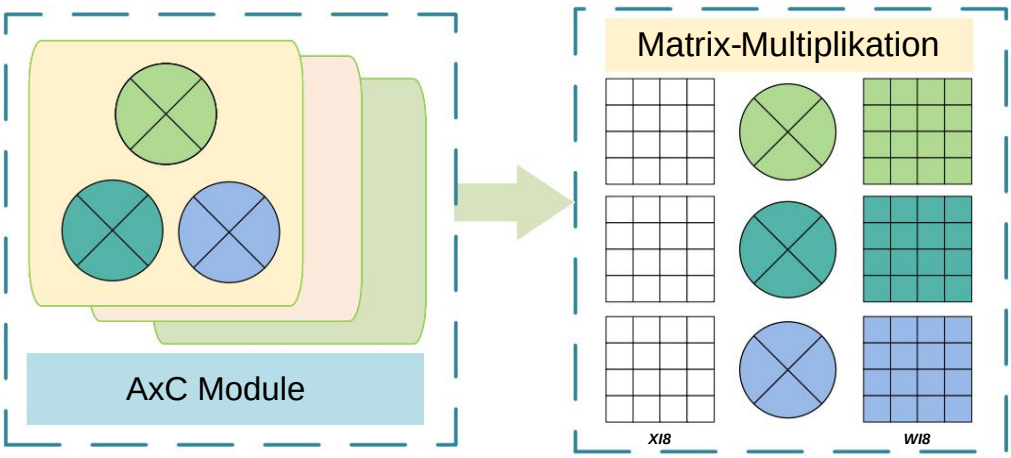


Abbildung 10. Matrixmultiplikationsbeschleuniger.

Tabelle 5 zeigt die experimentellen Ergebnisse des Matrixmultiplikationsmoduls im Transformer-Beschleuniger und verdeutlicht die effektive Reduzierung des Hardware-Overheads, die mit dem vorgeschlagenen Multiplikator erreicht wurde. Die Ergebnisse zeigen die Fähigkeit des Multiplikators, Ressourcen- und Energieverbrauch bei der Bereitstellung von LLMs auf Edge-Geräten zu optimieren. Verglichen mit dem Basisergebnis, das mit dem exakten Multiplikator erzielt wurde, weisen die MSAMs erhebliche Reduzierungen bei Fläche, Stromverbrauch und Latenz auf, wobei MSAMZ-41 **die deutlichsten Verbesserungen zeigt . Der Einsatz von Näherungsberechnungen hat die Möglichkeit gezeigt, die Effizienz von Transformern durch Minimierung ihrer Rechenkomplexität und ihres Speicherbedarfs** zu verbessern **[25]. In Anbetracht der Tatsache, dass große Modelle höhere Rechenanforderungen haben,** Daher sollten wir uns in unserer zukünftigen Forschung weiterhin mit einer **angemesseneren und rationaleren Integration dieser MSAMs befassen ,** mit dem Ziel, den Energieverbrauch zu senken und gleichzeitig **eine akzeptable Genauigkeit beizubehalten.**

Machine Translated by Google

Tabelle 5. Hardwareleistung des vorgeschlagenen Designs mit dem Transformatorbeschleuniger.

	Fläche (um2)	Kamm. Leistung (mW)	Verzögerung (ns)
Ausgangslage	90.391,64	2.07	0,89
MSAM-41	**67.618,91 (ÿ25,19 %)**	**1,82 (ÿ12,08 %)**	**0,79 (-11,24 %)**
MSAMZ-41	**47.618,35 (ÿ47,32 %)**	**1,46 (ÿ29,47 %)**	**0,75 (-15,73 %)**
MSAMC-41	**51.618,67 (ÿ42,89 %)**	**1,55 (ÿ25,12 %)**	**0,76 (-14,61 %)**

4.4. Fallstudie: Bildmatrix-Multiplikationsbeschleuniger

In unserem Ansatz führen wir eine Pixel-für-Pixel-Matrixmultiplikation zwischen einem Eingangssignal durch. Bild und ein Maskenbild, wobei der Wert des entsprechenden Pixels im Ausgabebild wird nach dem resultierenden Produkt bestimmt. Abbildung 11 zeigt die Bildmultiplikation Ergebnisse, bei denen die Bildqualität positiv mit dem Spitzensignal-Rausch-Verhältnis korreliert Verhältnis (PSNR); daher sind die PSNR-Werte für die generierten Bilder in deren jeweiligen Untertiteln. Die mit den meisten unserer MSAMs erzielten Ergebnisse deuten auf eine hohe Qualität hin, und zeigt ein sehr gutes PSNR im Vergleich zum exakten Multiplikator in Abbildung 11c. Aufgrund der relativ gleichmäßigen Verteilung der Bilddaten, wie sie unser Design beinhaltet ESM, es hat einen hohen Nutzen und eine hohe Robustheit im Kontext der Bildverarbeitung.



(a) Originalbild **(b) Maskenbild** **(c) Genauer Multiplikator (100 dB)** **(d) MSAM-21 (57,65 dB)**

(e) MSAM-32 (51,71 dB) **(f) MSAM-31 (50,29 dB)** **(g) MSAM-42 (47,48 dB)** **(h) AxRM2 (47,28 dB)**

(i) MSAM-52 (44,46 dB) **(j) MSAM-41 (41,41 dB)** **(k) DR-ALM-7 (38,10 dB)** **(l) Mitchell (37,98 dB)**

Abbildung 11. Die Ergebnisse der Multiplikation (a) eines Eingabebildes mit (b) einem Maskenbild unter Verwendung der ausgewählten Multiplikatoren.

Machine Translated by Google

4.5. Diskussion der vorgeschlagenen MSAMs **4.5.1.**

Vorteile und Einschränkungen Die

vorgeschlagenen MSAMs bieten erhebliche Energieeinsparungen und begrenzen gleichzeitig die Fehler von Näherungsmultiplikatoren durch den Einsatz von höchstwertigen Bit-geführten Näherungsstrategien . Diese Strategien bieten gegenüber anderen Arbeiten mehrere wesentliche Vorteile.

Erstens verwenden MSAMs höchstwertige bitgeführte Näherungstechniken, die eine Multiplikation durch Verschiebung ermöglichen. Dies führt zu einer deutlichen Verringerung von Fläche und Energieverbrauch, was unseren Ansatz für neuronale Netzwerkanwendungen hocheffizient macht . Zweitens garantiert unser Ansatz im Gegensatz zu logarithmischen Ansätzen, dass die Näherungsausgabe nicht einseitig verzerrt ist. Diese Funktion ermöglicht eine Kompensation zwischen verschiedenen Näherungen, was zu einer verbesserten Gesamtgenauigkeit und Zuverlässigkeit führt. Ein weiterer wichtiger Vorteil unserer Näherungsstrategie besteht darin, dass sie sich bei der Multiplikationsoperation auf nur einen Operanden konzentriert und dabei die Präzision des anderen Operanden vollständig bewahrt. Dieser besondere Ansatz gewährleistet effektiv Präzision und bietet eine deutliche Überlegenheit gegenüber den meisten anderen Arbeiten auf diesem Gebiet. Die präzise Kontrolle über die Näherungsstrategien und die Fähigkeit, die Präzision eines Operanden zu bewahren, tragen zur verbesserten Leistung bei neuronalen Netzwerkanwendun

Eine wesentliche Einschränkung der vorgeschlagenen MSAMs ist die relativ hohe Fehlerwahrscheinlichkeit, die durch die Verwendung vordefinierter Komponenten zum Ersetzen bestimmter Bits entsteht. Dies würde unsere Multiplikatoren in Szenarien, in denen hohe Präzision erforderlich ist, unbrauchbar machen. Trotz dieser Einschränkung besitzen die vorgeschlagenen MSAMs einen erheblichen praktischen Wert, da das Ausmaß der Fehler nicht erheblich ist und der Gesamtfehler für Batch-Operationen akzeptabel ist, da die Fehler sowohl positiv als auch negativ sein können .

4.5.2. Zukünftige Arbeit

Die Matrixmultiplikation dient als grundlegendes Bauelement in der Transformer- Architektur [25]. In Zukunft wollen wir das Design von Näherungsmultiplikatoren mit konzentrierten Fehlern weiter erforschen und eine für LLMs angepasste Näherungsmultiplikatorbibliothek erstellen . Näherungen auf Softwareebene, wie Netzwerke mit geringerer Komplexität und Pruning, erleichtern die energieeffiziente Beschleunigung von DNNs [23]. Während sowohl Software- als auch Hardware- Näherungstechniken unabhängig voneinander erhebliche Energievorteile für ein DNN bieten, beabsichtigen wir, diese Techniken in zukünftigen Bemühungen zur Maximierung der Energieeinsparungen in einem synergetischen Rahmen zu kombinieren. Wir hoffen, dass dieser Artikel mehr Forscher dazu anregt, in diesem Bereich zu forschen, was die Anwendung von MSAMs in komplexeren Szenarien mit **knappen Energiebudgets und strengen Genauigkeitsanforderungen ermöglicht .**

5. Schlussfolgerungen

In dieser Studie haben wir einen energieeffizienten dynamischen Näherungsmultiplikator mit dem höchstwertigen Bit zur Beschleunigung neuronaler Netze vorgeschlagen. Die Verwendung einer auf NNs und LLMs zugeschnittenen Gewichtsnäherungsstrategie führte letztendlich zu einem hocheffizienten Näherungsmultiplikator, der auf MSO-Verschiebungsprinzipien basiert. Das resultierende MSAM-Design erreicht geschickt ein harmonisches Gleichgewicht zwischen Präzision und Hardwarekosten. Bei Anwendung sowohl auf LLMs als auch auf DNNs führte unser Ansatz zu erheblichen Reduzierungen des Hardwareressourcenaufwands. Im Vergleich zu den herkömmlichen exakten Multiplikatoren, die in LLMs und DNNs verwendet werden, erreichte unser Design eine bemerkenswerte Reduzierung des Energieverbrauchs um 60 % ohne nennenswerte negative Auswirkungen auf die Aufgabengenauigkeit. Darüber hinaus haben wir eine Fallstudie zur Bildzusammenführung durchgeführt , und die Ergebnisse zeigen, dass unsere Multiplikatoren zur Erzeugung **qualitativ hochwertiger Bilder verwendet werden können .** Im Vergleich zum exakten Multiplikator schn

Beiträge der Autoren: Konzeptualisierung, PH; Software, BG; Validierung, BG; Schreiben – Originalentwurf , PH; Schreiben – Überprüfung und Bearbeitung, KC; Betreuung, CW. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Machine Translated by Google

Finanzierung: Diese Arbeit wurde durch Zuschüsse der National Natural Science Foundation of China (62101252 und 62134002) unterstützt.

Erklärung zur Datenverfügbarkeit: Auf angemessene Anfrage kann der entsprechende Autor die Daten bereitstellen, die die Ergebnisse dieser Studie stützen.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Verweise

1. **Raiaan, MAK; Mukta, MSH; Fatema, K.; Fahad, NM; Sakib, S.; Mim, MMJ; Ahmad, J.; Ali, ME; Azam, S. Eine Übersicht über große Sprachmodelle: Architekturen, Anwendungen, Taxonomien, offene Fragen und Herausforderungen.** IEEE Access **2024,** 12, **26839–26874. [CrossRef]**
2. **Rasch, MJ; Mackin, C.; Gallo, ML; Chen, A.; Fasoli, A.; Odermatt, F.; Li, N.; Nandakumar, SR; Narayanan, P.; Tsai, H.; et al. Hardwarebewusstes Training für umfangreiche und vielfältige Deep Learning-Inferenz-Workloads mithilfe von In-Memory-Computing-basierten Beschleunigern.** Nat. Commun. **2023,** 14, **5282. [CrossRef] [PubMed]**
3. **Towhidy, A.; Omidi, R.; Mohammadi, K. Über den Entwurf iterativer approximativer Gleitkommamultiplikatoren.** IEEE Trans. Comput. **2023,** 72, **1623–1635. [CrossRef]**
4. **Zhang, H.; Ko, SB Effiziente ungefähre Posit-Multiplikatoren für Deep Learning-Berechnungen.** IEEE J. Emerg. Sel. Top. Circuits Syst. **2023,** 13, **201–211. [CrossRef]**
5. **Sayadi, L.; Timarchi, S.; Sheikh-Akbari, A. Zwei effiziente ungefähre vorzeichenlose Multiplikatoren durch Entwicklung einer neuen Konfiguration für ungefähre 4:2-Kompressoren.** IEEE Trans. Circuits Syst. I Regul. Pap. **2023,** 70, **1649–1659. [CrossRef]**
6. **Schaefer, CJ; Taheri, P.; Horeni, M.; Joshi, S. Der Hardware-Einfluss von Quantisierung und Beschneidung für Gewichte in Spiking Neural Networks.** IEEE Trans. Circuits Syst. II Express Briefs **2023,** 70, **1789–1793. [CrossRef]**
7. **Pinos, M.; Mrazek, V.; Vaverka, F.; Vasicek, Z.; Sekanina, L. Beschleunigungstechniken für den automatisierten Entwurf von Näherungs Faltungsneuronale Netze.** IEEE J. Emerg. Sel. Top. Circuits Syst. **2023,** 13, **212–224. [CrossRef]**
8. **Ahmadinejad, M.; Moaiyeri, MH Energie- und qualitätseffiziente Näherungsmultiplikatoren für neuronale Netze und Bilder Verarbeitungsanwendungen.** IEEE Trans. Emerg. Top. Comput. **2022,** 10, **1105–1116. [CrossRef]**
9. **Li, Z.; Zheng, S.; Zhang, J.; Lu, Y.; Gao, J.; Tao, J.; Wang, L. Anpassbares Design eines ungefähren Multiplikators basierend auf Eingangsverteilung und Polarität.** IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **2022,** 30, **1813–1826. [CrossRef]**
10. **Amirafshar, N.; Baroughi, AS; Shahhoseini, HS; TaheriNejad, N. Carry ignoriert ungefähre Multiplikatoren.** IEEE Trans. Schaltungen Syst. I Regul. Pap. **2023,** 70, **4840–4853. [CrossRef]**
11. **Lotriÿc, U.; Pilipovi´c, R.; Buli´c, P. Ein hybrider Radix-4- und ungefährer logarithmischer Multiplikator für energieeffiziente Bildverarbeitung.** Elektronik **2021,** 10, **1175. [CrossRef]**
12. **Faraone, J.; Kumm, M.; Hardieck, M.; Zipf, P.; Liu, X.; Boland, D.; Leong, PH AddNet: Tiefe neuronale Netzwerke mit FPGA-optimierte Multiplikatoren.** IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **2019,** 28, **115–128. [CrossRef]**
13. **Gong, B.; Chen, K.; Huang, P.; Wu, B.; Liu, W. Die bedeutendsten einsgetriebenen dynamischen effizienten Multiplikatoren für große Sprachmodelle. In Proceedings des 2024 IEEE International Symposium on Circuits and Systems (ISCAS), Singapur, 19.–22. Mai 2024; S. 1–5.**
14. **Mitchell, JN Computermultiplikation und -division mit binären Logarithmen.** IRE Trans. Electron. Comput. **1962,** EC-11, **512–517.**
15. **Yin, P.; Wang, C.; Waris, H.; Liu, W.; Han, Y.; Lombardi, F. Entwurf und Analyse energieeffizienter dynamischer Bereichsnäherungsmodelle logarithmische Multiplikatoren für maschinelles Lernen.** IEEE Trans. Sustain. Comput. **2020, 6, 612–625. [CrossRef]**
16. **Hashemi, S.; Bahar, RI; Reda, S. DRUM: Ein dynamischer, unvoreingenommener Multiplikator für Näherungsanwendungen. In Proceedings der 2015 IEEE/ACM International Conference on Computer-Aided Design (ICCAD), Austin, TX, USA, 2.–6. November 2015; S. 418–425.**
17. **Waris, H.; Wang, C.; Xu, C.; Liu, W. AxRMs: Ungefähre rekursive Multiplikatoren mit Hochleistungsbausteinen.** IEEE Trans. Emerg. Top. Comput. **2021,** 10, **1229–1235. [CrossRef]**
18. **Strollo, Hauptversammlung; Napoli, E.; De Caro, D.; Petra, N.; Saggese, G.; Di Meo, G. Näherungsmultiplikatoren mit statischer Segmentierung: Fehleranalyse und Verbesserungen.** IEEE Trans. Circuits Syst. I Regul. Pap. **2022,** 69, **2449–2462. [CrossRef]**
19. **Park, G.; Kung, J.; Lee, Y. Vereinfachte Kompressor- und Encoder-Designs für kostengünstige ungefähre Radix-4-Booth-Multiplikatoren.** IEEE Trans. Circuits Syst. II Express Briefs **2022,** 70, **1154–1158. [CrossRef]**
20. **Jiang, H.; Santiago, FJH; Mo, H.; Liu, L.; Han, J. Näherungsarithmetische Schaltungen: Eine Übersicht, Charakterisierung und aktuelle Anwendungen.** Proc. IEEE **2020,** 108, **2108–2135. [CrossRef]**
21. **Mo, H.; Wu, Y.; Jiang, H.; Ma, Z.; Lombardi, F.; Han, J.; Liu, L. Lernen der Fehlermerkmale von ungefähren Multiplikatoren für neuronale Netzwerkanwendungen.** IEEE Trans. Comput. **2024,** 73, **842–856. [CrossRef]**
22. **Danopoulos, D.; Zervakis, G.; Siozios, K.; Soudris, D.; Henkel, J. AdaPT: Schnelle Emulation von ungefähren DNN-Beschleunigern in PyTorch.** IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. **2023,** 42, **2074–2078. [CrossRef]**
23. **Sanyal, S.; Negi, S.; Raghunathan, A.; Roy, K. Approximate Computing für Machine Learning-Workloads: Eine Schaltkreis- und Systemperspektive . In** Approximate Computing; **Liu, W., Lombardi, F., Hrsg.; Springer International Publishing: Cham, Schweiz, 2022; S. 365–395.**

Machine Translated by Google

24. Dettmers, T.; Lewis, M.; Belkada, Y.; Zettlemoyer, L. GPT3.int8(): 8-Bit-Matrixmultiplikation für Transformatoren im großen Maßstab. In Proceedings of the Advances in Neural Information Processing Systems, New Orleans, LA, USA, 28. November–9. Dezember 2022.

25. Danopoulos, D.; Zervakis, G.; Soudris, D.; Henkel, J. TransAxx: Effiziente Transformatoren mit Näherungsberechnungen. arXiv **2024, arXiv:2402.07545.**

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Eine einstufige elektronische Beleuchtungstreiberschaltung mit SiC Schottky-Dioden zur Versorgung einer Tief-Ultraviolett-LED Desinfektions- und Sterilisationslampe

Chun-An Cheng, Chien-Hsuan Chang, Hung-Liang Cheng, En-Chih Chang *, Long-Fu Lan, Sheng-Hong Hou und Cheng-Kuan Lin

Fakultät für Elektrotechnik, I-Shou-Universität, Kaohsiung 84001, Taiwan;
cacheng@isu.edu.tw (C.-AC); chchang@isu.edu.tw (C.-HC); hlcheng@isu.edu.tw (H.-LC);
isu11001004m@cloud.isu.edu.tw (L.-FL); isu11001001m@cloud.isu.edu.tw (S.-HH);
isu11001014m@cloud.isu.edu.tw (C.-KL)
*** Korrespondenz: enchihchang@isu.edu.tw; Tel.: +886-7-6577711 (Durchwahl 6642)**

Zusammenfassung: Vor kurzem ist **auf den Märkten für Raumreinigung und Oberflächensterilisation eine neue Art von Lichtquelle erschienen, nämlich eine Leuchtdiode (LED) für tiefes Ultraviolett .** In diesem Artikel **wird ein neuer Typ eines elektronischen Beleuchtungstreibers zur Versorgung einer LED-Sterilisationslampe für tiefes Ultraviolett vorgeschlagen und entwickelt. Der Hauptschaltkreis kombiniert einen Abwärtswandler und einen Sperrwandler zu einem einstufigen, einstufigen Abwärts-Sperrwandler mit AC/DC-Stromrichter und Leistungsfaktorkorrektur. Darüber hinaus nutzt der vorgeschlagene elektronische Beleuchtungstreiber eine SiC-Schottky-Diode mit großem Bandabstand als Ausgangsdiode, um die Verluste der Leistungsdiode zu senken, und recycelt die in der Streuinduktivität des Transformators gespeicherte Energie, um die Schaltungseffizienz zu verbessern. Die Magnetisierungsspule im Inneren des vorgestellten AC/DC-Stromrichters ist für den Betrieb im diskontinuierlichen Leitungsmodus (DCM) ausgelegt, der auf natürliche Weise eine Leistungsfaktorkorrektur (PFC) ermöglicht. Ein einstufiger Prototyp-Treiber mit einer Nennleistung** von 3,6 W (90 V/40 mA) wurde entwickelt und implementiert, um eine LED- **Desinfektions- und Sterilisationslampe für tiefes Ultraviolett bereitzustellen . Versuchsergebnisse zeigen, dass der gemessene Leistungsfaktor (PF) größer als 0,9 ist und die gemessene Gesamtverzerrung (THD) des Eingangsstroms bei einer Eingangsnetzspannung von 110 V weniger als 18 % beträgt. Darüber hinaus beträgt der gemessene Welligkeitsfaktor der Ausgangsspannung weniger als 1 % und der Welligkeitsfaktor des Ausgangsstroms weniger als 4 %. Darüber hinaus erreicht der vorgeschlagene einstufige elektronische Beleuchtungstreiber zur Versorgung einer tief ultravioletten LED-Desinfektions-**

Schlüsselwörter: elektronischer Beleuchtungstreiberschaltkreis; tief ultraviolette LED; Desinfektions- und Sterilisationslampe; Leistungsfaktorkorrektur; gesamte harmonische Verzerrung

check for updates

Zitat: Cheng, C.-A.; Chang, C.-H.; Cheng, H.-L.; Chang, E.-C.; Lan, L.-F.; Hou, S.-H.; Lin, C.-K. Eine einstufige elektronische Beleuchtungstreiberschaltung mit SiC-Schottky-Dioden zur Versorgung einer Desinfektions- und Sterilisationslampe mit tiefem Ultraviolettlicht. Electronics **2024,** 13, **3048. https:// doi.org/ 10.3390/electronics13153048**

Akademische Herausgeber: Yimeng Zhang, Lejia Sun und Yuming Zhang

Empfangen: 24. Juni 2024
Überarbeitet: 27. Juli 2024
Akzeptiert: 29. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

Mikroorganismen wie Bakterien, Viren, Zysten und Schimmelpilze nutzen Keimung oder Sporenproduktion zur Vermehrung. Wenn Mikroorganismen bestrahlt oder Desinfektions- und Sterilisationslampen ausgesetzt werden, dringt das Licht durch die Zellwände und zerstört ihre DNA, wodurch ihre DNA verändert wird, sodass sie unschädlich und unfähig zur Vermehrung werden oder sogar sterben, wodurch die Wirkung der Desinfektion und Sterilisation erreicht wird. Sterilisations- und Desinfektionslampen werden in einer Vielzahl von Anwendungen in verschiedenen Sektoren eingesetzt, darunter Gesundheitseinrichtungen , die Lebensmittel- und Getränkeindustrie, öffentliche Räume und Transportmittel, Gewerbe- und Bürogebäude, Wohngebäude, Gastgewerbe, Bildungseinrichtungen, Einzelhandels- und Lebensmittelgeschäfte, Wasseraufbereitung und Landwirtschaft. Diese vielfältigen Anwendungen unterstreichen die Vielseitigkeit und Wirksamkeit von Desinfektions- und Sterilisationslampen bei der Förderung von Gesundheit und Sicherheit in verschiedenen Umgebungen [1–4]. Die traditionelle Lichtquelle für Sterilisation und Desinfektion ist die ultraviolette (UV) Quecksilberlampe, die die Vorteile eines niedrig

Machine Translated by Google

nämlich Siliziumkarbid (SiC) und Galliumnitrid (GaN), die sich durch hohe Energieeffizienz und geringen Energieverbrauch auszeichnen, sind entstanden. Darüber hinaus können Leistungsgeräte aus Halbleitermaterialien mit großem Abstand bei höheren Spannungen betrieben werden, Gen- Restaurants und Schulen. Die Nachteile von UV-Quecksilberlampen sind jedoch lange mehr Leistung und geringere Energieverluste, bei deutlich geringerer Aufwärmzeit, kurzer Lampenlebensdauer und einem Quecksilbergehalt, der die Anforderungen nicht erfüllt Leistungsbauelemente aus Silizium-Halbleitermaterialien. Siliziumkarbid (SiC) ist ein Kom- Umweltanforderungen. UV-Licht ist für das menschliche Auge unsichtbar und hat eine Wellenlänge Pfund Halbleitermaterial bestehend aus Silizium (Si) und Kohlenstoff (C). SiC hat eine untere von 100 bis 400 Nanometern. Je nach Wellenlänge kann UV in drei unterteilt werden Widerstand der Driftschicht als Si, erfordert keine Leitungsmodulation, hat eine Energiearten : langwelliges UV- A (Wellenlängen von 320–400 nm), mittelwelliges UV-B (Wellenlängen Spalt etwa 3-mal breiter als der von Si und hat eine Isolationsbruchfeldstärke von 280–320 nm) und kurzwelliges UV-C (Wellenlängen von 100–280 nm). Kurzwelliges UV-C, etwa 10-mal höher als die von Si. Im Vergleich zu den Hochspannungs-Si-PN-Dioden und kann auch als tiefes Ultraviolett bezeichnet werden, hat desinfizierende und keimtötende Eigenschaften. Hochgeschwindigkeits-Si-Fast-Recovery-Dioden, SiC-Schottky-Dioden aus Halbleitern mit großem Abstand Die Verwendung von kurzwelligem ultraviolettem UV-C- oder tiefem ultraviolettem Licht zerstört die DNA oder tor Materialien haben hervorragende Hochgeschwindigkeits- und Hochspannungseigenschaften. Da die umgekehrte RNA-Molekülstruktur von mikrobiellen Zellen das Zellwachstum oder die Regeneration verhindert, Die Erholungszeit von SiC-Schottky-Dioden ist extrem kurz, die Energie geht bei der Rückwärtsbewegung verloren und erreicht letztendlich den Effekt der Desinfektion und Sterilisation. Deep UV-LEDs sind Die Wiederherstellung kann erheblich reduziert werden, was für Anwendungen mit hohem Schaltbedarf geeignet ist, die sich durch Umweltschutz und Energieeinsparung, lange Lebensdauer, keine Aufwärmzeit, Frequenz und fördert die Miniaturisierung der gesamten Schaltung sowie die Un- kein Quecksilber und die Miniaturisierung der Lichtquelle. Herkömmliche UV-Quecksilberlampen haben Verbesserung der Leistungsdichte [21,22]. Ergänzende Hinweise zur Langzeitzuverlässigkeit und wurden nach und nach durch tiefe UV-LEDs ersetzt, die sich als neue Lichtquelle für Mögliche Degradation von Siliziumkarbid-Schottky-Dioden unter thermischen und elektrischen Raumreinigungs- und Oberflächensterilisationsmärkten wie Häusern, Schulen, Krankenhäusern, Büro Belastungen sind unten angegeben. Laut der Literatur [21] sind Siliziumkarbid-Schottkydioden Gebäude, Supermärkte oder Kaufhäuser. Darüber hinaus ist eine tiefe Ultraviolett-Desinfektion haben keine Minoritätsträgerinjektion und ihre Erholungszeit erhöht sich nicht mit der Temperatur. Sterilisationslampen können Bakterien leicht abtöten, aber aufgrund der geringen Durchdringungskraft von tiefen Temperatur im Vergleich zu herkömmlichen Siliziumdioden. Darüber hinaus sind SiC-Schottky-Dioden ultraviolette Desinfektions- und Sterilisationslampen, sie können durch Staubpartikel beeinträchtigt werden; wenn die Mehrheit Trägergeräte und daher ist der Rückstrom kleiner und Temperatur in- Zellen durch tiefe ultraviolette Desinfektion und Sterilisation Lampen beschädigt sind nicht vollständig abhängig. Darüber hinaus haben SiC-Schottkydioden im Vergleich zu Siliziumdioden die Fähigkeit, die Zellen zu reparieren und zu revitalisieren. Daher ist es wichtig, Vorteil der schnellen Wiederherstellung, die Schaltverluste und Rauschen reduziert, und niedrigere Vorwärts bieten den entsprechenden Abstand und ausreichende Bestrahlungszeit und Intensität der tiefen Spannung, die Leitungsverluste reduziert. Darüber hinaus wird in der Literatur [22] die UV-Desinfektions- und Sterilisationslampe für die desinfizierten Objekte erwähnt, um

Ausdau. Eine gute Sterilisationswirkung [3–9].

Langzeitzuverlässigkeit von SiC-Schottkydioden hinsichtlich dV/dt-Ausfall und dI/dt-

Darüberhinaus tritt bei Siliziumdioden bei großem dV/dt ein Muster von erhöhtem Erholungsstrom auf Abbildung 1 zeigt die zweistufige Schaltungstopologie eines vorhandenen elektronischen Beleuchtun (Irr) Leitung, was das Gerät durch Stromkonzentration beschädigen kann. Im Gegensatz dazu werden in Versorgung tief-ultravioletter LED-Sterilisations- und Desinfektionslampen verwendet. Die erste Stufe ist ein AC-DC-Wandler mit Leistungsfaktorkorrektur (PFC), und die zweite Stufe ist ein DC-DC-Wandler und liefert die erforderliche Gleichstromleistung an die Tief-UV-LED-Sterilisation einer tiefen Lampe [8] entwickelt. Probleme zu lösen, wurde ein Beleuchtungselektronischer Beleuchtungstreiber mit Einzelschalter Ultraviolett LED-Desinfektions- und Sterilisationslampe [8] entwickelt. In dieser Arbeit wird ein Steuerungsalgorithmus für die Verwendung primäre Einzelschaltung und Komponenten. Um die Anzahl der Leistungsschalter und Schaltungskomponenten zu reduzieren Spezifische Eingangsbedingungen, Zeitgeber in Version der Belastung, Der Treiber recycelt die in der Schaltkreis des Transformators gespeicherte Energie und nutzt sie effizient. eine Siliziumkarbid-Schottkydiode mit großem Bandabstand als Ausgangsdiode zur Verbesserung der Schaltung Effizienz.

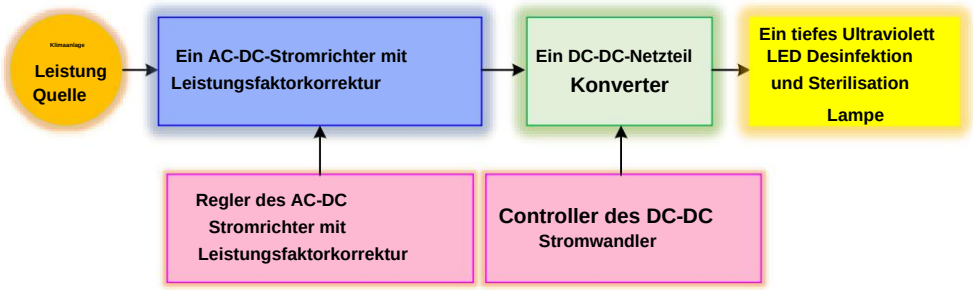


Abbildung 1. Blockdiagramm des vorhandenen zweistufigen elektronischen Beleuchtungstreibers zur Versorgung einer tief-ultravioletten LED-Desinfektions- und Sterilisationslampe [8].

In der Literatur [9–20] wurden eine Reihe von einstufigen LED-Treibern mit Einzelschalter entwickelt.

die einen AC-DC-Stromrichter und einen DC-DC-Stromrichter integrieren, die geeignet sind für die Stromversorgung von LED-Sterilisations- und Desinfektionslampen im tiefen Ultraviolettbereich.

Da die Umweltprobleme, die durch die globale Erwärmung und den Kohlenstoffausstoß entstehen, Dieses Dokument ist wie folgt organisiert und eingeführt. Abschnitt 2 beschreibt und analysiert, wie in den letzten Jahren immer ernster geworden ist, die Menschen haben Energieeinsparungen vorgenommen die Funktionsweise einer einstufigen elektronischen Beleuchtungstreiberschaltung unter Verwendung einer SiC- Kohlenstoffreduzierung und die Liebe zur Erde als gemeinsame primäre Entwicklungsrichtung Schottky-Diode zur Versorgung einer LED-Desinfektions- und Sterilisationslampe mit tiefem Ultraviolettlicht. Daher müssen fossile Brennstoffe schrittweise reduziert und schnell in grüne Energie eingeführt werden bietet Designrichtlinien für einige Schaltungsparameter. Abschnitt 3 zeigt die experimentellen und stromsparenden Anwendungen und damit hohe Energieeffizienz und niedrigen Energieverbrauch Ergebnisse eines Prototyps einer elektronischen Beleuchtungstreiberschaltung unter Verwendung einer SiC- Schottky-Diode zum Verbrauch werden auch allmählich als Ziel in Alltagsprodukten übernommen. In der Pariser

In der Vereinbarung der Konferenz der Vereinten Nationen über Klimaänderungen wird erklärt, dass die

Machine Translated by Google

Die globale Erwärmung muss auf 2 °C begrenzt werden. Basierend auf dem aktuellen wirtschaftlichen Entwicklungstrend werden die **CO2- Emissionen selbst bei einer Begrenzung der Erwärmung auf 2** °C im Jahr 2050 immer noch um **21 % steigen und bis zu 50 % mehr Strom erzeugt werden müssen, um verschiedene menschliche Aktivitäten zu bewältigen und zu unterstützen. Daher ist es zunehmend wichtig, bestehende Energiequellen deutlich aufzurüsten und zu verbessern. Der wichtigste Rohstoff für Halbleiter ist die Produktion von Silizium-Wafern der ersten Generation (Si). Die physikalischen Eigenschaften dieses Materials haben jedoch bei bestehenden Si-basierten Produkten ihre Grenzen erreicht und es ist nicht mehr möglich, die Leistung zu erhöhen, den Wärmeverlust zu reduzieren und die Geschwindigkeit zu erhöhen. Infolgedessen besteht die Notwendigkeit, andere Materialien zu entwickeln, die die Effizienz der Elektronenübertragung und den geringen Energieverbrauch besser nutzen können. In diesem Zusammenhang sind die derzeit auf dem Markt heiß diskutierten Halbleiter der dritten Generation mit großem Bandabstand entstanden, nämlich Siliziumkarbid (SiC) und Galliumnitrid (GaN), die sich durch hohe Energieeffizienz und geringen Energieverbrauch auszeichnen. Darüber hinaus können Leistungsgeräte aus Halbleitermaterialien mit großem Bandabstand bei höheren Spannungen betrieben werden, mehr Leistung erzeugen und geringere Energieverluste aufweisen, während sie deutlich kleiner sind als Leistungsgeräte aus Siliziumhalbleitermaterialien. Siliziumkarbid (SiC) ist ein zusammengesetztes Halbleitermaterial, das aus Silizium (Si) und Kohlenstoff (C) besteht. SiC hat einen niedrigeren Widerstand der Driftschicht als Si, erfordert keine Leitungsmodulation, hat einen etwa dreimal größeren Energieabstand als Si und eine etwa zehnmal höhere Isolationsbruchfeldstärke als Si. Im Vergleich zu den Hochspannungs-Si-PN-Dioden und den Hochgeschwindigkeits-Si-Fast-Recovery-Dioden weisen SiC-Schottky-Dioden aus Halbleitermaterialien mit großem Abstand hervorragende Hochgeschwindigkeits- und Hochspannungseigenschaften auf. Da die Sperrverzögerungszeit von SiC-Schottky- Dioden extrem kurz ist, kann der Energieverlust während der Ergänzende Hinweise zur Langzeitzuverlässigkeit und möglichen Degradation von Siliziumkarbid-Schottkydioden unter thermischen und elektrischen Belastungen finden Sie weiter unten. Laut der Literatur [21] haben Siliziumkarbid-Schottkydioden keine Minoritätsträgerinjektion und ihre Erholungszeit erhöht sich im Vergleich zu herkömmlichen Siliziumdioden nicht mit der Temperatur. Darüber hinaus sind SiC-Schottkydioden Majoritätsträgerbauelemente und daher ist der Sperrstrom kleiner und temperaturunabhängig. Darüber hinaus haben SiC-Schottkydioden im Vergleich zu Siliziumdioden den Vorteil einer schnelleren Erholung, wodurch Schaltverluste und Rauschen reduziert werden, sowie einer niedrigeren Durchlassspannung, wodurch Leitungsverluste r Darüber hinaus wird in der Literatur [22] die langfristige Zuverlässigkeit von SiC-Schottkydioden in Bezug auf dV/dt-Ausfälle und dI/dt-Ausfälle erwähnt. Darüber hinaus tritt bei Siliziumdioden bei großem dI/dt ein Muster erhöhten Erholungsstroms (Irr) auf, das das Gerät durch Stromkonzentration beschädigen kann. Da der Erholungsstrom bei SiC-Schottkydioden dagegen sehr gering ist, kann davon ausgegangen werden, dass ein solches**

Um die oben genannten Probleme zu lösen, wird in diesem Artikel ein neuartiger elektronischer Beleuchtungstreiber für eine tief ultraviolette LED-Desinfektions- und Sterilisationslampe vorgeschlagen und entwickelt, der eine Erweiterung und weitere Verbesserung der Literatur darstellt [20]. Die Hauptschaltung kombiniert einen Abwärtswandler und einen Sperrwandler zu einem einstufigen Einzelschalter- Abwärts-Sperrwandler-AC-DC-Leistungswandler mit Eingangsstromformung. Darüber hinaus recycelt der vorgeschlagene Treiber die in der Streuinduktivität des Transformators gespeicherte Energie und nutzt eine Schottky-Diode aus Siliziumkarbid mit großem Bandabstand als Ausgangsdiode, um die Schaltungseffizienz zu verbessern.

Dieses Dokument ist wie folgt aufgebaut und eingeleitet. Abschnitt 2 beschreibt und analysiert die Funktionsweise einer einstufigen elektronischen Beleuchtungstreiberschaltung, die eine SiC-Schottky-Diode zur Versorgung einer tief ultravioletten LED-Desinfektions- und Sterilisationslampe verwendet, und bietet Designrichtlinien für einige Schaltungsparameter. Abschnitt 3 zeigt die experimentellen Ergebnisse einer Prototyp-Elektronikbeleuchtungstreiberschaltung, die eine SiC-Schottky-Diode zur Versorgung einer tief ultravioletten LED-Sterilisations- und Desinfektionslampe verwendet. Abschließend werden in Abschnitt 4 einige Schlussfolgerungen gezogen.

Machine Translated by Google

liefern eine tief ultraviolette LED-Sterilisations- und Desinfektionslampe. Abschließend werden in Abschnitt 4 einige Schlussfolgerungen gegeben.

2. Beschreibung und Analyse der Betriebsmodi im vorgeschlagenen einstufigen

Elektronischer Beleuchtungstreiberschaltkreis mit SiC-Schottky-Dioden zur Versorgung eines elektronischen Beleuchtungstreiberschaltkreis mit SiC-Schottky-Dioden zur Versorgung eines tiefen Tiefe Ultraviolett-LED-Desinfektions- und Sterilisationslampe Ultraviolette LED-Desinfektions- und Sterilisationslampe

Der in diesem Dokument entwickelte elektronische Beleuchtungstreiber mit PFC zur Stromversorgung einer tief ultravioletten LED- Desinfektions- und Sterilisationslampe ist in Abbildung 2 dargestellt, die einen integrierten Chip enthält . Der in diesem Dokument entwickelte elektronische Beleuchtungstreiber mit PFC zur Stromversorgung einer tief ultravioletten LED-Desinfektions- und Sterilisationslampe ist in Abbildung 2 dargestellt, die einen int kombiniert einen Abwärtswandler und einen Sperrwandler zu einem einstufigen AC -DC-Leistungswandler und sion und umfasst eine Filterinduktivität *LF*, einen Filterkondensator *CF*, einen Vollbrückengleichrichter (einschließ einschließlich umfasst eine Filterinduktivität LF, einen Filterkondensator CF, einen Vollbrückengleichrichter D1, D2, D3 und D4), ein Leistungsschalter *SB*, zwei Dioden *DB* und *DF*, ein Transformator *TR* mit einem Magnetisie und CO2, Induktivität LM und eine Streuinduktivität Llk, zwei Ausgangskondensatoren *CO1* und *CO2* und die LED-Desinfektions- und Sterilisationslampe. Durch die Entwicklung des magnetischen Verbindungstransfors Netzinduktivität *LM* in der vorgeschlagenen elektronischen Beleuchtungstreiberschaltung, um in unterbrochener gesteuerten Modus (DCM) und CrM) kann PFC auf natürliche Weise erreicht werden. Darüber hinaus kann die gespeicherte Energie im Leck Energie in der Streuinduktivität *Llk* vom Transformator *TR* in der vorgeschlagenen elektronischen Beleuchtungstreiberschaltkreis Induktivität Die Diode wird durch die Schaltung effektiv durch die Nutzung der SiC-Schottky-Diode ersetzt werden. einer SiC-Schottky-Diode mit großem Bandabstand als Ausgangsdiode.

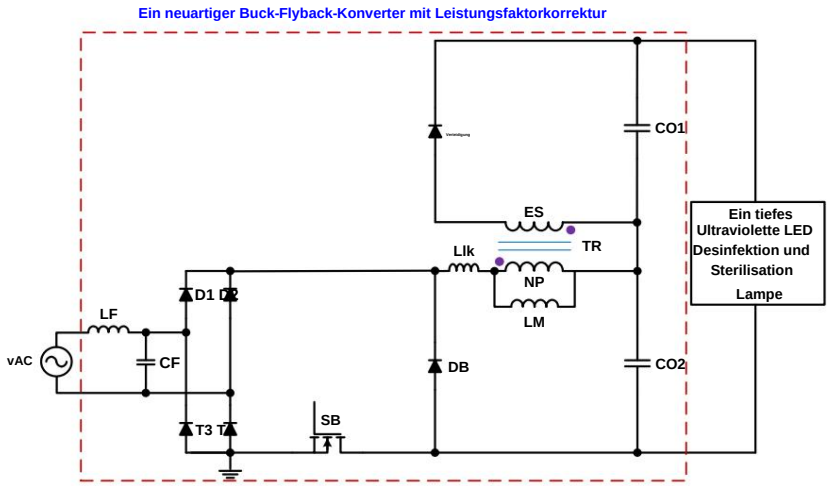


Abbildung 2. Der vorgeschlagene einstufige elektronische Beleuchtungstreiber mit SiC-Schottky-Dioden zur Unterstützung Abbildung 2. Der vorgeschlagene einstufige elektronische Beleuchtungstreiber mit SiC-Schottky-Dioden zur Unterstützung Verwendung einer LED-Desinfektions- und Sterilisationslampe mit tiefem Ultraviolettlicht. Versorgung mit einer LED-Desinfektions- und Sterilisationslampe mit tiefem Ultraviolettlicht.

Abbildung 3 ist das Äquivalentdiagramm der einstufigen elektronischen Beleuchtungstreiberschaltung. Abbildung 3 ist das Äquivalentdiagramm der einstufigen elektronischen Sterilisationslampe, die in diesem Dokument mit der Gleichstromkorrekturfunktion entwickelt wurde. Theoretische Verwendung von SiC-Schottky-Dioden zur Versorgung einer tief ultravioletten LED-Desinfektions- und Verwendung von SiC-Schottky-Dioden zur Versorgung einer tief ultravioletten LED-Desinfektions- und Sterilisationslampe, die in diesem Dokument mit der Gleichstromkorrekturfunktion entwickelt wurde. Theoretische Wellenformen des vorgeschlagenen einstufigen elektronischen Beleuchtungstreibers unter Verwendung von SiC Schottky Sind Dioden zur Versorgung einer tief ultravioletten LED-Desinfektions- und Sterilisationslampe. Dargestellt Siehe in Abbildung 4. Betrieb und Analyse des vorgeschlagenen einstufigen elektronischen Beleuchtungstreibers Abbildung 4 Betrieb und Analyse des vorgeschlagenen einstufigen elektronischen Beleuchtungstreibers zur Desinfektion und Sterilisation mit tiefen Ultraviolettlampen, Annahmen werden während Beschreibungen und Erläuterungen des Betriebsmodus getroffen vorgestellten Beschreibungen und Erläuterungen des Betriebsmodus getroffen werden wie folgt dargestellt: Treiberschaltung werden wie folgt dargestellt:

(a) Die äquivalente Spannungsquelle der Eingangsnetzspannung nach dem Durchlaufen (a) Die äquivalente Spannungsquelle der Eingangsnetzspannung nach dem Durchlaufen des Vollweggleichrichterschaltkreises wird mit VREC bezeichnet. Da die Schaltfrequenz fs viel größer ist als die Da die Gleichspannung VREC viel größer als die Netzfrequenz fAC, kann die gleichgerichtete Spannung VREC betrachtet werden als Netzfrequenz fAC, kann die gleichgerichtete Spannung VREC betrachtet werden als während eines Schaltzyklus als konstanter Wert betrachtet werden. einen konstanten Wert während eines Schaltzyklus in der Schaltungsmodusanalyse.

(b) Die Streuinduktivität Llk des Transformators TR wird berücksichtigt. (b) Die Streuinduktivität *Llk* des Transformators *TR* wird berücksichtigt.

(c) Die Magnetisierungsspule LM des Transformators TR ist für den Betrieb im diskontinuierlichen Leitungsmodus (DCM) ausgelegt. (c) Die Magnetisierungsspule *LM* des Transformators *TR* ist für den Betrieb im diskontinuierlichen Leitungsmodus (DCM) ausgelegt.

(d) Vernachlässigung der Leitungsspannungsabfälle und ihrer äquivalenten Widerstände aller Dioden.

(e) Die restlichen Schaltungskomponenten werden als ideale Komponenten betrachtet.

Machine Translated by Google

(d) Vernachlässigt man die Leitungsspannungsabfälle und ihre äquivalenten Widerstände aller di-
Oden.

(e) Die restlichen Schaltungskomponenten werden als ideale Komponenten betrachtet.

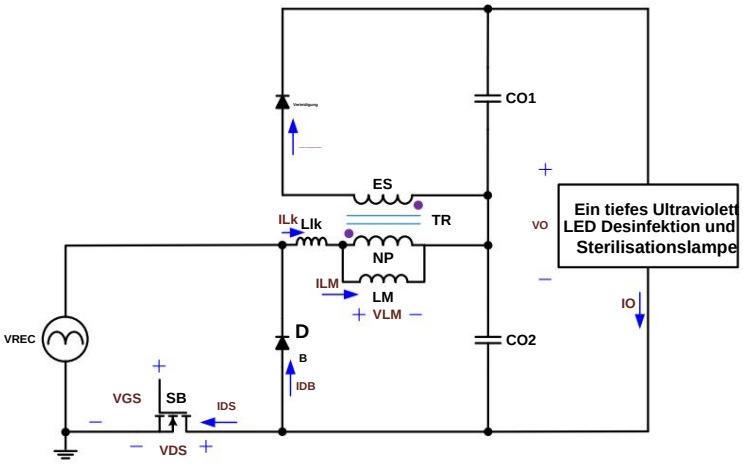


Abbildung 3. Äquivalente Schaltung der vorgeschlagenen elektronischen Beleuchtungstreiberschaltung unter Verwendung von SiC-Schottky-Dioden zur Versorgung einer tief ultravioletten LED-Desinfektions- und Sterilisationslampe im tiefen Ultraviolettbereich.

Betriebsmodi.

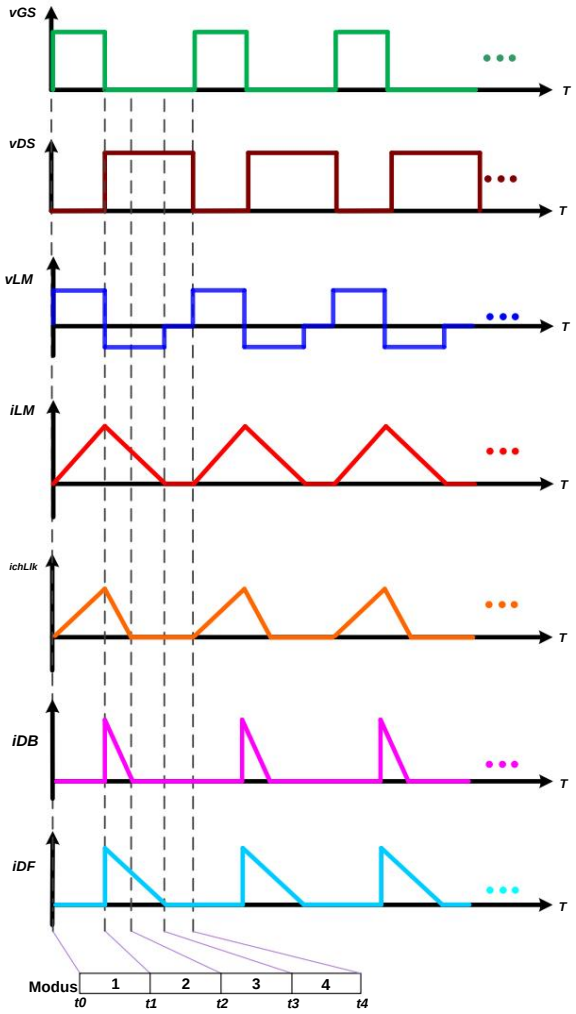


Abbildung 4. Theoretische Wellenformen des vorgeschlagenen einstufigen elektronischen Beleuchtungstreibers unter Verwendung von SiC-Schottky-Dioden zur Versorgung einer LED-Desinfektions- und Sterilisationslampe mit tiefem Ultraviolettlicht.

Betriebsart 1 (t0 ÿ t < t1): Abbildung 5 zeigt das Ersatzschaltbild des vorgeschlagenen einstufiger elektronischer Beleuchtungstreiberschaltkreis zur Versorgung einer tief ultravioletten LED-Desinfektions- und Sterilisationslampe im Modus 1. Der Netzschalter SB wird einges

Machine Translated by Google

Betriebsart 1 (t0 ÿ t < t1): Abbildung 5 zeigt das Ersatzschaltbild des vorgeschlagenen einstufiger elektronischer Beleuchtungstreiberschaltkreis zur Versorgung einer tief ultravioletten LED-Desinfektions- und Sterilisationslampe im Modus 1. Der Netzschalter *SB* wird eingeschaltet und die äquivalente Spannungsquelle VREC **wird an beide Enden der Diode** DB **angeschlossen, wodurch die Diode *DB* ein in Sperrrichtung vorgespannter nichtleitender Zustand. Die äquivalente Spannungsquelle** VREC **versorgt den Ausgangskondensator** CO2 **und die Magnetisierungsspule** LM **des Transformators** TR **mit Energie auf dem Ausgangskondensator CO2 und die Magnetisierungsspule *LM* des Transformators *TR* sowie die Streuinduktivität durch den Schalter *SB*, was zu einem linearen Anstieg des Stroms der Magnetisierungsinduktivität** LM **führt und die Spannungspolarität auf beiden Seiten der Magnetisierungsinduktivität *LM* und Streuinduktivität *Llk* sind links positiv und rechts negativ. *Die* Ausgangskondensatoren CO1 **und** CO2 **versorgen die tief ultraviolette LED- Desinfektions- und Sterilisationslampe mit Energie. Wenn der Schalter *SB* ausgeschaltet wird, wird der Strom der Magnetisierungslampe und Sterilisationslampe. Wenn der Schalter *SB* ausgeschaltet wird, erreicht der Strom der Magnetisierungsspule** LM **seinen Maximalwert und der Modus endet. Der Leiter *LM* erreicht seinen Maximalwert und der Modus wird beendet.**

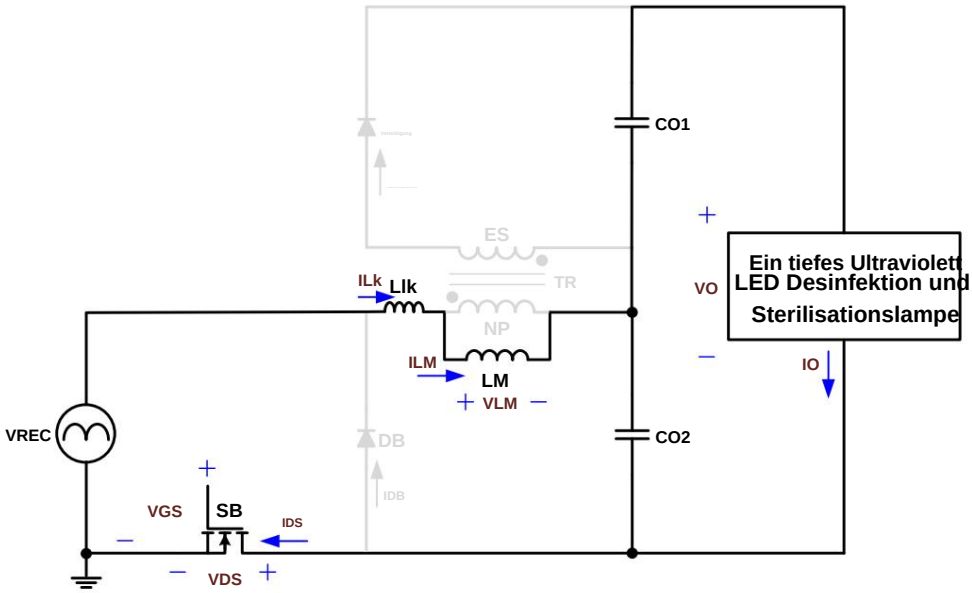


Abbildung 5. Ersatzschaltbild der vorgeschlagenen einstufigen elektronischen Beleuchtungstreiberschaltung unter Verwendung einer Variantenlampe während Sie Schalter Diode Versorgung eine tief ultraviolette LED-Desinfektions- und Sterilisationslampe im Modus 1.

Betriebsart 2 (t1 ÿ t < t2): Abbildung 6 zeigt das Ersatzschaltbild des vorgeschlagenen einstufiger elektronischer Beleuchtungstreiberschaltkreis zur Versorgung einer tief ultravioletten LED-Desinfektions- und Sterilisationslampe im Modus 2. Wenn der Netzschalter *SB* ausgeschaltet wird, wird die Spannung Die Spannungspolarität der Magnetisierungsspule *LM* und der Streuspule *Llk* wird gemäß dem Lenzschen Gesetz auf positiv geändert und negativ gelassen, wodurch die Diode DB **zur Durchlasszustand versetzt wird. Der Strom der Magnetisierungsspule** LM **fließt in die Position, an der sich die primärseitige Wicklung** NP **des Transformators** TR **kein Punkteende befindet, und die Spannungspolarität der beiden Enden der Wicklung** NP **ist rechtspositiv und linksnegativ. Daher ist die Polarität der an beiden Enden des Sekundär-** NS **reflektierten Spannung linkspositiv und rechtsnegativ, was die Diode** DF **in der Vorwärtsrichtung arbeiten lässt, vorwärts vorgespannter Leitungszustand. Zu diesem Zeitpunkt befinden sich die Induktivität *Llk* befindet sich im Zustand der Energieabgabe, wodurch der Strom der Magnetisierung Induktor *LM* linear abnimmt. Da die Diode *DB* in der Vorwärtsleitung des Magnetisierungsinduktors** LM **ist und der Streuinduktor** Llk **dem Ausgang zugeführt wird, Zustand wird die Energie der Magnetisierungsinduktivität *LM* und der Streuinduktivität *Llk* dem Kondensator** CO2 **über die Diode** DB **zugeführt, und der Strom der Magnetisierungsinduktivität** LM **zeigt zum Ausgangskondensator *CO2* über die Diode *DB,* und der Strom des Magnetisierungseingangs- eine lineare Abnahme. Aufgrund der vorwärts vorgespannten Leitung der Diode** DF **wird die Energie von der Diode *DF* wird der Magnetisierungsinduktor** LM **über den Transformator** TR **und die Diode** DF **versorgt, Die Energie der Magnetisierungsspule *LM* wird über den Transformator *TR* und die Ausgangskapazitäten** CO1 **und** CO2 **sowie zur tiefen ultravioletten LED-Desinfektion geliefert. Diode *DF* zu den Ausgangskondensatoren *CO1* und *CO2* sowie zur tief ultravioletten LED** und Sterilisationslampe. Darüber hinaus **liefern die Ausgangskondensatoren** CO1 **und** CO2 weiterhin **Desinfektions- und Sterilisationslampe. Darüber hinaus liefern die Ausgangskondensatoren *CO1* und *CO2* weiterhin Energie an die tief ultraviolette LED-Desinfektions- und Sterilisationslampe. Der Modus endet wenn die Streuinduktivität** Llk **die Energieabgabe beendet hat und der Strom** ILlk **der Streuinduktivität Induktor** Llk **fällt linear auf Null ab.**

Machine Translated by Google

um die tief-ultraviolette LED-Desinfektions- und Sterilisationslampe mit Energie zu versorgen. Der Modus endet, wenn die Streuinduktivität *Llk* keine *Energie* mehr freigibt und der Strom *ILlk der* Streuinduktivität *Llk* linear auf Null abfällt.

die Streuinduktivität *Llk* sinkt linear auf Null.

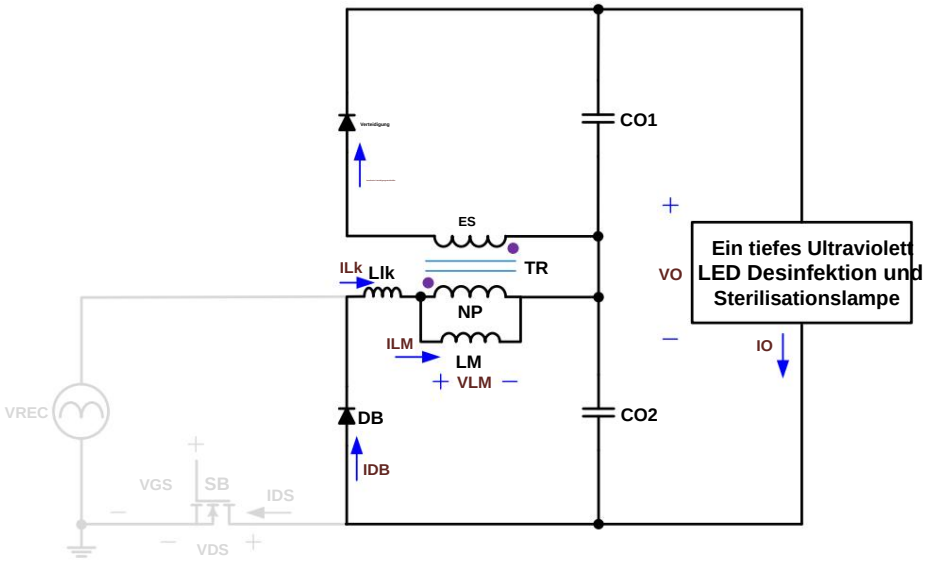


Abbildung 6. Ersatzschaltbild der vorgeschlagenen einstufigen elektronischen Beleuchtungstreiberschaltung unter Verwendung von SiC Schottky-Dioden zur Versorgung einer tief-ultravioletten LED-Desinfektions- und Sterilisationslampe während Modus 2.

Betriebsart 3 (t2 ÿ t < t3): Abbildung 7 zeigt das Ersatzschaltbild des vorgeschlagenen einstufigen elektronischen Beleuchtungstreiberschaltung zur Versorgung einer tief-ultravioletten LED-Desinfektions- und Sterilisationslampe im Modus 3. Die Energie des Magnetisierungsinduktors *LM* wird kontinuierlich über den Transformator *TR* und die Diode *DF* an die Ausgangskondensatoren CO1 und CO2 und zur tiefen Ultraviolett-LED-Desinfektions- und Sterilisationslampe. Darüber hinaus liefern die Ausgangskondensatoren CO1 und CO2 weiterhin Energie an die Tiefentladung. Ultraviolett-LED-Desinfektions- und Sterilisationslampe. Dieser Modus endet, wenn der Magnetisierungsinduktor LM keine Energie mehr freigibt und der Magnetisierungsinduktivitätsstrom *ILM* sinkt linear auf Null.

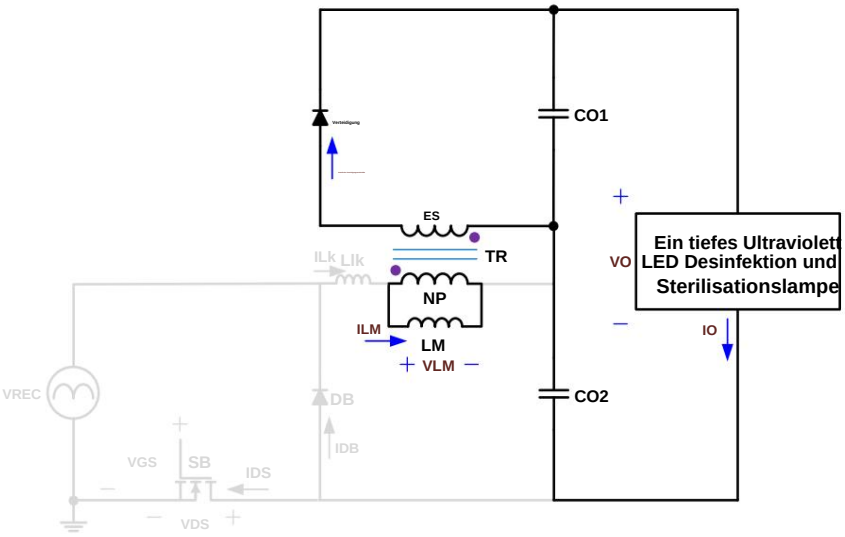


Abbildung 7. Ersatzschaltbild der vorgeschlagenen einstufigen elektronischen Beleuchtungstreiberschaltung unter Verwendung von SiC Schottky-Dioden zur Versorgung einer tief-ultravioletten LED-Desinfektions- und Sterilisationslampe während Modus 3.

Betriebsmodus 4 (t3 ÿ t < t4): Abbildung 8 zeigt das Ersatzschaltbild der vorgeschlagenen einstufigen elektronischen Beleuchtungstreiberschaltung zur Versorgung einer tief-ultravioletten LED Desinfektions- und Sterilisationslampe im Modus 4. Die Ausgangskondensatoren CO1 **und** CO2

Machine Translated by Google

Betriebsart 4 (t3 ÿ t < t4): Abbildung 8 zeigt das Ersatzschaltbild des vorgeschlagenen einstufiger elektronischer Beleuchtungstreiberkreis zur Versorgung einer tief ultravioletten LED-Desinfektions- und Sterilisationslampe im Modus 4. Die Ausgangskondensatoren *CO1* und *CO2* versorgen die tief ultraviolette LED-Desinfektions- und Sterilisationslampe kontinuierlich mit Energie.

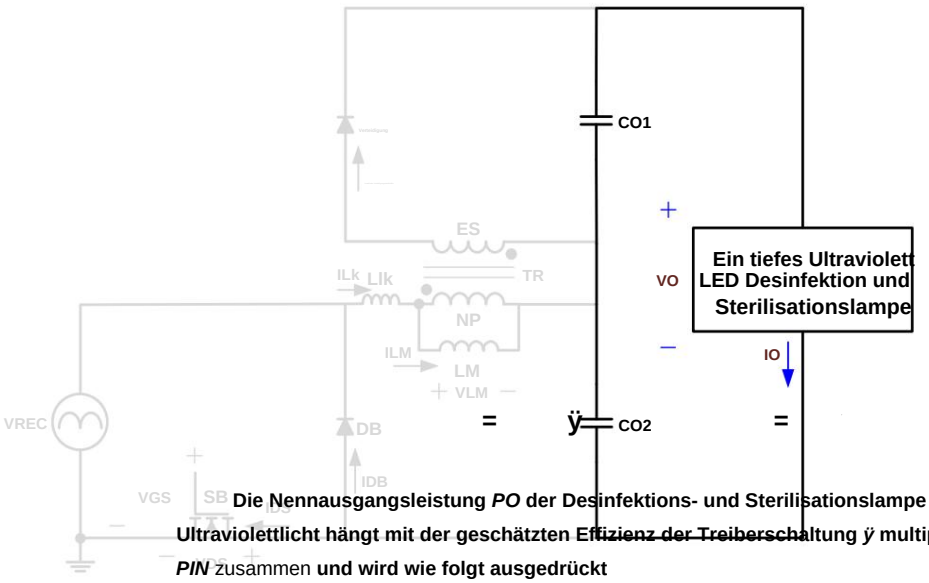
versorgen die tief ultraviolette LED-Desinfektions- und Sterilisationslampe mit Energie. Lampe.

Wenn der Netzschalter SB **wieder eingeschaltet wird, endet dieser Modus und der Schaltkreis kehrt zur**

Beim **erneuten Einschalten des Netzschalters *SB*** wird dieser Modus beendet und die Schaltung kehrt **in den ersten**

Betriebsmodus zurück.

der erste Betriebsmodus.



Die Nennausgangsleistung *PO* der Desinfektions- und Sterilisationslampe mit tiefem Ultraviolettlicht hängt mit der geschätzten Effizienz der Treiberschaltung *ÿ* multipliziert mit der Leistung *PIN* zusammen und wird wie folgt ausgedrückt

Abbildung 8. Ersatzschaltbild der vorgeschlagenen einstufigen elektronischen Beleuchtungstreiberschaltung unter Verwendung von SiC-Schottky-Dioden zur Versorgung einer tief ultravioletten LED-Desinfektions- und Sterilisationslampe im Modus 4.

Abbildung 8. Äquivalente Schaltung der vorgeschlagenen einstufigen elektronischen Beleuchtungstreiberschaltung **SiC-Schottky-Dioden zur Versorgung einer tief ultravioletten LED-Desinfektions- und Sterilisationslampe im Modus 4.**

Umstellung (5) ergibt die Entwurfsformel für die Induktivität der Magnetisierungsspule i 2.1. Entwurfsrichtlinie für die Magnetisierungsspule LM im Transformator TR

2.1. Entwurfsrichtlinie für die Magnetisierungsspule LM im Transformator TR

Um die Konstruktionsrichtlinie für die Magnetisierungsinduktivität LM im Transformator TR zu veranschaulichen , zeigt Abbildung 9 theoretische

Abbildung 9 zeigt ***die* theoretischen Wellenformen des Magnetisierungsinduktivitätsstroms .**

Magnetisierungsinduktorstroms iLM(t), **den Spitzenwert von** iLM-pk(t) **und den Eingangs**

***iLM(t)*, der Spitzenpegel von *iLM-pk(t)* und der Eingangsnetzstrom *iAC(t)* in der positiven Halbwelle der**

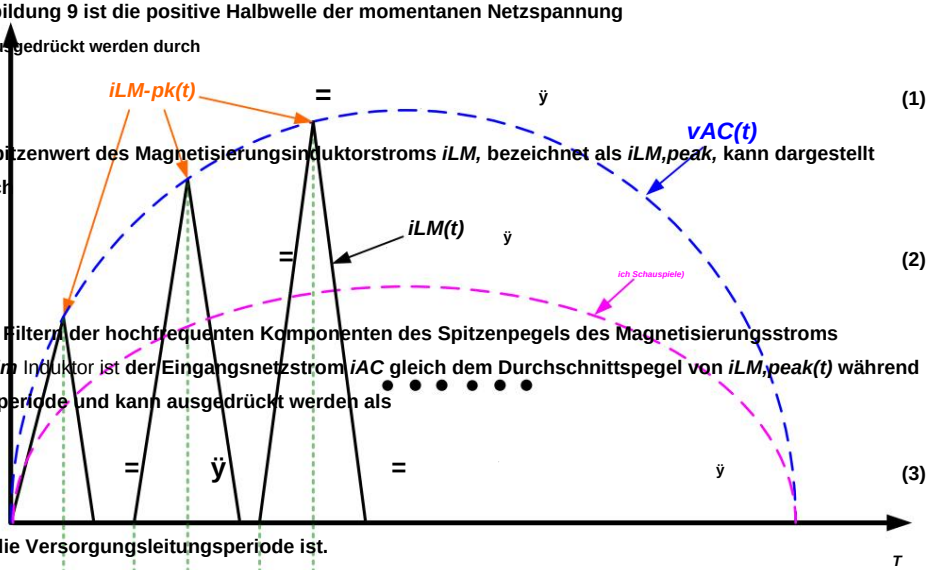
Netzspannung vAC(t)

der Netzspannung *vAC(t)*.

In Abbildung 9 ist die positive Halbwelle der momentanen Netzspannung

Aktuell

***vAC(t)* kann ausgedrückt werden durch**



(1)

Der Spitzenwert des Magnetisierungsinduktorstroms *iLM,* bezeichnet als *iLM,peak,* kann dargestellt werden durch

(2)

Durch Filtern der hochfrequenten Komponenten des Spitzenpegels des Magnetisierungsstroms *iLM,peak(t) im* Induktor ist der Eingangsnetzstrom *iAC* gleich dem Durchschnittspegel von *iLM,peak(t)* während einer Schaltperiode und kann ausgedrückt werden als

(3)

wobei *TAC* die Versorgungsleitungsperiode ist.

Der Durchschnittswert der Eingangsleistung *PIN* ergibt sich aus der Multiplikation des Momentanwertes der Netzspannung *vAC(t)* mit dem Momentanwert des Netzstroms *iAC(t)* und der anschließenden Mittelung über einen Zykluszeitraum. Dieser berechnet sich wie folgt:

vGS

Abbildung 9. Illustrative Wellenformen des Magnetisierungsinduktorstroms *iLM(t)*, des Spitzenpegels *iLM-pk(t)* und dem Eingangsnetzstrom *iAC(t)* in der positiven Halbwelle der Netzspannung *vAC*.

Abbildung 9. Illustrative Wellenformen des Magnetisierungsinduktorstroms iLM(t), der Spitzenpegel von iLM-pk(t) und des Eingangsnetzstroms *iAC(t)* in der positiven Halbwelle der Netzspannung vAC(t).

2.2. Entwurfsrichtlinie für die Ausgangskondensatoren CO1 und CO2

Die Auslegung der Ausgangskondensatoren *CO1* und CO2 hängt von der Ausgangsspannung, der zulässigen Überspannung, der Ausgangsleistung und der gewünschten Spannung ab. Die Spannungswelligkeit *ÿVO* beträgt die Hälfte des Spitzen-Spitzen-Werts der Ausgangsspannung bei der Netzfrequenz und ist eine Funktion der Kapazität *CO* und des Spitzenkondensators.

Machine Translated by Google

In Abbildung 9 **kann die positive Halbwelle der momentanen Netzspannung** vAC(t) wie folgt ausgedrückt werden:

$$|vAC(t)| = \ddot{y}\ 2vAC\ddot{y}rms|sin(2\ddot{y}\ fACt)| \quad (1)$$

Der Spitzenwert des Magnetisierungsinduktorstroms iLM, **bezeichnet als** iLM,peak, **kann dargestellt werden durch**

$$iLM,peak(t) = \frac{\ddot{y}\ 2vAC\ddot{y}rms|sin(2\ddot{y}\ fACt)|DutyTS}{2LM} \quad (2)$$

Durch Filtern der hochfrequenten Komponenten des Spitzenpegels des Magnetisierungsinduktorstroms iLM ,peak(t) ist **der Eingangsnetzstrom** iAC **gleich dem Durchschnittspegel von** iLM,peak(t) **während einer Schaltperiode und kann ausgedrückt werden als**

$$iAC(t) = \frac{1}{TAC}\int_0^{T_{Einsatz}} iLM,peak(t)dt = \frac{\ddot{y}\ 2vAC\ddot{y}rmsDuty2TS}{4LM}\sin(2\ddot{y}\ fACt) \quad (3)$$

wobei TAC **die Versorgungsleitungsperiode ist.**

Der Durchschnittswert der Eingangsleistung PIN **ergibt sich aus der Multiplikation des Momentanwertes der Netzspannung** vAC(t) **mit dem Momentanwert des Netzstroms iAC** (t) **und der anschließenden Mittelung über einen Zykluszeitraum. Dieser berechnet sich wie folgt:**

$$PIN = \frac{1}{TAC}\int_0^{T_{Einsatz}} vAC(t)iAC(t)dt = \frac{\mathbf{vAC\ddot{y}rms}\ 2Duty2TS}{4LM} \quad (4)$$

Die Nennausgangsleistung PO **der LED-Desinfektions- und Sterilisationslampe mit tiefem Ultraviolett hängt mit der geschätzten Effizienz der Treiberschaltung** *ÿ* **multipliziert mit der Eingangsleistung PIN** zusammen **und wird wie folgt ausgedrückt**

$$PO = \ddot{y}PIN = \frac{\ddot{y}\mathbf{vAC\ddot{y}rms}\ 2Betriebsspannung2TS}{4LM} \quad (5)$$

Umstellung (5) der Entwurfsformel für die Induktivität der Magnetisierungsspule

LM **ist gegeben durch**

$$LM = \frac{\ddot{y}\mathbf{vAC\ddot{y}rms}\ 2Betriebsspannung2TS}{4PO} \quad (6)$$

2.2. Entwurfsrichtlinie für die Ausgangskondensatoren CO1 und CO2

Die Auslegung der Ausgangskondensatoren CO1 **und CO2 hängt von der DC-Ausgangsspannung, der zulässigen Überspannung, der Ausgangsleistung und der gewünschten Spannungswelligkeit** ab . **Die Spannungswelligkeit** ÿVO **ist der halbe Spitze-Spitze-Wert der Ausgangsspannung bei doppelter Netzfrequenz und ist eine Funktion der Kapazität** CO **und des Spitzenkondensatorstroms (der gleich dem Ausgangsstrom** IO ist) **und kann durch die folgende Gleichung** ausgedrückt werden **[23].**

$$\ddot{y}VO = IO\sqrt{\left(\frac{1}{2\ddot{y} \times \mathbf{2}\ fAC \times CO}\right)^2 + RESR2} \quad (7)$$

Unter Vernachlässigung des äquivalenten Serienwiderstandes, der als RESR bezeichnet wird, **des Ausgangs Kondensator, die Kapazität** CO **des Ausgangskondensators kann ausgedrückt werden als**

$$CO\ \ddot{y}\ \frac{IO}{4\ddot{y}\ fAC \times \ddot{y}VO} = \frac{}{PO\ 4\ddot{y}\ fAC \times VO \times \ddot{y}VO} \quad (8)$$

Darüber hinaus wird die Spannungswelligkeit ÿVO **typischerweise zwischen 1 % und 5 % der Ausgangsspannung gewählt.**

Machine Translated by Google

3. Experimentelle Ergebnisse der vorgeschlagenen elektronischen Beleuchtungstreiberschaltung unter Verwendung von SiC Schottky-Didos zur Versorgung einer tiefen Ultraviolett-LED-Desinfektion und Sterilisationslampe

Abbildung 10 zeigt ein Foto der tiefen ultravioletten LED-Desinfektion und Sterilisation

Das in diesem Dokument verwendete Modul mit Spezifikationen wie einer Wellenlänge von 275 nm und einer Punktwinkel von 120 Grad. Darüber hinaus hat der Hersteller einen Konstantstrom angeschlossen Steuer-IC in Reihe und eine Schottky-Diode parallel zum Verpolungsschutz auf der Aluminiumsubstrat in jedem Tief-Ultraviolett-LED-Desinfektions- und Sterilisationsmodul.
Die Spezifikation der verwendeten tiefen ultravioletten LED-Desinfektions- und Sterilisationslampe in diesem Papier besteht aus zehn tiefen ultravioletten LED-Desinfektions- und Sterilisationsmodulen in Reihe geschaltet mit einer Nennspannung von 90 Volt, einem Nennstrom von 40 Milliampere und eine Nennleistung von 3,6 Watt. Tabelle 1 zeigt einen Vergleich zwischen dem tiefen Ultraviolett LED-Desinfektions- und Sterilisationsmodul in [24] und das in diesem Papier verwendete Modul.
Aus Tabelle 1 ist ersichtlich , dass das Desinfektions- und Sterilisationsmodul mit tief ultravioletter LED Die in diesem Dokument verwendete eine etwas größere optische Ausgangsleistung und spektrale Halbwertsbreite als das Modul in [24].

Tabelle 2 zeigt die elektrischen Parameter des einstufigen elektronischen Beleuchtungsprototyps Treiberschaltung für die tief ultraviolette LED-Desinfektions- und Sterilisationslampe entwickelt in dieses Papier. Die RMS AC-Eingangsspannung betrug 110 Volt, und die Spezifikationen des tiefen Ultra

Als Ausgabeparameter wurden LED-Desinfektions- und Sterilisationslampen verwendet, darunter: Ausgangsleistung von 3,6 Watt, Ausgangsspannung von 90 Volt und Ausgangsstrom von 40 Milliampere.

Darüber hinaus sind Sicherheitsbedenken und die Sicherstellung der Einhaltung gesetzlicher Normen für UV Zur Exposition und elektrischen Sicherheit gehören: Schauen Sie nicht direkt in die UV-Lichtquelle, wenn Verwendung einer tiefen ultravioletten Desinfektions- und Sterilisationslampe, da tiefe UV-Strahlung verursachen kann Hautrötung, Bindehautreizung und Müdigkeit. Darüber hinaus sollten Sie tiefes UV nicht überbeanspruchen Desinfektions- und Sterilisationslampen, da längerer Gebrauch oder Kontakt mit UV-Lampen oder direkter Die Einwirkung von UV-Strahlung kann zu Haut- und Augenschäden führen.

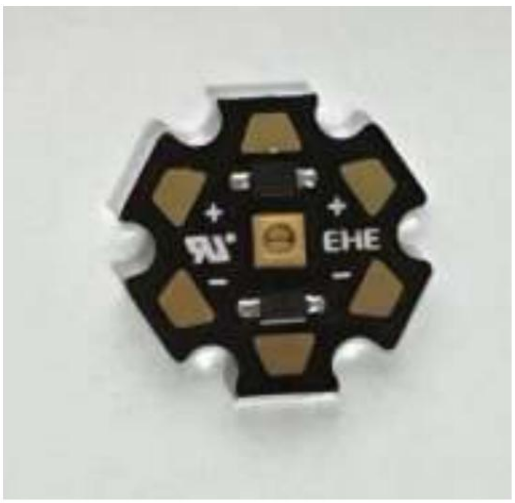


Abbildung 10. Foto des in diesem Produkt verwendeten Desinfektions- und Sterilisationsmoduls mit tief ultravioletter LED-Strahlung Abbildung 10. Foto des in diesem Produkt verwendeten Desinfektions- und Sterilisationsmoduls mit tief ultravioletter LED-S Papier mit Spezifikationen, die eine Wellenlänge von 275 nm und einen Spitzenwinkel von 120 Grad umfassen. Papier mit Spezifikationen, die eine Wellenlänge von 275 nm und einen Spitzenwinkel von 120 Grad umfassen.

3.1. Berechnung der Magnetisierungsspule LM im Transformator TR

Unter Bezugnahme auf (6) mit einem ÿ von 0,8, einem Duty von 0,25, einem PO von 3,6 W, einer Schaltfrequenz fS von 50 kHz (das ist der Kehrwert der Schaltperiode TS) und einem VAC rms von 110 V, der Die Induktivität der Magnetisierungsspule LM berechnet sich aus

$$LM = \frac{\ddot{y}\,\text{ÿvACÿrms}^{2}\,\text{Betriebsspannung2TS}}{4PO} \equiv \frac{\ddot{y}\,0{,}8 \times 1102 \ddot{y}\,0{,}252 \times \ddot{y}\ddot{y}\,\frac{1}{50.000}\,\ddot{y}}{4 \times 3{,}6\,\ddot{y}} \equiv \ddot{y}\,834\,\mu H\,\ddot{y} \quad (9)$$

Um den Magnetisierungsinduktorstrom in einem diskontinuierlichen Leitungsmodus zu betreiben und so der Treiberschaltung eine Eingangsstromformungsfunktion zu ermöglichen und den Leistungsfaktor zu verbessern, wird für die Prototyp-Treiberschaltung der Magnetisierungsinduktor ***LM*** mit 800 ÿH gewählt.

Machine Translated by Google

Tabelle 1. Vergleiche zwischen dem Tief-Ultraviolett-LED-Desinfektions- und Sterilisationsmodul in [24] und die in diesem Dokument verwendete.

Artikel	Die tiefe Ultraviolett-LED-Desinfektion und Sterilisationsmodul in [24]	Die tiefe Ultraviolett-LED-Desinfektion und Sterilisationsmodul Verwendet in dieses Papier
Typischer Wert der Peakwellenlänge	255 nm	275 nm
Typischer Wert des optischen	3,5 mW	4 mW
Ausgangsleistungsbetrachtungswinkels	125 Grad 11	120 Grad
Halbe Spektrumbreite	nm	12 nm

Tabelle 2. Spezifikationen der vorgeschlagenen elektronischen Beleuchtungstreiberschaltung unter Verwendung von SiC-Schottky-Dioden zur Versorgung einer tief-ultravioletten LED-Desinfektions- und Sterilisationslampe.

Parameter	Wert
Eingangswechselspannung vAC	110 V
Nennausgangsleistung PO	3,6 Watt
Nennausgangsspannung VO	90 V
Nennausgangsstrom IO	40 mA

Um den Magnetisierungsinduktorstrom in einer diskontinuierlichen Leitung zu betreiben Modus, um der Treiberschaltung eine Eingangsstromformungsfunktion zu ermöglichen und Um den Leistungsfaktor zu verbessern, wird die Magnetisierungsspule LM mit 800 **µH für die Prototyp einer Treiberschaltung.**

3.2. Berechnung der Ausgangskondensatoren CO1 und CO2

Wenn wir die Schaltungsparameter in Gleichung (8) einsetzen, erhalten wir bei einer PO **von 3,6 W eine** fAC **von 60 Hz, eine** VO **von 90 V, eine** ÿVO **von 2,7 V (das sind 3 % der Ausgangsspannung VO), die Kapazität** CO **kann gewonnen werden durch:**

$$CO\ \ddot{y} \quad \frac{\text{Bestellung}}{4\ddot{y}\ fAC \times VO \times \ddot{y}VO} = \frac{3.6}{\times 60 \times 90 \times 2,7} = 196,49\ \mu F\ 4\ddot{y} \qquad (10)$$

Bei der Implementierung der Prototypschaltung werden die Kapazitätswerte der Ausgangskondensatoren Für CO1 **und** CO2 wurden 220 **µF und eine Spannungsfestigkeit von 250 V** gewählt.

Tabelle 3 zeigt die Komponentenspezifikationen der einstufigen elektronischen Beleuchtung Prototyp einer Treiberschaltung für die tief ultraviolette LED-Desinfektions- und Sterilisationslampe in diesem Papier entwickelt. Darüber hinaus Mehrheitsträgerdioden mit Schottky-Technologie auf Siliziumkarbid-Material mit großem Bandabstand hat eine höhere Leistung und wird als Ausgangsmaterial verwendet Dioden DB **und** DF. **Durch Variation der Anzahl der in Reihe geschalteten tiefen Ultraviolett-Desinfektions und Sterilisationsmodule verwendet, ein skalierbares Design des vorgeschlagenen elektronischen Beleuchtungstreibers Schaltungen können für Anwendungen mit höherer Leistung und variierender Last realisiert werden Bedingungen. Dadurch können** die Ausgangsleistung PO **und die Ausgangsspannung** VO variiert werden, während **Der Ausgangsstrom** IO **bleibt konstant und kann in einer Vielzahl praktischer Anwendungen eingesetzt werden Szenarien. Wichtige Schaltungsparameter der vorgeschlagenen Antriebsschaltung, wie die Magnetisierungs-Induktor und Ausgangskondensator können dann neu gestaltet werden.**

Abbildung 11 zeigt die gemessene Wellenform des Magnetisierungsinduktorstroms iLM. **Abbildung 12 zeigt die gemessene entfaltete Wellenform des Magnetisierungsinduktorstroms** iLM. **Aus Abbildung 12 ist ersichtlich, dass der Magnetisierungsinduktorstrom** iLM **in der diskontinuierliche Leitungsmodus. Abbildung 13 zeigt die gemessenen Wellenformen der Ausgangs Spannung** VO **und Ausgangsstrom** IO **sowie die Durchschnittswerte von Ausgangsspannung und -strom waren 91,35 Volt bzw. 38,57 Milliampere. Abbildung 14 zeigt die Wellenformen von Ausgangsspannung** VO **und Ausgangsstrom** IO.

Machine Translated by Google

Abbildung 11 zeigt die gemessene Wellenform des Magnetisierungsinduktorstroms *iLM.*

Abbildung 12 zeigt die gemessene entfaltete Wellenform des Magnetisierungsinduktorstroms *iLM*. Aus Abbildung 12 ist ersichtlich, dass der magnetisierende Induktorstrom *iLM* in der diskontinuierliche Leitungsmodus. Abbildung 13 zeigt die gemessenen Wellenformen der Ausgangskomponente

Spannung *VO* und Ausgangsstrom *IO*, und die Durchschnittswerte von Ausgangsspannung und

Ausgangsstrom betrugen 91,35 Volt bzw. 38,57 Milliampere. Abbildung 14 zeigt die

Welligkeitswellenformen Filterinduktor Lf **Filterkondensator** Cf **der Ausgangsspannung *VO* und des Ausgangsst**

Tabelle 4 zeigt den Welligkeitsfaktor der Ausgangsspannung und den Welligkeitsfaktor des Ausgangsstroms

Leistungsschalters SB, Prototyp einer elektronischen Beleuchtungstreiberschaltung für eine LED-Desinfektion im tiefen Ultraviolettbereich und die Magnetisierungsspule LM **800 µH**

Sterilisationslampe bei 110 Volt RMS AC Eingangsspannung. Wie aus Tabelle 4 ersichtlich, dividiert Streuinduktor Llk

Die Spitzen-Spitzen-Ausgangsspannung von 143,8 Millivolt wird durch den Durchschnittswert von 91,35 Volt der

bestimmt und ergibt einen Spannungswelligkeitsfaktor von 0,16 %, dividiert man den Spitzen-Spitzen-Ausgangsstrom

Milliampere mal dem Durchschnittswert von 38,57 Milliampere ergibt einen Stromwelligkeitsfaktor von 3,02 %.

		800 ÿH
		12 ÿH
		220 ÿF/200 V

Komponenten, die in der vorgeschlagenen elektronischen Beleuchtungstreiberschaltung unter Verwendung von SiC Schottky verw

Versorgung einer tief ultravioletten LED-Desinfektions- und Sterilisationslampe arbeitet.

Tabelle 3. Wichtige

	Wert
	1 mH
	220 nF
Dioden D1, D2, D3, D4	**MUR460**
des	**IRF730**
Dioden DB1,DB2	**VS-SCHOTTKY-MS**
	220 µF/200 V

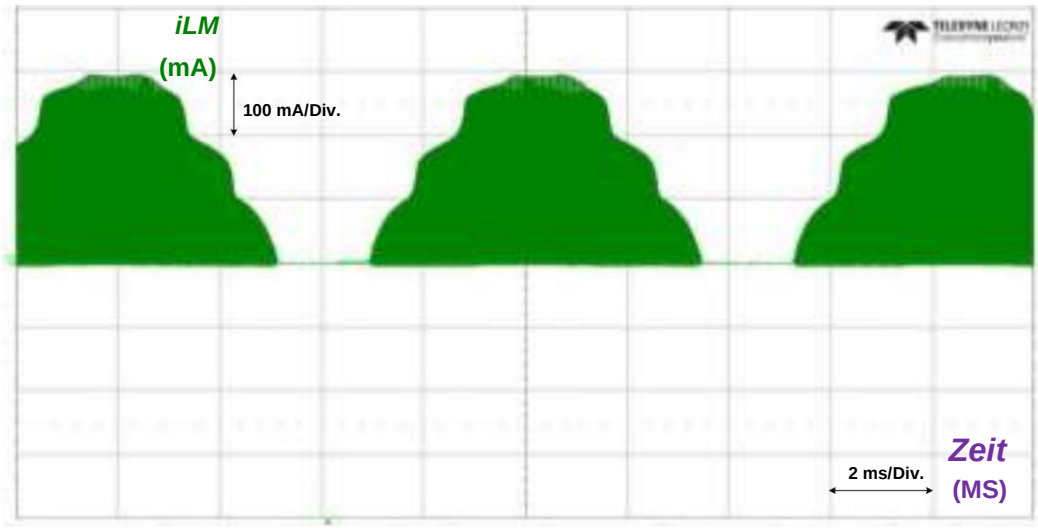


Abbildung 11. Gemessener Magnetisierungsinduktorstrom *iLM* (100 mA/div.), Zeitskala: 2 ms/div.

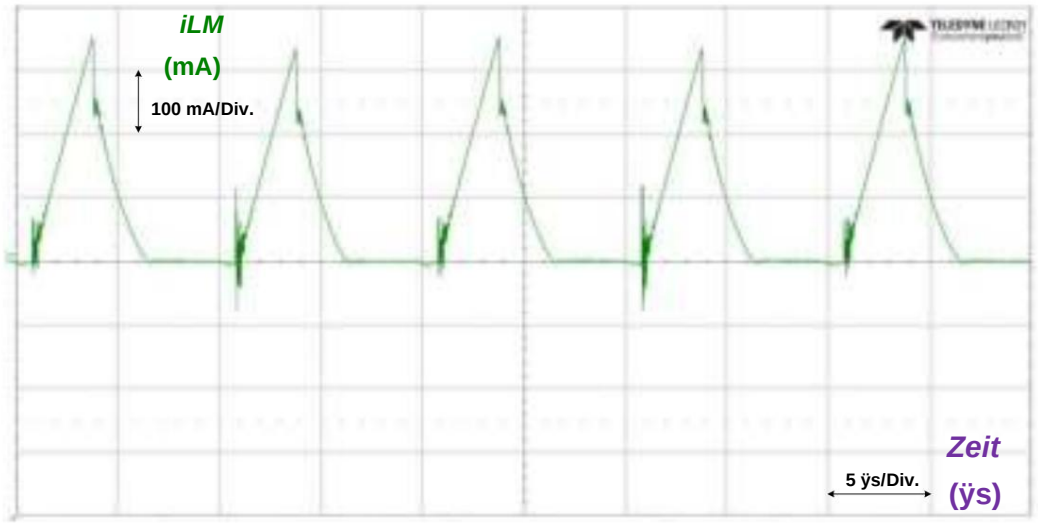


Abbildung 12. Gemessene entfaltete Wellenform des Magnetisierungsinduktorstroms *iLM* (100 mA/div.), Zeitskala: 5 µs/div.

Machine Translated by Google

Abbildung 12. Gemessene entfaltete Wellenform des Magnetisierungsinduktorstroms *iLM* (100 mA/div); Zeit (100 mA/d, Maßstab: 5 ÿs/div., Maßstab: 5 ÿs/div.

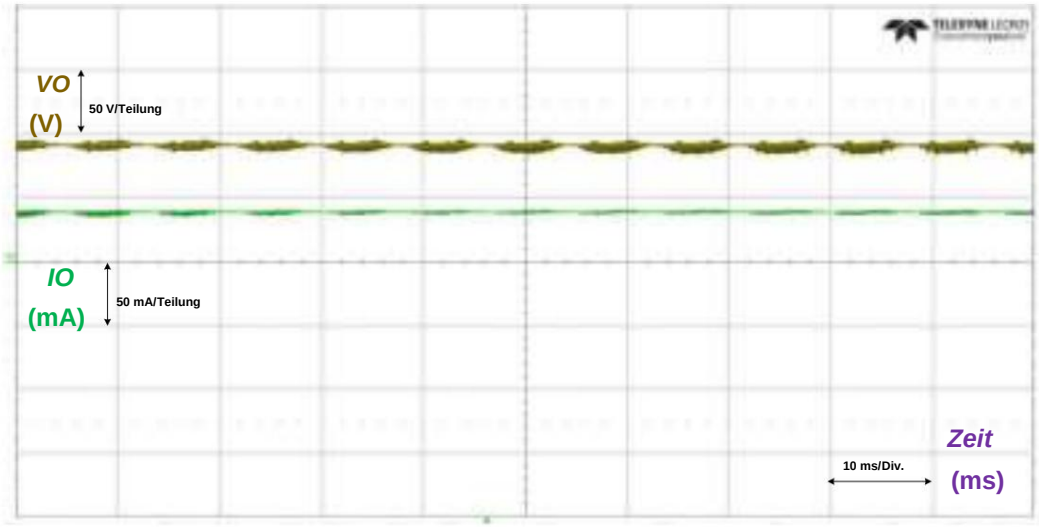


Abbildung 13. Gemessene Wellenformen der Ausgangsspannung *Vo* (50 V/div) und des Ausgangsstroms Io (50 mA/div); Zeitskala: 10 ms/div. mA/div); Zeitskala: 10 ms/div.

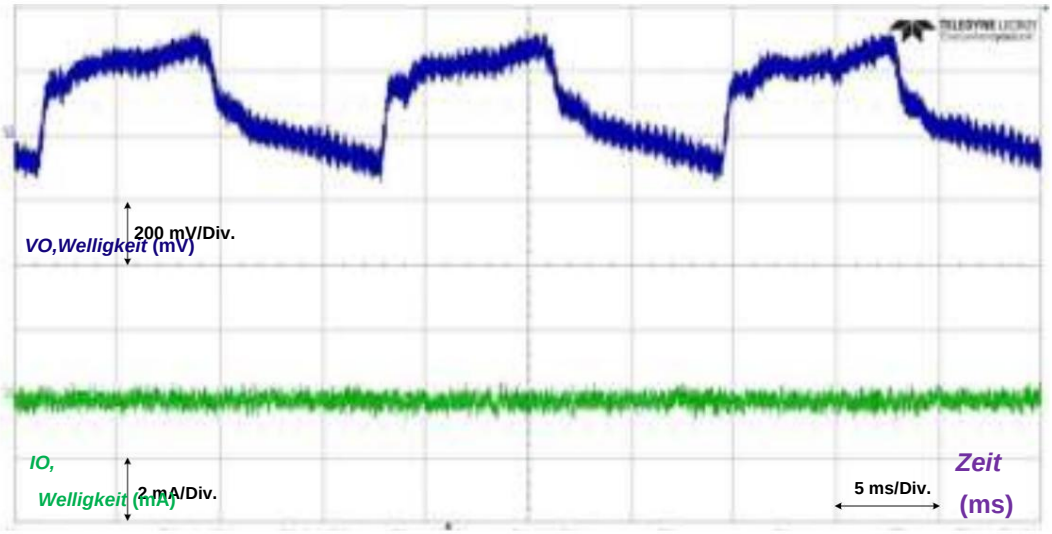


Abbildung 14. Gemessene Wellenformen der Ausgangsspannungswelligkeit Vo,ripple **(200 mV/div) und der Ausgangsstromwelligkeit Io ,ripple (2 mA/div); Zeitskala: 5 ms/div.**

Tabelle 4 zeigt den Ausgangsspannungswelligkeitsfaktor und den Ausgangsstromwelligkeitsfaktor des Prototyps einer elektronischen Beleuchtungstreiberschaltung für eine tief ultraviolette LED-Desinfektions- und Sterilisationslampe bei 110 Volt RMS AC-Eingangsspannung. Wie aus Tabelle 4 ersichtlich , ergibt sich aus der Division der Spitzen-Spitzen-Ausgangsspannung von 143,8 Millivolt durch den Durchschnittswert von 91,35 Volt ein Spannungswelligkeitsfaktor von 0,16 %; aus der Division des Spitzen-Spitzen-Ausgangsstroms von 1,1658 Milliampere durch den Durchschnittswert von 38,57 Milliampere ergibt

Abbildung 15 zeigt die gemessenen Wellenformen der AC-Eingangsspannung vAC **und des Eingangsstroms** iAC. **Aus der Abbildung ist ersichtlich, dass der AC-Eingangsstrom der Eingangsspannung folgt und die Phasen der beiden Wellenformen gleich sind; daher hat der Prototyp der vorgeschlagenen einstufigen Treiberschaltung zur Versorgung der tief ultravioletten LED-Desinfektions- und Sterilisationslampe den Effekt einer Leistungsfaktorkorrektur. Der Leistungsfaktor und der gesamte harmonische Verzerrungsfaktor des Eingangsstroms wurden mithilfe eines Leistungsanalysators mit 0,9236 bzw. 17,401 % gemessen. Darüber hinaus beträgt die gemessene Eingangsleistung der Prototyp-Treiberschaltung 3,901 W und die Ausgangsleistung 3,583 W, und der Wirkungsgrad der Prototyp- Schaltung beträgt 91,85 %.**

Machine Translated by Google

Parameter	Werte
Spitze-Spitze-Wert der Ausgangsspannung	143,8 mV
Mittelwert der Ausgangsspannung	91,35 V
Welligkeitsfaktor der Ausgangsspannung	0,16 %
Spitze-Spitze-Wert des Ausgangsstroms	1,1658 mA
Mittelwert des Ausgangsstroms	38,57 mA
Welligkeitsfaktor des Ausgangsstroms	3,02 %

Tabelle 4. Messung des Welligkeitsfaktors der Ausgangsspannung und des Welligkeitsfaktors des Ausgangsstroms von elektronischen Beleuchtungs-Prototyp-Treiberschaltung unter Verwendung von SiC-Schottky-Dioden für eine LED-Desinfektion im tiefen Ultraviolettbereich

Abbildung 15 zeigt gemessene Wellenformen der AC-Eingangsspannung *vAC* und des Eingangsstroms sowie der Serilisationslampe *iAC*. Aus der Abbildung ist ersichtlich, dass der AC-Eingangsstrom der Eingangsspannung folgt und die Phasen der beiden Wellenformen sind gleich; daher hat der Prototyp der einstufigen Treiberschaltung Sterilisationslampe mit der ultravioletten LED-Strahlung den Effekt einer Leistungsfaktorkorrektur. Der Gesamtwelligkeitsfaktor der Ausgangsspannung werden jeweils mit einem Leistungsanalysator gemessen. Der harmonische Verzerrungsfaktor des Eingangsstroms wurde mit 0,9236 und 17,401 % gemessen. Darüber hinaus beträgt die gemessene Eingangsleistung der Prototyp - Treiberschaltung 3,901 W und die Ausgangsleistung 3,583 W, und die Effizienz des Welligkeitsfaktors der Ausgangsstrom- Prototypschaltung beträgt 91,85 %.

Parameter	Werte
Zur Versorgung der Desinfektions- und Leistungsfaktor und der Gesamtwelligkeitsfaktor der Ausgangsspannung Darüber hinaus beträgt die gemessene Eingangsleistung der Prototyp - Ausgangsstrom- Prototypschaltung	143,8 mV 91,35 V 0,16 % 1,1658 mA 38,57 mA 3,02 %

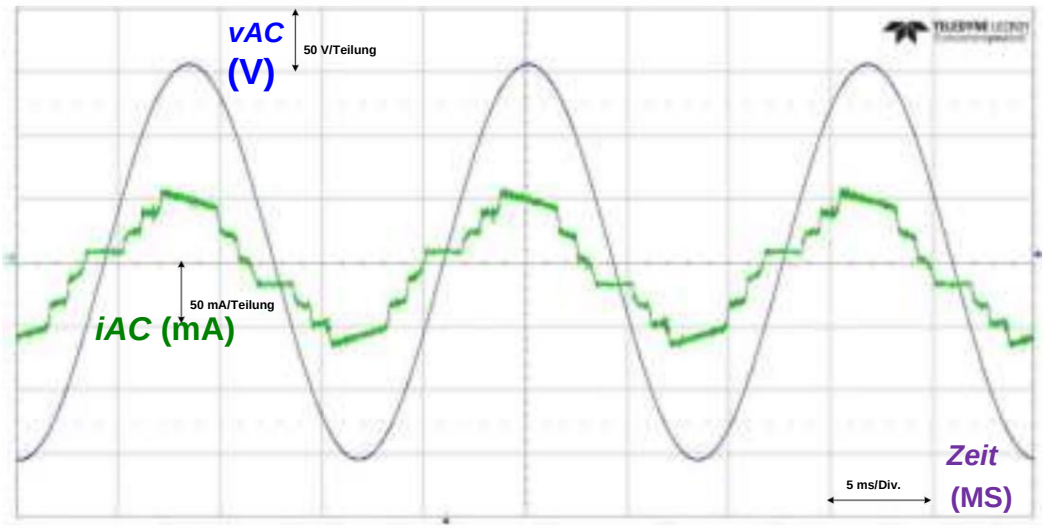


Abbildung 15. Gemessene Netzeingangsspannung *vAC* (50 V/div) und Stromstärke *iAC* (50 mA/div); Zeitskala: Abbildung 15. Gemessene Netzeingangsspannung vAC (50 V/div) und Stromstärke iAC (50 mA/div); 5 ms/Div.
Skala: 5 ms/div.

Tabelle 5 zeigt einen Vergleich zwischen dem bestehenden AC-DC-LED-Treiber in [19], der einen inversen Buck-Boost-Wandler mit einem verlustfreien Snubber integriert und eine Leistung von 18 W liefert, und dem vorgeschlagenen Treiber, der einen Abwärtswandler und einen Flyback-Wandler integriert und eine Nennleistung von 3,6 W (90 V/0,04 A) liefert. Wie aus Tabelle 3 ersichtlich, liefert der Wandler (60 V/0,3 A) und die vorgeschlagene Version, die einen Abwärtswandler und einen Flyback-Wandler integriert und eine Nennleistung von 3,6 W (90 V/0,04 A) liefert. Wie aus Tabelle 3 ersichtlich, liefert der Wandler

Beide LED-Treiber werden mit einer Eingangswechselspannung von 110 V versorgt und verwenden ein einziges Netzteil. Netzschalter. Die vorgeschlagene AC-DC-LED-Treiberschaltung spart einen Kondensator, ein magnetisches Element und eine Diode im Vergleich zu denen in [19]. Darüber hinaus ist der Strom THD und die Schaltung des Wirkungsgrads des vorgeschlagenen AC-DC-LED-Treibers besser als die des vorhandenen Treibers.

Beide LED-Treiber werden mit einer Eingangswechselspannung von 110 V versorgt und verwenden ein einziges Netzteil. Netzschalter. Die vorgeschlagene AC-DC-LED-Treiberschaltung spart einen Kondensator, ein magnetisches Element und eine Diode im Vergleich zu denen in [19]. Darüber hinaus sind der Strom THD und die Schaltung des Wirkungsgrads des vorgeschlagenen AC-DC-LED-Treibers besser als die des vorhandenen Treibers.

Tabelle 5. Vergleiche zwischen dem bestehenden AC-DC-LED-Treiber in [19] und dem vorgeschlagenen.

Artikel	Vorhandener AC-DC-LED-Treiber in Referenz [19]	Vorgeschlagener AC-DC-LED-Treiber
Schaltungstopologie	Integration eines inversen Buck-Boost Konverter mit verlustfreiem Snubber	Integration eines Abwärtswandlers und eines Flyback-Wandler
Eingangswechselspannung	110 V	110 V
Ausgangsleistung	18 W (60 V/0,3 A) 1	3,6 W (90 V/0,04 A)
Anzahl der benötigten Schalter		1
Anzahl der benötigten Kondensatoren	4	3
Anzahl der benötigten Magnetelemente	3	2
Anzahl der benötigten Dioden	7	6
Gemessener Leistungsfaktor	0,9737	0,9236
Gemessener Strom THD	18,422 %	17,401 %
Gemessener Schaltungswirkungsgrad	85,01 %	91,85 %

Machine Translated by Google

4. Schlussfolgerung

In diesem Artikel wird ein Prototyp einer einstufigen Treiberschaltung für elektronische Beleuchtung unter Verwendung von SiC-Schottky-Dioden für eine Desinfektions- und Sterilisationslampe mit tief ultravioletter LED vorgeschlagen, die einen Abwärtswandler und einen Sperrwandler integriert, um eine einstufige AC-DC-Stromrichterschaltungsarchitektur mit einem Leistungsschalter und einer Faktorkorrekturfunktion zu bilden. Durch die Ausgestaltung der Induktivität im AC-DC-Stromrichter für den Betrieb im diskontinuierlichen Leitungsmodus kann der Effekt der Leistungsfaktorkorrektur auf natürliche Weise erzielt werden. Außerdem verwendet die vorgeschlagene einstufige Treiberschaltung für elektronische Beleuchtung eine Mehrheitsträger-Schottky- Diode aus Siliziumkarbid-Halbleitermaterial mit großem Bandabstand und gewinnt die Energie aus der Streuinduktivität des Transformators zurück, wodurch die Umwandlungseffizienz der Treiberschaltung verbessert wird. In diesem Artikel wurde ein Prototyp einer Treiberschaltung für eine elektronische Beleuchtung mit tief ultravioletter LED-Beleuchtung mit 3,6 Watt für eine Desinfektions- und Sterilisationslampe entwickelt und getestet. Bei einer AC-Eingangsspannung von 110 Volt RMS und Nennausgangsleistung waren die Ergebnisse: Ausgangsspannungswelligkeitsfaktor unter 1 %, Ausgangsstromwelligkeitsfaktor unter 4 %, Leistungsfaktor über 0,9, Gesamtklirrfaktor des Eingangsstroms unter 18 % und Schaltungseffizienz über 90 %. Der in diesem Dokument vorgeschlagene Prototyp einer einstufigen Treiberschaltung für tief ultr

Beiträge der Autoren: C.-AC entwickelte die Schaltungstopologie des elektronischen Beleuchtungstreibers und entwarf die Methodik; C.-HC, H.-LC und E.-CC arrangierten die Softwareressourcen und richteten die Simulation ein und führten Schaltungssimulationen durch; L.-FL, S.-HH und C.-KL implementierten den Prototyp der elektronischen Beleuchtungstreiberschaltung und führten die Messungen der Schaltung durch sowie analysierten experimentelle Ergebnisse unter Anleitung von C.-AC; C.-AC erstellte den ersten Entwurf des Artikels und überarbeitete das Manuskript für die Einreichung. Alle Autoren haben die veröffentlichte **Version des Manuskripts gelesen und ihr zugestimmt .**

Finanzierung: Diese Forschung wurde vom National Science and Technology Council (NSTC) von Taiwan mit den Zuschussnummern MOST 110-2221-E-214-014, NSTC 111-2221-E-214-011 und NSTC 112-2221 -E-214-005 finanziert.

Erklärung zur Datenverfügbarkeit: Die in dieser Studie vorgestellten Daten sind im Artikel verfügbar.

Danksagungen: Die Autoren möchten dem National Science and Technology Council (NSTC) von Taiwan ihren aufrichtigen Dank für die Zuschüsse mit den Nummern MOST 110-2221-E-214-014, NSTC 111-2221-E-214-011 und NSTC 112-2221-E-214-005 aussprechen.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Verweise

1. **Kebbi, Y.; Muhammad, AI; Sant'Ana, AS; Prado-Silva, LD; Liu, D.; Ding, T. Jüngste Fortschritte bei der Anwendung der UV-LED-Technologie zur mikrobiellen Inaktivierung: Fortschritt und Mechanismus.** Compr. Rev. Food Sci. Food Saf. **2020,** 19, **3501–3527. [Cross [PubMed]**
2. **Shatalov, M.; Jain, R.; Saxena, T.; Dobrinsky, A.; Shur, M. Entwicklung von Deep-UV-LEDs und aktuelle Probleme in der Material- und Gerätetechnologie. In** Semiconductors and Semimetals; **Elsevier: Amsterdam, Niederlande, 2017; Band 96, S. 45–83.**
3. **Koutchma, T.; Popovi´c, V.; Green, A. Überblick über die Ultraviolett-(UV)-LED-Technologie für Anwendungen in der Lebensmittelproduktion.** Ultraviolette LED-Technologie für Lebensmittelanwendungen; **Academic Press: Cambridge, MA, USA, 2019; S. 1–23.**
4. **Akgün, MP; Ünlütürk, S. Auswirkungen von ultravioletten Leuchtdioden (LEDs) auf die mikrobielle und enzymatische Inaktivierung von Apfelsa** Int. J. Food Microbiol. **2017,** 260, **65–74. [CrossRef] [PubMed]**
5. **Kaya, Z.; Yýldýz, S.; Ünlütürk, S. Einfluss von UV-C-Bestrahlung und Wärmebehandlung auf die Haltbarkeit von Zitronen-Melonensaft Mischung: Multivariater statistischer Ansatz.** Innov. Food Sci. Emerg. Technol. **2015,** 29, **230–239. [CrossRef]**
6. **Guan, Z.; Liu, P.; Zhou, T.; Zhou, L.; Zhang, D.; Xie, Q.; Yu, Q.; He, Y.; Wang, S.; Wang, X.; et al. Studie zur Lichtfeldregulierung der UVC-LED-Desinfektion für den Kühlkettentransport.** Appl. Sci. **2022,** 12, **1285. [CrossRef]**
7. **Liang, JJ; Liao, CC; Chang, CS; Lee, CY; Chen, SY; Huang, SB; Yeh, YF; Singh, KJ; Kuo, HC; Lin, YL; et al. Die Wirksamkeit von Prototypgeräten für fernes Ultraviolettlicht (UVC) mit unterschiedlichen Wellenlängen bei der Desinfektion von SARS-CoV-2.** A **2021,** 11, **10661. [CrossRef]**
8. **Juarez-Leon, FA; Soriano-Sanchez, AG; Rodriguez-Licea, MA; Perez-Pinal, FJ Design und Implementierung eines keimtötenden UVC-LED-Lampe.** IEEE Access **2020, 8, 196951–196962. [CrossRef] [PubMed]**

Machine Translated by Google

9. Alonso, JM; Costa, MAD; Ordiz, C. Integrierter Buck-Flyback-Wandler als Offline-Stromversorgung mit hohem Leistungsfaktor. IEEE Trans. Ind. Electron. **2008,** 55, **1090–1100. [CrossRef]**

10. Gacio, D.; Alonso, JM; Calleja, AJ; Garcia, J.; Rico-Secades, M. Ein universelles, einstufiges Hochleistungs-Netzteil für HB-LEDs basierend auf einem integrierten Abwärts-Flyback-Wandler. IEEE Trans. Ind. Electron. **2010,** 58, **589–599. [CrossRef]**

11. Abdelmessih, GZ; Alonso, JM; Costa, MAD Verlustanalyse zur Verbesserung der Effizienz des integrierten Buck-Flyback-LED- Treibers. IEEE Trans. Ind. Appl. **2018,** 54, **6543–6553. [CrossRef]**

12. Abdelmessih, GZ; Alonso, JM; Tsai, WT Analyse und Experimente an einem neuen High-Power-Factor-Offline-LED-Treiber basierend über verschachtelte integrierte Abwärts-Flyback-Wandler. IEEE Trans. Ind. Appl. **2019,** 55, **4359–4369. [CrossRef]**

13. Luz, PC; Bolán, **PE; Maboni, T.; Cosetin, MR; da Silva, MF; do Prado, RN Ein integrierter isolierter Buck-Boost-Flyback-Konverter zur Versorgung von LED-Lampen mit Straßenbeleuchtung mit reduzierten Kapazitäten. In Proceedings der 11. IEEE/IAS International Conference on Industry Applications 2014, Juiz de Fora, Brasilien, 7.–10. Dezember 2014; S. 1–6.**

14. Divya, KM; Parackal, R. Integrierter Abwärts-/Aufwärts-Flyback-Wandler mit hohem Leistungsfaktor, der mehrere Ausgänge ansteuert. In Proceedings der 2015 Online International Conference on Green Engineering and Technologies (IC-GET), Coimbatore, Indien, 27. November 2015; S. 1–5.

15. Cheng, CA; Chung, TY Ein neuartiger integrierter Treiber mit PFC zur Versorgung von Leuchtdioden. In Proceedings der 41. Jahreskonferenz der IEEE Industrial Electronics Society (IECON 2015), Yokohama, Japan, 9.–12. November 2015; S. 3514–3518.

16. Poorali, B.; Adib, E. Analyse des integrierten SEPIC-Flyback-Konverters als einstufiger Einschalter-Leistungsfaktorkorrektur LED-Treiber. IEEE Trans. Ind. Electron. **2016,** 63, **3562–3570. [CrossRef]**

17. Cheng, C.-A.; Chang, C.-H.; Cheng, H.-L.; Chang, M.-T. Ein neuartiger LED-Treiber mit Leistungsfaktorkorrektur, geeignet für Straßenlaternenanwendungen. In Proceedings der International Future Energy Electronics Conference 2017 (IFEEC 2017)—ECCE Asia, Kaohsiung, Taiwan, 3.–7. Juni 2017; S. 1–4.

18. Cheng, CA; Chang, CH; Cheng, HL; Chang, EC; Lai, CC Ein neuartiger LED-Röhrenlampentreiber mit Leistungsfaktorkorrektur für Innenbeleuchtungsanwendungen. In Proceedings der 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE), Nara, Japan, 9.–12. Oktober 2018; S. 700–701.

19. Cheng, CA; Chen, YL; Chen, YJ; Lai, SW Ein neuartiges kostengünstiges AC-DC-Netzteil für LED-Beleuchtungsanwendungen im Innenbereich. In Proceedings der 2019 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Macao, China, 1.–4. Dezember 2019; S. 1–3.

20. Cheng, CA; Lan, LF; Hou, SH; Lin, CK Ein einstufiger elektronischer Beleuchtungstreiber für eine tief ultraviolette LED-Desinfektions- und Sterilisationslampe. In Proceedings des 2023 IEEE Workshop on Wide Bandgap Power Devices and Applications in Asia (WiPDA Asia), Hsinchu, Taiwan, 27.–29. August 2023; S. 1–3.

21. Effektive Nutzung von Leistungsbauelementen – Verständnis von Leistungsbauelementen aus Siliziumkarbid und Anwendungsbeispiele unter Nutzung der Vorteil ROHM Semiconductor, 2017. Online verfügbar: https://pages.rohm.com.tw/rs/247-PYD-578/images/TechWeb_E_DL_SiC.pdf (abgerufen am 31. Juli 2024).

22. Anwendungshinweis zu SiC-Leistungsgeräten und -Modulen, ROHM Semiconductor, 2020. Online verfügbar: https://fscdn.rohm.com/en produkte/datenbuch/applinote/diskret/sic/common/sic_appli-e.pdf (abgerufen am 31. Juli 2024).

23. AN966 APPLICATION NOTE, STMicroelectronics, „L6561, ENHANCED TRANSITION MODE POWER FACTOR CORRECTOR“, S. 1–23, März 2003. Online verfügbar: https://www.st.com/resource/en/application_note/cd00004002-l6561-enhanced- transition-mode-power-factor-corrector-stmicroelectronics.pdf (abgerufen am 31. Juli 2024).

24. SEOULVIOSYS. UV CA3535-Serie (CUD5GF1B). Online verfügbar: https://media.digikey.com/pdf/Data%20Sheets/Seoul%20 Halbleiter/CUD5GF1B_R0.1_Prellimiminary_2-3-20.pdf (abgerufen am 31. Juli 2024).

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultieren.

Machine Translated by Google

Elektronik

MDPI

Artikel

Effiziente Sprachsignal-Dimensionsreduktion durch Komplexwertige Techniken

Sungkyun Ko [1] **und Minho Park 2,***

[1] Abteilung für KI-IT-Konvergenz, Soongsil University, Seoul 06978, Republik Korea; kosungkyun@naver.com
[2] Fakultät für Elektrotechnik, Soongsil University, Seoul 06978, Republik Korea
* Korrespondenz: mhp@ssu.ac.kr

Zusammenfassung: In dieser Studie schlagen wir den CVMFCC-DR- Algorithmus (Complex-Valued Mel-Frequency Cepstral Co-efficients Dimensionality Reduction) als effiziente Methode zur Reduzierung der Dimensionalität von Sprachsignalen vor. Durch die Verwendung der komplexwertigen MFCC-Technik, die sowohl reale als auch imaginäre Komponenten berücksichtigt, ermöglicht unser Algorithmus eine Dimensionsreduzierung ohne Informationsverlust bei gleichzeitiger Senkung des Rechenaufwands. Die Wirksamkeit des vorgeschlagenen Algorithmus wird durch Experimente bestätigt, die seine Effektivität beim Aufbau eines Spracherkennungsmodells unter Verwendung eines komplexwertigen neuronalen Netzwerks demonstrieren. Zusätzlich wird eine komplexwertige Softmax-Interpretationsmethode für komplexe Zahlen eingeführt. Die experimentellen Ergebnisse zeigen, dass der Ansatz im Vergleich zu herkömmlichen MFCC-basierten Techniken eine verbesserte Leistung bietet, was sein Potenzial im Bereich der Spra

Schlüsselwörter: Mel-Frequenz-Cepstrum-Koeffizienten; Fehlerkorrektur bei komplexwertigen neuronalen Netzwerken; Backpropagation; komplexwertiger Softmax

check for updates

Zitat: Ko, S.; Park, M. Effiziente Reduzierung der Sprachsignaldimensionalität mithilfe komplexwertiger Techniken. Electronics **2024,** 13, **3046. https://doi.org/10.3390/Elektronik13153046**

Wissenschaftliche Herausgeber: Xin Geng, Padma Iyenghar und Elke Pulvermüller

Erhalten: 3. Juni 2024
Überarbeitet: 31. Juli 2024
Akzeptiert: 31. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Einleitung

Heutzutage spielen Sprachsignalverarbeitung und -erkennung in zahlreichen Anwendungsbereichen eine entscheidende Rolle . Diese Technologien werden für unterschiedliche Zwecke eingesetzt, beispielsweise um Sprachbefehle in Text umzuwandeln oder den Besitzer einer Stimme zu identifizieren. Die Weiterentwicklung der Spracherkennungstechnologie kann die Mensch-Maschine-Interaktion verbessern und so die Produktivität und Effizienz in zahlreichen Branchen steigern. Sprachsignale weisen jedoch sehr unterschiedliche Formen und hochdimensionale Eigenschaften auf, was eine effektive Verarbeitung und Erkennung zu einer ständigen Herausforderung macht. Die Vielfalt und hohe Dimensionalität von Sprachsignalen erschweren die Dimensionsreduzierung und Mustererkennung, was zu erhöhten Rechenkosten führt. Daher ist die effiziente Reduzierung der Dimensionalität von Sprachsignalen und die Verbesserung der Leistung neuronaler netzbasierter S

Technologien zur Sprachverarbeitung und -erkennung sind für zahlreiche Anwendungen wie Sprachschnittstellen, virtuelle Assistenten und automatische Übersetzungen unverzichtbar. Genauere und effizientere Ergebnisse bei der Sprachsignalverarbeitung und -erkennung können das Benutzererlebnis bei entsprechenden Diensten erheblich verbessern. Aus diesem Grund wird im Bereich der Sprachsignalverarbeitung und -erkennung kontinuierlich geforscht. Sprachsignale weisen hochdimensionale Eigenschaften auf, was eine effektive Dimensionsreduzierung Darüber hinaus erfordert die genaue Erkennung unterschiedlicher Sprachmuster komplexe neuronale Netzwerkmodelle , was wiederum den Rechenaufwand erhöht [1]. Daher ist es eine entscheidende Aufgabe, die Dimensionalität von Sprachsignalen ohne Informationsverlust effektiv zu reduzieren und gleichzeitig die Genauigkeit und Effizienz neuronaler Netzwerk-basierter Spracherkennungsmodelle zu verbessern .

Trotz der umfangreichen Literatur zu komplexwertigen Netzwerken gibt es nur wenige Beispiele für deren Anwendung in der Audio- oder Sprachverarbeitung. Da das menschliche Hörsys

Machine Translated by Google

Da neuronale Netzwerke Schall im Frequenzbereich verarbeiten, ist es auch sinnvoll, komplexwertige neuronale Netzwerke für die Audioverarbeitung zu verwenden, anstatt sich ausschließlich auf realwertige neuronale Netzwerke im Zeitbereich zu verlassen **[2].**

Durch die Kombination **traditioneller Methoden wie der Mel-Frequency-Cepstral-Koeffizienten (MFCCs) mit neuronalen Netzwerken wurden in der jüngsten Vergangenheit erhebliche Fortschritte bei der Spracherkennung erzi Obwohl die MFCC-Technik häufig zur Extraktion von Sprachmerkmalen verwendet wird, hat sie den Nachteil, dass sie keine verlustfreie Dimensionsreduzierung durchführen kann, da sie nur reale Werte verwendet. Mit anderen Worten: Die Reduzierung der Anzahl der MFCCs führt zu Informationsverlusten, was die Genauigkeit der Spracherkennung verringert. Diese Methoden sind immer noch nur begrenzt effektiv und es sind verbesserte Ansätze erforderlich, die über traditionelle Lernmethoden hinausgehen.**

Die Verwendung der komplexen Darstellung von MFCC ermöglicht die Berücksichtigung sowohl des Real- als auch des Imaginärteils und ermöglicht so eine verlustfreie Dimensionsreduktion. Darüber hinaus können durch den Einsatz komplexwertiger neuronaler Netze Phasenänderungen durch komplexe Multiplikationsoperationen abgebildet werden [4], wodurch die Eigenschaften von Sprachsignalen besser dargestellt werden. Dieser Ansatz dürfte im Vergleich zu bestehenden Methoden eine effizientere Dimensionsreduktion und eine auf neuronalen Netzen basierende Spracherkennung ermögliche

In dieser Studie schlagen wir den CVMFCC-DR-Algorithmus zur effizienten und verlustfreien Reduzierung der Sprachsignaldimensionalität vor. Durch die Kombination von CVMFCC-DR mit komplexwertigen neuronalen Netzwerken kann diese Technik **im Vergleich zu bestehenden Methoden eine bessere Spracherkennungsleistung aufweisen .**

Die wichtigsten Beiträge dieser Studie sind wie folgt:

- **Vorschlag einer effizienten Dimensionsreduktionstechnik (CVMFCC-DR) unter Verwendung der komplexen Darstellung von MFCCs.**
- **Verbesserung der Spracherkennungsleistung durch Kombination von komplexwertigen MFCCs und komplexwertigen neuronalen Netzwerken.**
- **Vorschlag einer neuen Interpretationsmethode für eine komplexwertige Softmax-Funktion mit komplexe Eingaben.**

Dieser Artikel beginnt mit einer Einführung in die bestehende MFCC-Technik und komplexwertige neuronale Netze. Anschließend werden der vorgeschlagene CVMFCC-DR- Algorithmus und die Lernmethode für komplexwertige neuronale Netze [5] ausführlich erläutert. Anschließend wird die Wirksamkeit der vorgeschlagenen Methode experimentell validiert und die Ergebnisse analysiert . Abschließend schließt der Artikel mit einer Zusammenfassung der Ergebnisse der Studie und Vorschlägen für zukünftige Forschungen.

2. Verwandte Werke

Im Bereich der Spracherkennung werden verschiedene Ansätze und Techniken auf Basis neuronaler Netze eingesetzt. Dabei kommen unterschiedliche Arten neuronaler Netze zum Einsatz, wie CNNs, RNNs, LSTM und DNNs, die jeweils ihre eigenen Vor- und Nachteile haben [3]. Darüber hinaus werden auch Hybridmodelle erforscht, die traditionelle Methoden mit neuronalen Netzstrukturen kombinieren . Hybridmodelle können die Leistung optimieren, indem sie die Stärken jeder Struktur kombinieren, haben aber auch den Nachteil einer erhöhten Komplexität bei der Implementierung und Interpretation [6].

Inmitten dieser Entwicklungen werden traditionelle Techniken zur Merkmalsextraktion, wie die MFCC- Technik, noch immer häufig als Eingaben für neuronale Netzwerke verwendet [3]. Darüber hinaus könnte die Erforschung neuartiger neuronaler Architekturen, wie etwa komplexwertiger neuronaler Netzwerke, neue Perspektiven und potenzielle Durchbrüche bei der Spracherkennung eröffnen. Jüngste Studien haben gezeigt, dass sich die Berechnungsgeschwindigkeit komplexer Zahlen mithilfe von FPGAs deutlich verbessern lässt [7]. Diese Forschungsergebnisse zeigen empirisch, dass die Hardwareimplementierung von Operationen mit komplexen Zahlen praktische Leistungsverbesserungen mit sich bringen kann . Darüber hinaus wurden in einer kürzlich durchgeführten Studie komplexwertige neuronale Netzwerke zur Klassifizierung auf FPGAs implementiert. Dabei wurde eine höhere Genauigkeit als realwertige neuronale Netzwerke nachgewiesen und im Vergleich zu CPU-basierten **Systemen eine deutliche Geschwindigkeitsverbesserung erreicht [8]. Dies lässt erwarten, dass CPUs, die heute nur Operationen mit reellen Zahlen unterstützen, in Zukunft vielle**

Machine Translated by Google

Basierend auf diesen Studien ermöglicht die Verwendung komplexer Zahlendarstellungen die Kodierung von reellen Zahlenpaaren als einzelne komplexe Zahlen, wodurch die Eingabedimension effektiv um die Hälfte reduziert wird. Diese Reduzierung der Dimensionalität verringert die Anzahl der Modellparameter und verbessert die Rechenleistung. Folglich zeigen experimentelle Beweise, dass komplexwertige neuronale Netzwerke im Vergleich zu herkömmlichen realwertigen neuronalen Netzwerken mit weniger Parametern eine gleiche oder bessere Leistung erzielen können.

Während die Forschung immer neue Grenzen erweitert, könnte eine synergetische Kombination aus bewährten Methoden und innovativen Ansätzen den Weg für die nächste Generation robuster und leistungsstarker Spracherkennungssysteme ebnen.

3. Hintergrundwissen

Es werden derzeit verschiedene Studien im Bereich der Spracherkennung und -klassifizierung durchgeführt. In diesem Abschnitt stellen wir MFCCs (Mel-Frequency Cepstral Coefficients) vor [3]. Die MFCC-Technik ist eine der in den Bereichen Sprachsignalverarbeitung und Spracherkennung häufig verwendeten Methoden zur Merkmalsextraktion. Die MFCC-Technik stellt die Frequenzeigenschaften von Sprachsignalen dar und wird hauptsächlich zur effektiven Darstellung und Erkennung von Sprachsignalen eingesetzt. Das Grundprinzip der MFCC-Technik ist in Abbildung 1 dargestellt.

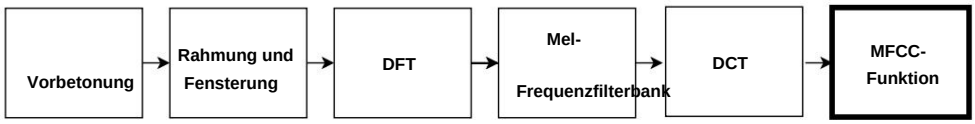


Abbildung 1. Ein Überblick über den Prozess der MFCC-Technik.

Vorverzerrung: Bei der Vorverzerrung handelt es sich im Wesentlichen um einen Hochpassfilter. Aufgrund der Struktur des menschlichen Körpers neigen hochfrequente Komponenten dazu, bei der Sprachproduktion im tatsächlich abgegebenen Ton abzunehmen. Daher handelt es sich bei der Vorverzerrung um den Prozess, bei dem zunächst ein Hochpassfilter angewendet wird, um die gedämpften hochfrequenten Komponenten zu verstärken.

Framing und Windowing: Sprachsignale werden zur Analyse in kurze Segmente, sogenannte Frames, unterteilt. Dies liegt daran, dass die Spracheigenschaften über kurze Zeiträume relativ stabil bleiben. Diese Frames überlappen sich leicht, um Sprachänderungen effektiver zu erfassen.
Jeder Frame wird mithilfe einer Fenstertechnik verarbeitet, die die Analyse verbessert. Zu den gängigen Fenstertypen gehören Hamming- und Hanning-Fenster. Diese Framing-Methode ermöglicht eine genauere Analyse komplexer Sprachdaten, was für die Entwicklung von Spracherkennungs- und -synthesetechnologien von entscheidender Bedeutung ist. Während die spezifischen Implementierungsdetails variieren können, werden diese grundlegenden Prinzipien in Sprachverarbeitungssystemen häufig verwendet.

Diskrete Fourier-Transformation: Zur Berechnung des Leistungsspektrums wird die DFT (Diskrete Fourier-Transformation) verwendet, um das Leistungsspektrum für jeden zuvor erhaltenen Frame zu berechnen. Dieser Prozess ermöglicht die Unterscheidung der frequenzspezifischen Energieverteilung des Sprachsignals über jeden Frame hinweg.

Mel-Frequenz-Filterbank: Der Mel-Bandpassfilter ist eine Reihe von Filtern, die dazu dienen, Sprachsignale auf nichtlineare Weise zu extrahieren und so das menschliche Hörsystem nachzuahmen.
Es besteht normalerweise aus 40 Dreieckfiltern und analysiert die Frequenzeigenschaften von Sprachsignalen.

Diskrete Kosinustransformation: DCT (Diskrete Kosinustransformation) drückt eine endliche Folge von Datenpunkten als Summe von Kosinusfunktionen aus, die mit unterschiedlichen Frequenzen schwingen. Im MFCC-Prozess wird DCT auf die Mel-Filterbank angewendet, um Beschleunigungskoeffizienten auszuwählen oder die Beziehung in log-spektralen Größen von der Filterbank zu trennen.

Spracherkennung mit MFCCs kann in verschiedenen Umgebungen effektiv eingesetzt werden [9], und diese Methode kann in Kombination mit Deep-Learning-Modellen eine hohe Genauigkeit erreichen . Daher dient MFCC als entscheidender Input für Deep-Learning-Modelle zur Spracherkennung [10].

Machine Translated by Google

4. Vorgeschlagene

Methode In diesem Abschnitt schlagen wir eine Methode zur effizienten Durchführung einer Dimensionsreduzierung von Sprachsignalen unter Verwendung komplexwertiger MFCCs und komplexwertiger neuro Abbildung 2 bietet einen Überblick über den vorgeschlagenen Ansatz. Diese Abbildung veranschaulicht den vorgeschlagenen Ansatz zur Sprachklassifizierung mithilfe komplexwertiger neuronaler Netzwerke. Der Prozess beginnt mit der Spracheingabe, gefolgt von der Merkmalsextraktion mithilfe von Mel-Frequency- Cepstrum-Koeffizienten (MFCCs). Diese Merkmale werden dann einer Dimensionsreduzierung unterzogen und in komplexwertige Darstellungen u Die verarbeiteten Merkmale werden anschließend in ein komplexwertiges neuronales Netzwerk eingegeben. Die Leistung des Netzwerks wird durch die Auswahl geeigneter Aktivierungsfunktionen und Lernalgorithmen verbessert. Schließlich wird die Ausgabe des neuronalen Netzwerks für Klassifizierungszwecke verwendet.

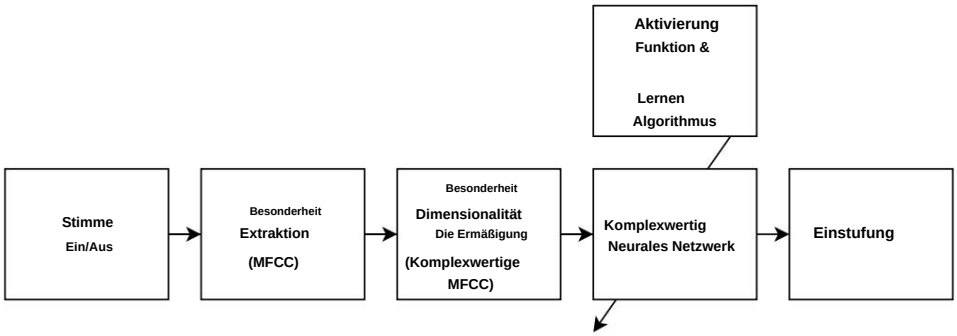


Abbildung 2. Übersicht über den vorgeschlagenen Ansatz.

Komplexwertige MFCCs: Die Anzahl der bei der Spracherkennung verwendeten MFCCs kann je nach Anwendung und Modell variieren. Normalerweise werden jedoch MFCCs im Bereich von 13 bis 40 verwendet. Dieser Bereich funktioniert nachweislich in verschiedenen Spracherkennungssystemen effektiv . Im Allgemeinen kann eine höhere Anzahl von MFCCs dem Modell ermöglichen, mehr Frequenz- und Spektraldetails zu erfassen, aber dies erhöht auch den Re Daher wird die ausgewählte Anzahl von MFCCs unter Berücksichtigung eines Gleichgewichts zwischen Rechenleistung und Modellleistung bestimmt.

Bei der Darstellung physikalischer Probleme mit komplexen Zahlen gibt es zwei Hauptmethoden . Eine davon ist eine einfache Substitutionsmethode und die andere ist ein eleganterer Ansatz, der als mathematische Äquivalenz bekannt ist, wie die Fourier-Transformation. Bei der Substitution werden zwei reale physikalische Parameter verwendet und einer davon in den Realteil der komplexen Zahl und der andere in den Imaginärteil gesetzt. Dadurch können die beiden Werte als eine einzige Entität, d. h. eine einzige komplexe Zahl, verarbeitet werden **[11]. Die in diesem Dokument zur Dimensionsreduzierung vorgeschlagene Methode umfasst die Umwandlung von n MFCC-Werten in komplexe Zahlen. Der erste MFCC wird als Re Anschließend wechselt jeder MFCC zwischen dem Real- und Imaginärteil innerhalb der komplexen Zahl. Dieses Muster setzt sich bis zum n-ten MFCC fort, was zu n/2 komplexen Zahlen führt. Wenn wir beispielsweise** n **MFCC-Werte haben, könnten diese wie folgt lauten:**

[x1, x2, x3, x4, . . . , xnÿ1, xn]

In diesem Fall wird der erste MFCC x1 als Realteil der ersten komplexen Zahl verwendet. Der zweite MFCC x2 wird als Imaginärteil der ersten komplexen Zahl verwendet. Der dritte MFCC x3 wird als Realteil der zweiten komplexen Zahl verwendet. Der vierte MFCC x4 wird als Imaginärteil der zweiten komplexen Zahl verwendet. Dieses Muster wechselt zwischen Real- und Imaginärteilen bis zum n-ten MFCC xn, **der zum Imaginärteil der** (n/2)-ten **komplexen Zahl wird. Daher werden mit n MFCC-Werten n/2 komplexe Zahlen erzeugt. Die erste komplexe Zahl ist x1** + **jx2, die zweite komplexe**

Machine Translated by Google

Zahl ist x3 + **jx4, und so weiter bis zur letzten komplexen Zahl, die** xnÿ1 + jxn ist. **Wir haben diese Erklärung in Abbildung 3 illustriert.**

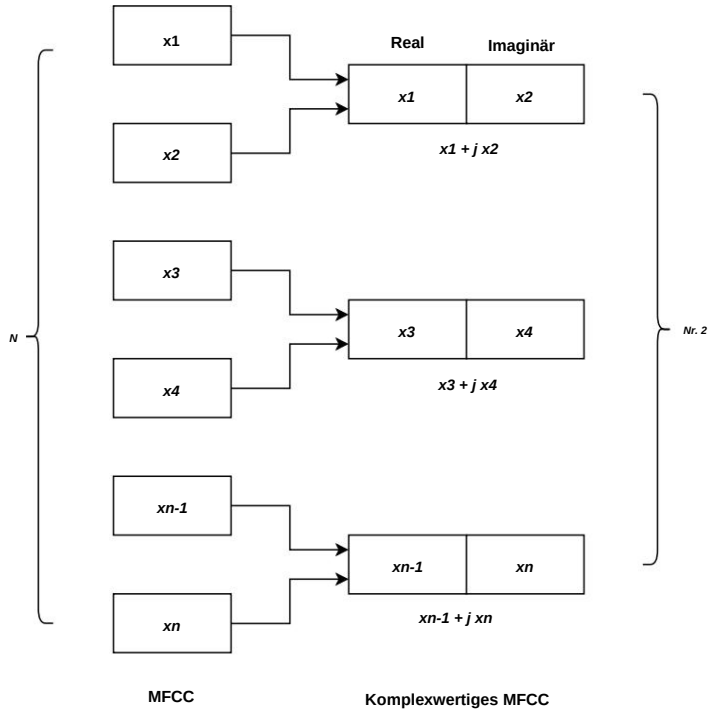


Abbildung 3. Prozess der Generierung komplexwertiger MFCCs.

Komplexwertiges neuronales Netzwerk: Komplexwertige neuronale Netzwerke sind neuronale Netzwerkmodelle, die komplexe Zahlen als Eingabewerte oder Gewichte verwenden, und sie haben den Vorteil, dass sie im Vergleich zu realwertigen neuronalen Netzwerken umfangreichere Informationen verarbeiten können. Darüber hinaus wurden in verschiedenen Bereichen verschiedene Forschungsstudien durchgeführt, und komplexwertige Methoden zur Dimensionsreduzierung oder Klassifizierung und Erkennung mithilfe komplexer Zahlen werden effektiv eingesetzt. Zu diesem Zweck wird die Multiplikation komplexer Eingaben und komplexer Gewichte durchgeführt, was eine Rotation in der komplexen Ebene impliziert und somit Phasenänderungen einführt [5]. Die Summe der Multiplikation von Eingabe X und Gewicht W kann als Gleic

$$Z = \mathbf{W0} + W1X1 + W2X2 + \ldots + WnXn \quad \textbf{(1)}$$

$$Z = \ddot{y}n\ Wn \cdot Xn \quad \textbf{(2)}$$

Aktivierungsfunktion und Lernalgorithmus: In komplexwertigen neuronalen Netzwerken wird eine phasenabhängige Aktivierungsfunktion verwendet, die eine Normalisierung der Betragsinformation auf einen Einheitskreis von 1 und nur die Extraktion der Phaseninformation erzwingt, wie in Gleichung (3) gezeigt.

$$P(Z) = \frac{Z}{|Z|} = e^{j \cdot \arg(Z)} \quad \textbf{(3)}$$

Der Lernalgorithmus komplexwertiger neuronaler Netze funktioniert wie folgt: Wenn ein gewünschtes Ziel t und ein Ausgabewert z vorliegen, wie in Abbildung 4 dargestellt, werden die Gewichte gemäß Gleichung (4) angepasst , **um sich schrittweise in Richtung des gewünschten Ziels zu be**

Machine Translated by Google

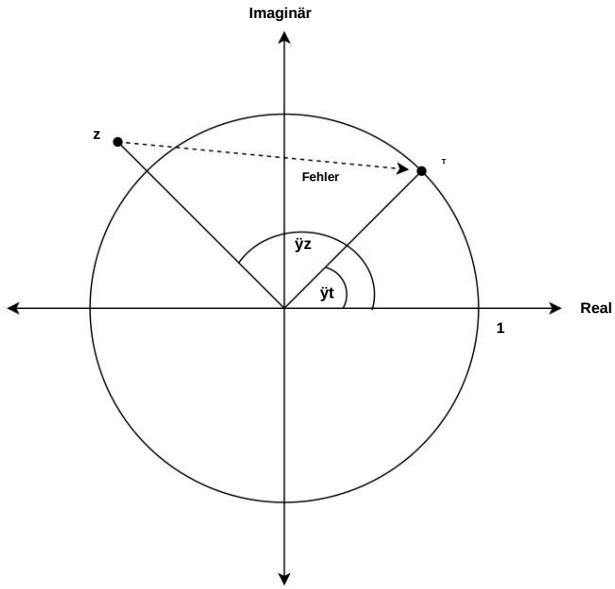


Abbildung 4. Lernalgorithmus für CVNN.

$$t = ÿn\ (Wn + ÿWn)\cdot Xn \tag{4}$$

Der Fehler in Bezug auf das Trainingsziel t **beträgt** $e = t ÿ z$ **wie in Gleichung (5) gezeigt.**

$$e = ÿn\ (Wn + ÿWn)\cdot Xn ÿ ÿn\ (Wn)\cdot Xn = ÿn\ (ÿWn)\cdot Xn \tag{5}$$

Um die Berechnung des Fehlers in Bezug auf Gewichtsänderungen zu vereinfachen und zu beschleunigen, wird unter der Annahme, dass jeder Knoten gleichermaßen zum Fehler beiträgt, das Summationssymbol *ÿ* **aus der Gleichung entfernt und beide Seiten werden zur Vereinfachung der Berechnung durch die Anzahl der beitragenden Knoten N im komplexwertigen neuronalen Netzwerk geteilt. Die resultierende Gleichung ist Gleichung (6).**

$$ÿWn \cdot Xn = \frac{t}{N} \tag{6}$$

Um ÿWn **zu erhalten und damit die Gewichte zu ändern, multiplizieren Sie beide Seiten der** $X_{N}^{ÿ1}$ $X_{N}^{ÿ1}$ **Gleichung (6) mit Xn , wenn Sie es mit der komplex konjugierten** Zahl Xn multiplizieren, **wobei n sowohl Zähler als auch Nenner ist. Somit ist** $\frac{Xn}{|Xn|} = \frac{Xn}{|Xn|^2}$ **wobei** |Xn| = **1 für** Xn **innerhalb der Einheit Kreis. Daher ist** $X_{N}^{ÿ1} = \overline{Xn}$, **und Gleichung (6)** wird erneut als Gleichungen **(7) und (8) hergeleitet .**

$$ÿWn \cdot Xn \cdot X_{N}^{ÿ1} = \frac{t}{N} \cdot X_{N}^{ÿ1} = \frac{t}{N} \cdot \overline{Xn} \tag{7}$$

$$ÿWn = \frac{t}{N} \cdot \overline{Xn} \tag{8}$$

Auf diese Weise kann die erzielte Änderung der Gewichte in Gleichung (8) beobachtet werden. Diese Änderung stellt die Anpassung der komplexen Gewichte für jeden Knoten dar, der zum Fehler in Bezug auf das Trainingsziel beiträgt, und wird einheitlich auf das gesamte Netzwerk angewendet. Durch diesen Prozess passen sich komplexwertige neuronale Netzwerke unter Verwendung phasenabhängiger Aktivierungsfunktionen und effizienter Lernalgorithmen näher an das Ziel an, während sie reichhaltige Informationen durch komplexe Multiplikationsoperationen zwischen Eingaben und G

Komplexe Domänen-Backpropagation: Bei der Verwendung von versteckten Schichten wird der Fehler von der Ausgabeschicht an die versteckten Schichten zurückpropagiert, um zu berechnen, wie viel jeder

Machine Translated by Google

in den verborgenen Schichten haben zu diesem Fehler beigetragen. Darauf basierend müssen die Gewichte zwischen den verborgenen und den Ausgabeschichten entsprechend angepasst werden. Dieser Prozess wird als Fehlerrückausbreitung bezeichnet , durch den das neuronale Netzwerk die Gewichte anpassen kann, um den Fehler in den Ausgabewerten während des Lernens zu minimieren . Die Struktur eines mehrschichtigen komplexwertigen neuronalen Netzwerks ist in Abbildung 5 dargestellt.

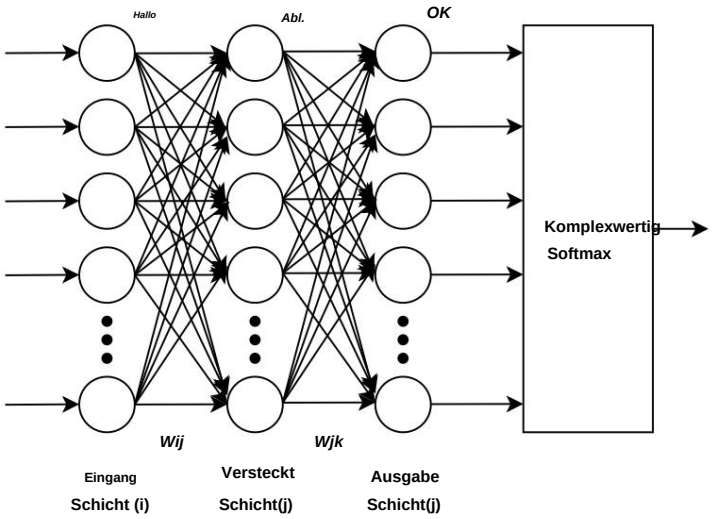


Abbildung 5. Mehrschichtiges CVNN.

In diesem Artikel wird eine Fehlerkorrekturmethode für die Ausgabeschicht k verwendet, um die Gewichte zu aktualisieren, während eine Fehlerrückpropagationsmethode für die verborgene Schicht j verwendet wird, um ihren Einfluss auf die Ausgabe zu untersuchen. Mit anderen Worten: Die Ausgabeschicht k verwendet die Fehlerkorrekturmethode, um das Erreichen lokaler Minima zu vermeiden, während die Fehlerrückpropagationsmethode auf die verborgene Schicht j angewendet wird, um eine bessere Verteilung der Fehler von der Ausgabeschicht auf die verborgene Schicht zu gewährleisten. Diese Kombination aus Fehlerrückpropagation in der verborgenen Schicht und Fehlerkorrektur in der Ausgabeschicht dürfte das Problem lösen, dass das gesamte neuronale Netzwerk in lokalen Minima hängen bleibt, Die Fehler-Backpropagation-Methode für komplexwertige neuronale Netzwerke ist wie folgt.

Bei der Verwendung des Gradientenabstiegs zur Gewichtsaktualisierung in komplexwertigen neuronalen Netzen wird die Wirtinger-Rechnung zur Differenzierung komplexer Funktionen verwendet. Da komplexe Funktionen aus Real- und Imaginärteilen bestehen, führt die Wirtinger-Rechnung die Differenzierung getrennt für Real- und Imaginärteile durch. Für eine komplexe Zahl z = x + jy , **die mit Real- und Imaginärteilen** x **bzw. y definiert ist, und eine komplexe Funktion** f(z) = u(x, y) + jv(x, y) werden **die Wirtinger-Ableitungen wie folgt definiert [14].**

$$\frac{\partial f}{\partial z} = \frac{1}{2}\left(\frac{\partial f}{\partial x} - j\frac{\partial f}{\partial y}\right) = \frac{1}{2}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + j\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)\right) \quad (9)$$

$$\frac{\partial f}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y}\right) = \frac{1}{2}\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} + j\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)\right) \quad (10)$$

$$+ \partial z \, \partial \bar{z} \, df = \quad (11)$$

In diesem Zusammenhang bezeichnet $\bar{z}$ **die komplex konjugierte Funktion von z. Bei der Wirtinger-Ableitungsmethode können komplexe Funktionen als Erweiterungen reeller Funktionen behandelt werden, indem man sie zur Differenzierung in Real- und Imaginärteile aufteilt. Im Folgenden wird der Aktualisierungsprozess komplexer Variablen unter Verwendung der Schrittweite ÿ/2 und der Verlustfunktion L** beso

Machine Translated by Google

$$xn+1 = xn \ddot{y} \frac{\ddot{y}}{2} \frac{\ddot{y}L}{\ddot{y}x} \quad (12)$$

$$yn+1 = yn \ddot{y} \frac{\ddot{y}}{2} \frac{\ddot{y}L}{\ddot{y}y} \quad (13)$$

Durch Einsetzen von zn+1 = xn+1 + j yn+1 **in die Gleichungen (12) und (13) und anschließendes Einsetzen von die Wirtinger-Ableitung am Ende des Ableitungsprozesses, sieht sie wie folgt aus.**

$$zn+1 = xn \ddot{y} \frac{\ddot{y}}{2} \frac{\ddot{y}L}{} + j\, yn \ddot{y} \ddot{y}x \frac{\ddot{y}}{2} \frac{\ddot{y}L}{\ddot{y}y}$$

$$= xn \ddot{y} \frac{\ddot{y}}{2} \frac{\ddot{y}L}{} + jyn \ddot{y} j \ddot{y}x \frac{\ddot{y}}{2} \frac{\ddot{y}L}{\ddot{y}y}$$

$$= xn + jyn \ddot{y} \frac{\ddot{y}}{2} \frac{\ddot{y}L \ddot{y}L + j \ddot{y}x}{\ddot{y}y}$$

$$\mathbf{1} = zn \ddot{y} \ddot{y} \mathbf{2} - \frac{\ddot{y}L \ddot{y}L + j \ddot{y}x}{\ddot{y}y}$$

$$= zn \ddot{y} \ddot{y} \ddot{y}z \frac{\ddot{y}L}{} \quad (14)$$

Die partielle Ableitung von L in Bezug auf die komplex konjugierte Wirtinger-Ableitung. $\frac{\ddot{y}L}{\ddot{y}z}$ **bezeichnet Dies zeigt, wie die konjugierte Wirtinger- Ableitung verwendet werden kann, um die Aktualisierungsformel für komplexe Variablen zu vereinfachen. Bei Verwendung der obigen Methode ergibt das Umschreiben von Gleichung (5) als Ausgabeschichtfehler Gleichung (15), wobei** tk **den Zielwert und** ok **den tatsächlichen Ausgabewert der Ausgabeschicht k** darstellt .

$$ek = tk \ddot{y} ok \quad (15)$$

Und die Verlustfunktion wird als Gleichung (16) definiert.

$$L = ek^{2} = |tk \ddot{y} ok| \overline{|tk \ddot{y} ok|} \quad (16)$$

Bei großen Gewichten ist es wahrscheinlicher, dass sich der Fehler auf die verborgenen Schichten ausbreitet. Der verborgene Schichtfehler ej **kann durch die folgende Gleichung ausgedrückt werden.**

$$ej = WTjk \cdot ek \quad (17)$$

Mit dem Ergebnis der Gleichung (14) lässt sich **der Gradient der Fehlerfunktion in Bezug auf das Gewicht** Wij **zwischen der Eingabeschicht** i **und der verborgenen Schicht** j **wie folgt ausdrücken. Dabei ist** oi **der Ausgabewert der Eingabeschicht i und** oj **der Ausgabewert der verborgenen Schicht j.**

$$\frac{\ddot{y}L}{\ddot{y}Wij} = \frac{\ddot{y}L}{\ddot{y}oj} \cdot \frac{\ddot{y}oj}{\ddot{y}Wij} \quad (18)$$

$$\frac{\ddot{y}L}{\ddot{y}Wij} \ddot{y}ej \cdot \frac{\ddot{y}}{\ddot{y}} P \ddot{y} Wij \cdot oi = \quad (19)$$

Machine Translated by Google

$$\frac{\ddot{y}L}{} = \ddot{y}ej \cdot P\,\dot{\ddot{y}}Wij \cdot \bar{o}h \quad (20)$$

In den Gleichungen (19) und (20) stellt P **die phasenabhängige Aktivierungsfunktion dar, die in der Ausgabe der verborgenen Schicht verwendet wird,** $\dot{\ddot{y}}$ **ist die Ableitung der abhängigen Aktivierungsfunktion.** und P. **Die Aktualisierung der Gewichte der verborgenen Schicht kann wie folgt ausgedrückt werden.**

$$\text{neues Wij} = \text{altes Wij } \ddot{y}\ lr \cdot \ddot{y}Wij \frac{\ddot{y}L}{} \quad (21)$$

Beim Umgang mit komplexen Variablen muss anstelle von $\ddot{y}Wij$ $\ddot{y}Wij$ **zum Aktualisieren verwendet werden. Dabei** spielt lr $\frac{\ddot{y}L}{}$ **eine Rolle bei der Anpassung der Intensität der Änderung, um ein Überschwingen zu verhindern, und wird als Lernrate bezeichnet. Darüber hinaus ist die Ableitung der Aktivierungsfunktion P** eine **komplexe Funktion, und um die Ableitung einer komplexen Funktion zu finden, wird die Wirtinger-Ableitung verwendet [15].**

Komplexwertiges Softmax: Die Softmax-Funktion, die als Aktivierungsfunktion in der Ausgabeschicht verwendet wird, nimmt Eingabewerte und normalisiert jeden Wert auf einen Wert zwischen 0 und 1. Sie stellt sicher, dass die Summe der normalisierten Werte immer 1 ist. Diese Funktion ermöglicht die Interpretation der Ergebnisse als Wahrscheinlichkeiten. Sie kann als Gleichung (22) ausgedrückt werden.

$$\text{Also f tmax(Z)i} = e\ Zi\ \ddot{y}\ i=1 \frac{e^{Zi}}{N} \quad (22)$$

Die Softmax-Funktion kann Ergebnisse jedoch nur als Wahrscheinlichkeiten für reelle Zahleneingaben interpretieren, und für komplexe Zahleneingaben ist eine andere Interpretationsmethode erforderlich. In diesem Artikel schlagen wir eine komplexwertige Softmax-Interpretationsmethode für komplexe Zahleneingaben vor. Das Ergebnis der Softmax-Funktion für komplexe Zahleneingaben ergibt komplexe Ausgabewerte, und da die Summe aller Ausgabewerte 1+0j ist, schlagen wir einen Algorithmus vor , der den Abstand zwischen den Ausgabewerten und 1 + 0j **in der komplexen Ebene berechnet und den Index des nächstgelegenen Werts auswählt. Dies kann als Gleichung (24) ausgedrückt werden.**

$$\text{distancei} = |\text{So f tmax(Z)i } \ddot{y}\ \ddot{y}| \quad (23)$$

$$\text{nächstgelegener_Index} = \text{argmin(Entfernung i)} \quad (24)$$

Hier ist $\ddot{y}$ = **1** + 0j **definiert. Folglich entscheiden wir uns, die komplexen Zahlen entspricht closest_index als Klasse 1.**

5. Leistungsanalyse In unseren

Experimenten zur Sprachsignalerkennung wurde ein detaillierter Vergleich zwischen den Ergebnissen von realwertigen neuronalen Netzwerken und komplexwertigen neuronalen Netzwerken durchgeführt.
Die Leistungsanalyse wurde parallel zur Vorverarbeitung der komplexwertigen Mel- Frequency-Cepstrum-Koeffizienten (MFCCs) zur Dimensionsreduzierung durchgeführt. Zusätzlich wurde der Einfluss der komplexwertigen MFCCs auf die Ergebnisse der Spracherkennung beobachtet. Insbesondere bei der Erkennung vorgegebener Sprachsignale mithilfe neuronaler Netzwerke wurde der Einfluss der MFCC-Werte und ihr Komplexifizierungseffekt auf die Modellleistung ausgewertet und verglichen, um zu ermitteln, ob eine Leistungssteigerung bei gleichzeitiger Reduzierung der Parameteranzahl möglich ist .

Experimentelle Methode: Abbildung 6 zeigt die Architektur eines mehrschichtigen komplexwertigen neuronalen Netzwerks, das zum Trainieren und Testen mit einem Datensatz aus 8000 kurzen Sprachproben verwendet wird, von denen jede acht unterschiedliche Wörter enthält und weniger als eine Sekunde dauert. Die in dieser Studie verwendeten Sprachdaten sind Teil des von Google gesammelten und unter der CC BY-Lizenz verbreiteten Datensatzes von Sprachbefehlen.

Machine Translated by Google

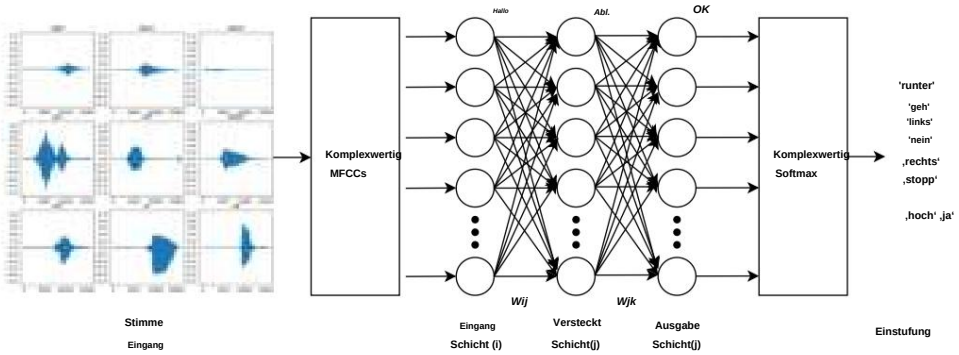


Abbildung 6. Experimenteller Aufbau eines mehrschichtigen CVNN.

Komplexwertige neuronale Netze sind strukturell realwertigen neuronalen Netzen ähnlich, der Unterschied liegt jedoch in der Verwendung komplexer Zahlen, die Phasenänderungen einführen, da die Summe der Produkte zwischen Eingaben und Gewichten in der komplexen Ebene rotiert. Diese Phasenänderungen spielen in komplexwertigen neuronalen Netzen eine wichtige Rolle. Das komplexwertige neuronale Netzwerk besteht aus einer Eingabeschicht mit dimensionsreduzierten komplexwertigen MFCCs, die aus den Sprachdaten durch den in Abbildung 3 gezeigten Komplexifizierungsprozess gewonnen wurden, einer verborgenen Schicht mit 500 Knoten und einer Ausgabeschicht mit 8 Knoten zur Klassifizierung. Alle Knoten und Gewichte verwenden komplexe Zahlen. Die endgültigen Ergebnisse wurden in der Ausgabeschicht klassifiziert. Der Datensatz umfasste 7000 Proben zum Training und 1000 Proben zum Testen. Als Daten wurden Audioclips mit acht Wörtern verwendet und die Experimente wurden durchgeführt, indem die Eingabedaten danach aufgeteilt wurden, ob eine komplexwertige MFCC-Transformation angewendet wurde oder nicht. Im Fall der Verwendung realwertiger MFCCs wurde ein realwertiges neuronales Netzwerk eingesetzt, während ein komplexwertiges neuronales Netzwerk mit komplexwertigen MFCCs verwendet wurde. Sowohl die realwertigen als auch die komplexwertigen neuronalen Netzwerke wurden unter denselben Bedingungen trainiert: lr = 0,01, Batchgröße = 50 und Anzahl der Epochen = 50. Für die Eingabedaten wurden die Ergebnisse von drei verschiedenen Vorverarbeitungsschritten zur Dimensionsreduzierung mit dem realwertigen neuronalen Netzwerk verglichen. Diese Vorverarbeitungsschritte umfassten MFCC- Dimensionen von 20, 10 und 5. Die Länge des Eingangssignals für das neuronale Netzwerk kann wie in Gleichung (25) gezeigt bestimmt werden

$$\text{Die Länge der Eingangssignaldaten} = \text{MFCCs} \times \text{Frames} \quad (25)$$

Beispielsweise wird im Fall des komplexwertigen MFCC 20 zuerst der realwertige MFCC 40 berechnet. Mit Gleichung (25) ergibt sich für den realwertigen MFCC 40 eine Länge des Eingangssignals von 40 × 32 = 1280. Wenn wir diese 40 MFCCs dann in komplexe Zahlen umwandeln, wie in Abbildung 3 dargestellt, wird eine Dimensionsreduzierung auf die Hälfte der Länge erreicht, was zu 16 Frames führt. Folglich wird die Eingangssignallänge 40 × 16 = 640, was **der Eingangssignallänge des realwertigen MFCC 20 entspricht .** Eine ähnliche Dimensionsreduzierung kann **mit derselben Methode für den komplexwertigen MFCC 10 und den komplexwertigen MFCC 5 durchgeführt werden . Die Gesamtzahl der Parameter für das gesamte neuronale Netzwerk, die durch diesen Prozess erhalten wurde, ist in Tabelle 1 aufgeführt. In dieser Tabelle stellt I(n) die Anzahl der Neuronen in der Eingabeschicht dar, H(n) die Anzahl der Neuronen in der verborgenen Schicht und O(n) die Anzahl der Neuronen in der Ausgabeschicht.**

Tabelle 1. Gesamtzahl der Parameter im neuronalen Netzwerk.

MFCC	Realwertiger MFCC	Komplexwertiges MFCC
20	I(20 × 32 = 640) × H(500) + H(500) × O(8) = 324.000	I(40 × 16 = 640) × H(500) + H(500) × O(8) = 324.000
10	I(10 × 32 = 320) × H(500) + H(500) × O(8) = 164.000	I(20 × 16 = 320) × H(500) + H(500) × O(8) = 164.000
5	I(5 × 32 = 160) × H(500) + H(500) × O(8) = 84.000	I(10 × 16 = 160) × H(500) + H(500) × O(8) = 84.000

Machine Translated by Google

Komplexwertiger MFCC 20 vs. realwertiger MFCC 20: Wie in Abbildung 7 dargestellt, erreichte der komplexwertige MFCC 20 eine maximale Trainingsgenauigkeit von 99,40 % und eine Testgenauigkeit von 84,90 %, obwohl er die gleiche Eingabedatenlänge wie der realwertige MFCC 20 hatte, der eine maximale Trainingsgenauigkeit von 98,13 % und eine Testgenauigkeit von 83,10 % erreichte. Die mit komplexwertigen MFCCs erzielten Ergebnisse übertrafen diejenigen mit realwertigen MFCCs sowohl hinsichtlich der Trainings- als auch der Testgenauigkeit.

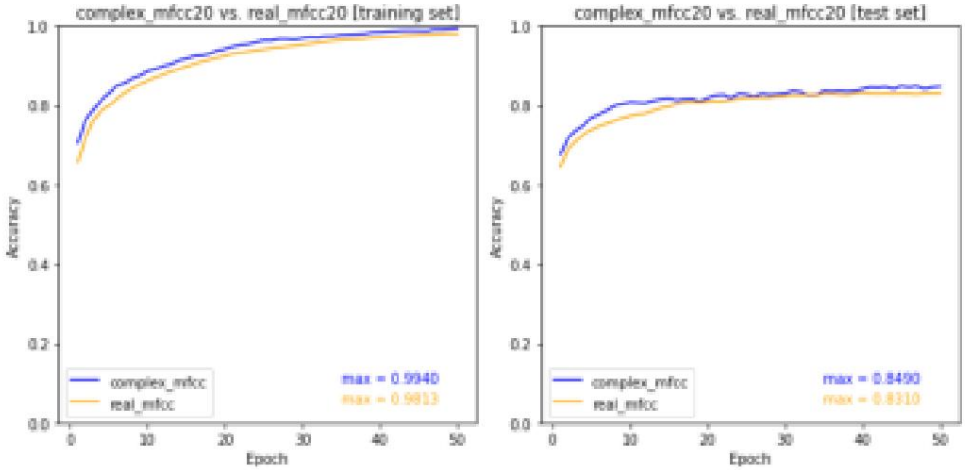


Abbildung 7. Vergleich der Ergebnisse des komplexwertigen MFCC 20 und des realwertigen MFCC 20.

Komplexwertiger MFCC 10 vs. realwertiger MFCC 10: Wie in Abbildung 8 dargestellt, erreichte der komplexwertige MFCC 10 eine maximale Trainingsgenauigkeit von 96,90 % und eine Testgenauigkeit von 83,80 %, obwohl er die gleiche Eingabedatenlänge wie der realwertige MFCC 10 hatte, der eine maximale Trainingsgenauigkeit von 94,24 % und eine Testgenauigkeit von 81,20 % erreichte. Die mit komplexwertigen MFCCs erzielten Ergebnisse übertrafen diejenigen mit realwertigen MFCCs sowohl hinsichtlich der Trainings- als auch der Testgenauigkeit.

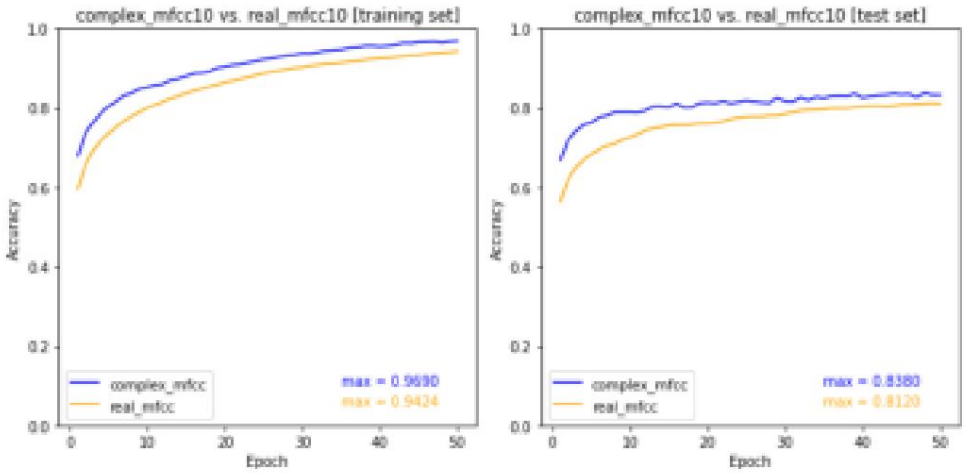


Abbildung 8. Vergleich der Ergebnisse des komplexwertigen MFCC 10 und des realwertigen MFCC 10.

Komplexwertiger MFCC5 vs. realwertiger MFCC5: Wie in Abbildung 9 dargestellt, erreichte der komplexwertige MFCC 5 eine maximale Trainingsgenauigkeit von 92,33 % und eine Testgenauigkeit von 81,40 %, obwohl er die gleiche Eingabedatenlänge hatte wie der realwertige MFCC 5, der eine maximale Trainingsgenauigkeit von 86,34 % und eine Testgenauigkeit von 74,90 % erreichte. Die mit komplexwertigen MFCCs erzielten Ergebnisse übertrafen diejenigen mit realwertigen MFCCs sowohl hinsichtlich der Trainings- als auch der Testger

Machine Translated by Google

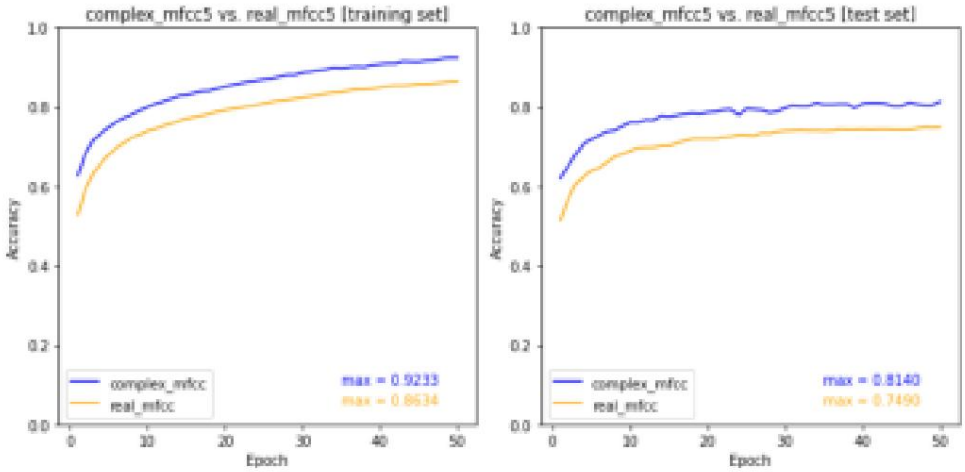


Abbildung 9. Vergleich der Ergebnisse von komplexwertigem MFCC 5 und realwertigem MFCC 5.

Gesamtergebnisse der Analyse: Bei der Überprüfung der experimentellen Ergebnisse wurde festgestellt, dass die Verwendung von MFCCs mit komplexen Werten im Vergleich zur Verwendung von MFCCs mit realen Werten eine höhere Leistung bringt, selbst bei gleicher Eingabedatenlänge. Daher wurde bestätigt, dass die Verwendung von MFCCs mit komplexen Werten und komplexwertigen neuronalen Netzwerken zu besseren Ergebnissen führt . Die Variation der Genauigkeit je nach Anzahl der MFCC-Funktionen ist **in Abbildung 10 dargestellt. Obwohl hier nicht gezeigt, wurden mit dem F1-Score dies**

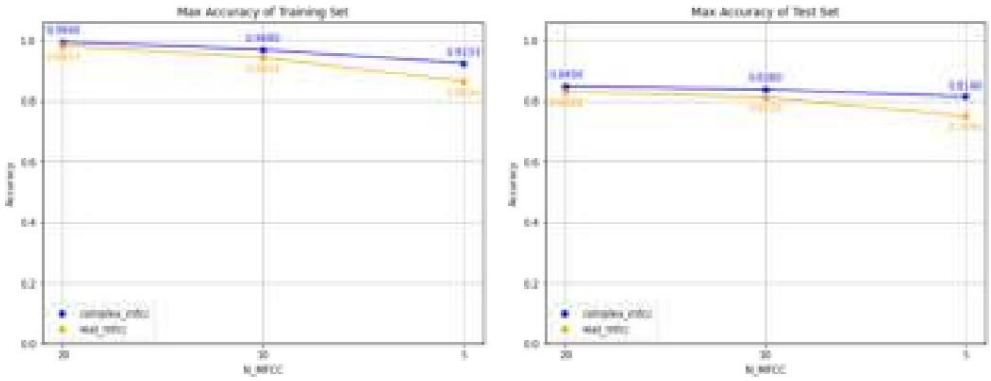


Abbildung 10. Genauigkeitsvariation mit der Anzahl der MFCCs.

Wenn eine ausreichende Anzahl von MFCCs vorhanden ist, scheint sich die Leistung bei Verwendung reeller Zahlen mit zunehmenden Epochen der Leistung bei Verwendung komplexer Zahlen anz Abbildung 9 zeigt jedoch ein anderes Ergebnis. Komplexwertiger MFCC 5 verfügt über 5 dimensionsreduzierte komplexe MFCCs, die durch den in Abbildung 3 gezeigten Komplexifizierungsprozess aus 10 MFCCs gewonnen wurden. Ein realwertiger MFCC 5 mit nur 5 MFCCs kann komplexwertigen MFCC 5 auch mit zunehmenden Epochen nicht einholen. Mit anderen Worten: Wenn die Anzahl der MFCCs gering ist, erreichen die Ergebnisse realwertiger MFCCs selbst mit zunehmenden Epochen nicht die Leistung komplexwertiger MFCCs. Wie in Abbildung 10 gezeigt, sind die Ergebnisse von komplexwertigem MFCC 5 denen von realwertigem MFCC 5 trotz derselben Anzahl von Parametern überlegen. Es wurde also beobachtet, dass bei Verwendung komplexwertiger MFCCs **die Änderung der Genauigkeit mit abnehmender Anzahl von MFCC-Funktionen gleichmäßiger verläuft als bei Verwendung realwertiger MFCCs.**

6. Schlussfolgerungen

In diesem Artikel haben wir komplexwertige neuronale Netzwerke eingeführt, um das Problem der Spracherkennung zu lösen. Unsere Experimente nutzten komplexwertige MFCCs als Eingab

Machine Translated by Google

komplexwertiges neuronales Netzwerk. Die Ergebnisse zeigten eine bessere Leistung beim Einsatz komplexwertiger MFCCs, mit einer allmählicheren Änderung der Genauigkeit bei abnehmender Anzahl der MFCC-Funktionen im Vergleich zur Verwendung realwertiger MFCCs. Dies deutet darauf hin, dass komplexwertige MFCCs dazu beitragen können, die Leistung mit weniger Funktionen aufrechtzuerhalten. Diese Erkenntnisse liefern wertvolle Einblicke in das Potenzial komplexwertiger neuronaler Netzwerke und zeigen die Wirksamkeit der Kombination komplexwertiger MFCCs mit komplexwertigen neuronalen Netzwerken bei der Spracherkennung.

Trotz dieser vielversprechenden Ergebnisse sind wir uns der gewissen Einschränkungen unseres Ansatzes bewusst . Diese Studie wurde mit einem begrenzten Datensatz durchgeführt, der möglicherweise nicht die gesamte Vielfalt realer Sprachmuster abbildet. Darüber hinaus ist die in dieser Studie beschriebene Architektur des komplexwertigen neuronalen Netzwerks relativ einfach und erfasst möglicherweise nicht alle Feinheiten von Sprachsignalen.

Mit Blick auf die Zukunft möchten wir auf dieser Arbeit aufbauen, um die Genauigkeit und Leistung von Sprachsignalverarbeitungs- und -erkennungsmodellen weiter zu verbessern. Unsere zukünftige Forschung wird sich auf die Entwicklung ausgefeilterer komplexwertiger neuronaler Netzwerkmodelle und die Durchführung von Experimenten mit unterschiedlichen Datensätzen zur Verbesserung der Leistung konzentrieren. Darüber hinaus planen wir, die Anwendung komplexwertiger neuronaler Netzwerke in anderen Bereichen als der Spracherkennung zu untersuchen und ihren Einsatz möglicherweise auf verschiedene Be

Beiträge der Autoren: Konzeptualisierung, SK; Projektverwaltung, MP; Überwachung, MP; Schreiben – Originalentwurf, SK. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Arbeit wurde gemeinsam von der National Research Foundation of Korea (NRF) über ein Stipendium der koreanischen Regierung (MSIT) (Stipendium-Nr. NRF-2023R1A2C1005461) und vom MSIT (Ministerium für Wissenschaft und IKT), Korea, im Rahmen des Convergence Security Core Talent Training Business Support Program (IITP-2024-RS-2024-00426853) unterstützt, das vom IITP (Institute of Information & Communications Technology Planning & Evaluation) beaufsichtigt wird .

Datenverfügbarkeitserklärung: Die in der Studie vorgestellten Originalbeiträge sind im Artikel enthalten. Weitere Anfragen können an die entsprechenden Autoren gerichtet werden.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte vorliegen.

Verweise

1. **Tebelskis, J. Spracherkennung mit neuronalen Netzwerken. Dissertation, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA, 1995.**
2. **Sarroff, AM Komplexe neuronale Netzwerke für Audio. Doktorarbeit, Dartmouth College, Hanover, NH, USA, 2018.**
3. **Abdul, ZK; Al-Talabani, AK Mel Frequency Cepstral Coefficient und seine Anwendungen: Eine Übersicht.** IEEE Access **2022,** 10, **122136–122158. [CrossRef]**
4. **Barrachina, JA; Ren, C.; Vieillard, G.; Morisseau, C.; Ovarlez, JP Theorie und Implementierung komplexwertiger neuronaler Netzwerke.** arXiv **2023, arXiv:2302.08286.**
5. **Aizenberg, I.** Komplexwertige neuronale Netze mit mehrwertigen Neuronen; **Springer: Berlin/Heidelberg, Deutschland, 2016.**
6. **Graves, A.; Jaitly, N.; Mohamed, AR Hybride Spracherkennung mit Deep Bidirectional LSTM. In Proceedings des 2013 IEEE Workshop on Automatic Speech Recognition and Understanding, Olomouc, Tschechische Republik, 8.–12. Dezember 2013; S. 273–278. [Querverweis]**
7. **Paz, P.; Garrido, M. Effiziente Implementierung komplexer Multiplikatoren auf FPGAs mit DSP-Slices.** J. Signal Process. Syst. **2023,** 95, **543–550. [CrossRef]**
8. **Ahmad, M.; Zhang, L.; Chowdhury, MEH FPGA-Implementierung eines komplexwertigen neuronalen Netzwerks für polar dargestellte Bildklassifizierung.** Sensors **2024,** 23, **2627. [CrossRef] [PubMed]**
9. **Anggraeni, D.; Sanjaya, WS; Nurasyidiek, MY; Munawwaroh, M. Die Implementierung der Spracherkennung mithilfe von Mel- Frequency-Cepstrum-Koeffizienten (MFCC) und der auf Python basierenden Support Vector Machine (SVM)-Methode zur Steuerung eines Roboterarms.** IOP Conf. Ser. Mater. Sci. Eng. **2018,** 288, **012042. [CrossRef]**
10. **Dhanjal, AS; Singh, W.** Eine umfassende Untersuchung zur automatischen Spracherkennung mithilfe neuronaler Netzwerke; **Springer Nature: Berlin/ Heidelberg, Deutschland, 2023.**
11. **Smith, SW** The Scientist and Engineer's Guide to Digital Signal Processing, **2. Auflage; California Technical Publishing: San Diego, CA,**

^ "Die Macht der Freiheit". In: Die Macht der Freiheit. Berlin 1999, S. 557–558.

12. **Bassey, J.; Li, X.; Qian, L. Eine Untersuchung komplexwertiger neuronaler Netzwerke.** arXiv **2021, arXiv:2101.12249v1.**
13. **MYONeuralNet. Komplexwertige neuronale Netzwerke – Experimente. 2016. Online verfügbar: http://makeyourownneuralnetwork. blogspot.com/2016/05/complex-valued-neural-networks.html (abgerufen am 19. Februar 2021).**

Machine Translated by Google

14. **Fischer, R. Wirtinger Calculus. In** Precoding and Signal Shaping for Digital Transmission; **John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2002; S. 405–413.**
15. **PyTorch. Automatisches Differenzierungspaket – torch.autograd, nd Online verfügbar: https://pytorch.org/docs/stable/notes/autograd.html (abgerufen am 3. November 2023).**
16. **Lee, C.; Hasegawa, H.; Gao, S. Komplexwertige neuronale Netzwerke: Eine umfassende Untersuchung.** IEEE/CAA J. Autom. Sin. **2022, 9, 1433–1454. [CrossRef]**

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Job-Empfehlungen: Benchmarking von Collaborative Filtering Methoden für Kleinanzeigen

Robert Kwieci ´nski 1,2, Tomasz Górecki 2,*, Agata Filipowska 3 und Viacheslav Dubrov 1

1 OLX Group, Królowej Jadwigi 43, 61-872 Posen, Polen; r.kwiecinskipl@gmail.com (RK); slavadubrov@gmail.com (VD)
2 Fakultät für Mathematik und Informatik, Adam-Mickiewicz-Universität, Uniwersytetu Pozna ´nskiego 4, 61-614 Pozna ´n, Polen; Institut
3 für Informationssysteme, Wirtschaftsuniversität Pozna ´n, Al. Niepodlegÿo´sci 10, 61-875 Pozna ´n, Polen; agata.filipowska@ue.poznan.pl
* Korrespondenz: tomasz.gorecki@amu.edu.pl

Zusammenfassung: Kleinanzeigen stellen zahlreiche Herausforderungen für Empfehlungsmethoden dar, darunter die temporäre Sichtbarkeit von Anzeigen, die Anonymität der meisten Benutzer und die Tatsache, dass normalerweise nur ein Benutzer einen beworbenen Artikel konsumieren kann. In dieser Arbeit begegnen wir diesen Herausforderungen, indem wir Modelle und Bewertungsverfahren auswählen, die als genau, vielfältig und effizient gelten (in Bezug auf Speicher und Zeitverbrauch während des Trainings und der Vorhersage). Ziel dieses Artikels ist es, verschiedene Empfehlungsmethoden für Stellenanzeigen am Beispiel von OLX Jobs zu vergleichen, um die Konversionsrate von Anzeigen und die Benutzerzufriedenheit zu verbessern. In unserer Forschung implementieren wir skalierbare Methoden und stellen verschiedene Ansätze für die Empfehlungen dar: Alternating Least Square (ALS), LightFM, Prod2Vec, RP3Beta und Sparse Linear Methods (SLIM). Wir führten A/B-Tests durch, indem wir Millionen von Nachrichten mit Empfehlungen versendeten, um Online-Bewertungen ausgewählter Methoden durchzuführen. Darüber hinaus haben wir den für unsere Forschung erstellten Datensatz veröffentlicht. Nach bestem Wissen ist dies der erste Datensatz dieser Art. Es enthält 65.502.201 Ereignisse, die auf OLX Jobs von 3.295.942 Benutzern durchgeführt wurden, die im Jahr 2020 innerhalb von zwei Wochen mit 185.395 Stellenanzeigen interagierten (sie anzeigten, beantworteten oder mit Lesezeichen versahen). Wir zeigen, dass RP3Beta, SLIM und ALS bei Tests in einer Laborumgebung deutlich besser abschneiden als Prod2Vec und LightFM. Online-A/B-Tests zeigen auch, dass das Senden von Nachrichten mit Empfehlungen, die von den ALS- und RP3Beta-Modellen generiert werden, die A

Schlüsselwörter: Stellenempfehlungen; Kleinanzeigen; kollaboratives Filtern; Skalierbarkeitstests; A/B-Tests; Empfehlungssysteme; Matrixfaktorisierung; graphenbasierte Empfehlungssysteme; OLX- Datensatz zu Stelleninteraktionen; Diversitätsmetriken

check for updates

Zitat: Kwieci ´nski, R.; Górecki, T.; Filipowska, A.; Dubrov, V. Job Empfehlungen: Benchmarking von kollaborativen Filtermethoden für Kleinanzeigen. Elektronik **2024**, 13, **3049**.

https://doi.org/10.3390/Elektronik13153049

Wissenschaftliche Herausgeber: Marco Lombardi, Domenico Santaniello, Zilu Liang und Muhammad Khan

Empfangen: 30. Mai 2024
Überarbeitet: 24. Juli 2024
Akzeptiert: 27. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Einleitung

Online-Kleinanzeigen sind Websites, auf denen Inserenten Anzeigen zum Verkauf oder zur Vermietung von Dienstleistungen oder Produkten veröffentlichen **[1]. In diesem Beitrag wird Online-Werbung am Beispiel von OLX Jobs behandelt.**

Kleinanzeigen zielen darauf ab, sowohl die Bedürfnisse von Werbetreibenden als auch von Online-Werbungsempfängern zu erfüllen . Um diese Gruppen zu erreichen, werden bessere Empfehlungsmethoden immer wichtiger. Im Vergleich zum traditionellen E-Commerce (z. B. Amazon (https://www.amazon.com/, abgerufen am 26. Juni 2024) oder eBay (https://www.ebay.com/, abgerufen am 26. Juni 2024)) stehen Kleinanzeigen vor verschiedenen Herausforderungen, die wir in unserem Bewertungsverfahren berücksichtigen. Eine davon ist die zeitliche Sichtbarkeit von Anzeigen, die wir durch häufiges Neutrainieren Daher ist die Effizienz des Modells, über die wir berichten, für die Bewertung von entscheidender Bedeutung. Eine weitere Herausforderung besteht darin, dass ein Werbetreibender normalerweise nur einen Benutzer benötigt, um die Transaktion a Diesem Aspekt haben wir Rechnung getragen, indem wir bei unserer Bewertung auch eine Diversitätsanalyse einbezogen haben.

Machine Translated by Google

In unserer Forschung konzentrieren wir uns auf Ansätze, die sich aus der kollaborativen Filterung ergeben. Kollaborative Filtermodelle haben ihre Genauigkeit bewiesen, obwohl sie keine zusätzlichen Informationen über Benutzer oder Artikel verwenden, die in der Praxis manchmal schwer zu erhalten oder zu verarbeiten sind. Ziel dieses Artikels ist es, verschiedene Empfehlungsmethoden für den Einsatz im Kleinanzeigensektor zu bewerten. Die hohe Spärlichkeit unseres Datensatzes sowie die beträchtliche Anzahl von Anzeigen und Benutzern tragen zur Komplexität des Forschungsproblems bei.

Nachfolgend listen wir unsere wichtigsten Beiträge auf.

- Wir bewerten in einer Laborumgebung umfassend verschiedene, skalierbare und kollaborative Filteransätze hinsichtlich Genauigkeit, Vielfalt und Effizienz. Wir bewerten die Genauigkeit für verschiedene Benutzergruppen, abhängig von der Anzahl der Elemente, mit denen sie interagiert haben, zusammen mit einer statistischen Analyse der Unterschiede. Die vorgeschlagene Bewertungsmethode befasst sich mit den wichtigsten Herausforderungen von Kleinanzeigen.
- Wir präsentieren die Ergebnisse zweier Online-A/B-Tests, die die Auswirkungen der ALS- und RP3Beta-Modelle auf OLX-Benutzer demonstrieren. An jedem Test nahmen mehr als 1 Million Benutzer teil. Die berichteten Auswirkungen können verwendet werden, um die Bedeutung von Empfehlungssystemen im Kleinanzeigenbereich abzuschätzen und anderen Unternehmen zu ermöglichen, fundiertere Entscheidungen über die Entwicklung solcher Modelle zu treffen.
- Wir veröffentlichen und präsentieren einen Datensatz zu Jobinterak öffentlich verfügbarer Datensatz zu Interaktionen am Arbeitsplatz.

Dieses Dokument besteht aus sieben Abschnitten. Abschnitt 2 ist ein Literaturüberblick über die zugrunde liegenden Entwicklungen. Abschnitt 3 beschreibt den Datensatz, der entwickelt und auf Kaggle öffentlich zugänglich gemacht wurde. Das Design des Laborexperiments sowie die Implementierung der Methoden werden in Abschnitt 4 beschrieben. Abschnitt 5 konzentriert sich auf die Ergebnisse der Laborauswertung unserer Implementierungen sowie die statistische Bedeutung der Ergebnisse. Abschnitt 6 präsentiert die Ergebnisse der Online-Implementierungen, die während A/B-Tests mit Benutzern durchgeführt wurden, sowie eine Diskussion der Ergebnisse. Schlussfolgerungen werden in Abschnitt 7 präsentiert.

2. Verwandte Arbeiten

In diesem Abschnitt beschreiben wir zunächst die besonderen Merkmale der Bereiche Kleinanzeigen und Stellenangebote . Anschließend geben wir einen Überblick über Empfehlungssysteme. Als Nächstes diskutieren wir die Faktoren, die unsere Modellauswahl beeinflussen, und geben eine detaillierte Beschreibung jedes Modells. Abschließend geben wir die Forschungslücken an.

2.1. Kleinanzeigen-Websites

Kleinanzeigen-Websites (Classifieds) sind Online-Versionen von Zeitungsanzeigenteilen mit Anzeigen zu Fahrzeugen, Immobilien, Stellenangeboten, Haustieren usw. Kleinanzeigen sind eine besondere Art von E-Commerce-Website, da sie nicht den Abschluss von Transaktionen zwischen Parteien unterstützen – sie werden stattdessen verwendet, um über das Produkt oder die Dienstleistung zu kommunizieren und Treffen zu vereinbaren, während die Transaktion offline durchgeführt wird; dies unterscheidet sie von traditionellen E-Commerce-Websites wie eBay oder Amazon [2].

Unsere Forschung konzentriert sich auf OLX, das in über 30 Ländern weltweit vertreten ist. Einer der Anzeigentypen, die in Kleinanzeigen veröffentlicht werden, sind Stellenangebote. In unserer Untersuchung haben wir mehrere Merkmale identifiziert, die Stellenanzeigen in Kleinanzeigen beispielsweise von Auto- oder Immobilienanzeigen unterscheiden:

- Die Anforderungen in einer Anzeige geben den Benutzertyp an, der auf die Anzeige reagieren kann – nicht alle Benutzer verfügen über die erforderlichen Kompetenzen.
- Intensivere Einbindung des Nutzers während des Suchvorgangs, da die Suche nach einem Arbeitsplatz ein vorrangiges Bedürfnis darstellt, verglichen mit beispielsweise dem Kauf eines Autos.
- Der Arbeitsort hat erhebliche Auswirkungen auf den Nutzer – um ein Mobiltelefon oder ein Auto zu kaufen, muss der Nutzer möglicherweise reisen oder einen Paketdienst für die Lieferung des Artikels nutzen. Dennoch erfordert ein Arbeitsplatz einen Umzug, wenn er weit vom Wohnort des Nutzers entfernt ist.
- Aus technischer Sicht weist eine Stellenbeschreibung viele qualitative Aspekte auf und je nach Position können die Unterschiede selbst zwischen Anzeigen desselben Unternehmens erheblich sein.

Machine Translated by Google

Zusätzlich zu diesen Merkmalen haben wir während unserer Untersuchungen auch die folgenden Herausforderungen für Empfehlungsmethoden im Kleinanzeigenbereich festgestellt, die unsere Implementierung der Methoden beeinflusst haben:

- **Die Anzahl der Benutzer und Angebote: Bei OLX Jobs hatten wir es mit Millionen von Benutzern und Zehntausenden von Angeboten zu tun, die sich beide ständig ändern.**
- **Ein Benutzer muss kein Profil erstellen, um mit Anzeigen zu interagieren. Dies bedeutet, dass Informationen aus einem Profil nicht für Empfehlungen verwendet werden können und Empfehlungen durch das Benutzerverhalten beim Anzeigen bestimmt werden.**
- **Eine Anzeige hat normalerweise eine begrenzte Anzahl von Besuchern (die mit der Anzeige interagieren), da es sich bei einer Anzeige normalerweise um ein einmaliges Angebot handelt. Nachdem der Bedarf gedeckt wurde, wird die Anzeige deaktiviert.**
- **Es gibt keine Informationen darüber, ob eine Transaktion stattgefunden hat; bei Kleinanzeigen bezieht sich die Konvertierung auf den Erhalt einer Antwort auf eine veröffentlichte Anzeige, da jede Transak**

Diese Herausforderungen erschweren nicht nur die Implementierung, sondern auch die Prüfung der Genauigkeit der vorgeschlagenen Methoden. Wir haben diese Herausforderungen bei der Vorbereitung des Datensatzes sowie bei der Implementierung und Bewertung der in diesem Dokument beschriebenen Methoden berücksi

2.2. Empfehlungssysteme

Es hat sich gezeigt, dass das Problem der Überauswahl [3] den Konsum von Produkten verringern **[4] und die Einnahmen der Site reduzieren** kann **[5]. Anbieter können ihre Benutzer nach zusätzlichen Informationen über ihre Erwartungen in Bezug auf Artikel fragen, um die Anzahl der angebotenen Auswahlmöglichkeiten zu reduzieren. Selbst wenn der Benutzer bereit ist, Zeit für die Bereitstellung dieser Informationen aufzuwenden, ist die Anzahl der Möglichkeiten normalerweise zu groß, um sie dem Benutzer anzuzeigen. Um dieses Problem zu lösen, haben viele Unternehmen wie Netflix [6] und Amazon [7] personalisierte Empfehlungssysteme entwickelt, deren Aufgabe es ist, Benutzern relevante Artikel v**

Empfehlungen werden in vielen Bereichen verwendet, darunter Musik, Tourismus, Mode und Lebensmittel. Zahlreiche Studien wurden in den Bereichen E-Commerce [9], Kleinanzeigen [10] und Stellenempfehlungen [11–13] durchgeführt. Li et al. [14] verbesserten ihre Empfehlungen auf LinkedIn erheblich, indem sie Deep Transfer Learning nutzten, um domänenspezifische Modelle zum Verständnis von Stellen zu erstellen. Zhu et al. [15] schlugen ein domänenübergreifendes Empfehlungsmodell vor, das einen heterogenen Graphen und einen Community-Erkennungsalgorithmus verwendet, um Kurse und Stellen zu empfehlen. Lacic et al. [16] bewerteten verschiedene Autoencoder-Architekturen für sitzungsbasierte Stellenempfehlungen und präsentierten die Ergebnisse anhand beliebter Datensätze zu Stellene

Typischerweise unterscheiden wir zwei Ansätze zur Bereitstellung von Empfehlungen: inhaltsbasierte Empfehlungen und kollaboratives Filtern [8].

Bei inhaltsbasierten Empfehlungen [17] werden Informationen über Nutzer (z. B. Hobby, Ausbildung, Fähigkeiten) und Artikel (z. B. Preis, Standort, Kategorie) berücksichtigt. Diese Systeme nutzen in der Regel die Interaktionshistorie (Bewertungen, Besuche, Käufe und Antworten) eines einzelnen Nutzers, berücksichtigen aber nicht die Bewertungen aller Nutzer gleichzeitig.

Das Sammeln/Verarbeiten von Informationen über Benutzer und Artikel ist nicht immer machbar. Beim kollaborativen Filteransatz [18,19] verwenden wir nur Informationen über Benutzerinteraktionen zwischen Artikeln. Auch ohne Informationen über Benutzermerkmale und Artikelmerkmale können diese Systeme sehr genaue Empfehlungen liefern. Wir haben uns für diesen Ansatz entschieden, weil Benutzermerkmale fehlen und es schwierig ist, kollaboratives Filtern durch die Verwendung von Artikelmerkmalen in einer Produktionsumgebung zu übertreffen (wir gehen näher darauf in Abschnitt 6.3 ein).

Einer der größten Nachteile des kollaborativen Filterns ist die Unmöglichkeit, Empfehlungen zu Benutzern und Elementen ohne Interaktionen bereitzustellen, das sogenannte Kaltstartproblem [20]. Eine mögliche Lösung zur Reduzierung dieses Problems ist das häufige Neutrainieren des Modells. Bei vielen Benutzerinteraktionen erfordert diese Strategie einen Fokus auf die Skalierbarkeit des Modells [21].

Ein weiterer wichtiger Faktor, der die Wahl des Modells bestimmt, ist die Art der Interaktion zwischen einem Benutzer und einem Artikel [20]. Wir nennen das Feedback explizit, wenn der Benutzer seine Präferenz für den Artikel ausdrücklich angibt (z. B. indem er ihn auf einer Skala von 1 bis 5 bewertet). Ein anderer Typ ist implizites Feedback, wenn wir die Präferenz des Benutzers für den Artikel nur anhand von

Machine Translated by Google

auf das Nutzerverhalten (z. B. beim Kauf eines Artikels oder beim Ansehen eines Films). Implizites Feedback ist weniger genau, häufiger und drückt in der Regel nur positive Präferenzen aus [22].

Laborbewertungen von Empfehlungssystemen basieren üblicherweise auf Genauigkeitsmetriken [23]. In letzter Zeit wurde viel Aufmerksamkeit auf andere Aspekte wie Vielfalt, Unsicherheit, Neuartigkeit und Reichweite gelegt [24,25]. Bei Kleinanzeigen kann ein Artikel normalerweise nur von einem Benutzer konsumiert werden, daher sollten wir es vermeiden, denselben Artikel einer signifikanten Anzahl von Benutzern zu empfehlen. Daher ist es wichtig, sich auf die Katalogabdeckung zu konzentrieren [26] . Eine der schwierigsten Aufgaben bei der Bewertung von Empfehlungssystemen besteht darin, die Offline-Metrik zu finden, die am stärksten mit dem Online-Ziel korreliert [27].

In diesem Artikel vergleichen wir kollaborative Filtermodelle anhand eines umfangreichen impliziten Feedback-Datensatzes. Wir vergleichen die Genauigkeit, Vielfalt und Effizienz ausgewählter Empfehlungsmethoden in einer Laborumgebung und beweisen anschließend die Qualität der Methoden in Onlinetests mit Benutzern.

2.3. Empfehlungsmethoden

Nachfolgend führen wir die wichtigsten Aspekte auf, die die Methodenauswahl für unsere Forschung beeinflusst haben. • Popularität in der Literatur. Wir haben Methoden ausgewählt, deren Leistung gründlich evaluiert wurde. Alle ausgewählten Methoden außer einer wurden mindestens hundertmal zitiert .

• Beliebtheit in der Branche. Die Bewertung dieses Aspekts war schwierig, da Unternehmen die Ergebnisse durchgeführter Experimente selten veröffentlichen. Einige Beliebtheitsindikatoren sind die Verfügbarkeit der Implementierung und die Anzahl der Benutzer (z. B. widergespiegelt in der Anzahl der Sterne und Forks auf GitHub). Wir haben auch die Erfahrung der in unserem Unternehmen tätigen Datenwissenschaftler berücksichtigt , die die meisten der ausgewählten Methoden **in der Vergangenheit für einige andere Anwendungsfälle eingesetzt haben .**

- **Skalierbarkeit. Wir haben nur Lösungen in Betracht gezogen, die Millionen von Benutzern und Elementen verarbeiten können . Darüber hinaus haben wir das Kaltstartproblem durch häufiges Neutrainieren des Modells reduziert. Daher mussten wir Techniken auswählen, die innerhalb weniger Stunden an unserem Datensatz trainiert werden konnten.**

• Vielfalt der Ansätze. Wir wollten Methoden verschiedener „Familien“ testen, um ihre Anwendbarkeit auf Kleinanzeigen zu beurteilen. Wir wollten die vielversprechendste Familie von Ansätzen identifizieren, um die Forschung in dieser Richtung fortzusetzen.

Schließlich wurden ALS, LightFM, Prod2Vec, RP3Beta und SLIM zum Testen und Evaluieren ausgewählt . In den folgenden Unterabschnitten werden die Gründe für die Auswahl der einzelnen Methoden und die Methodendefinition beschrieben. Eine gemeinsame Einschränkung dieser Methoden ist ihre Unfähigkeit, Elemente zu empfehlen, die von keinem Benutzer besucht wurden, und Empfehlungen für Benutzer zu generieren, die kein Element besucht haben (Kaltstartproblem, [21]). Mit unserem Datensatz ohne Benutzer- und Elementmerkmale ist es unmöglich, dieses Problem durch eine personalisierte Methode zu überwinden.

Ausgewählte Methoden wurden bereits in verschiedenen Bereichen verglichen. Anelli et al. [28] verglichen RP3Beta, SLIM, ALS und BPRMF (das bei Verwendung **der BPR-Verlustfunktion LightFM entspricht)** anhand von drei Datensätzen: Amazon Digital Music, Epinions und MovieLens **1M hinsichtlich Genauigkeit, Vielfalt und Effizienz. Die Anzahl der Interaktionen in diesen Datensätzen beträgt 145.523, 300.475 bzw. 571.531. Unsere Studie verwendete einen Datensatz mit mehreren zehn Millionen Interaktionen, ähnlich denen in OLX-Produktionssystemen.**

Zusätzlich stellen wir Ihnen die Ergebnisse einer Online-Auswertung zur Verfügung.

Dacrema et al. [29] verglichen die Genauigkeit von SLIM, RP3Beta und ALS anhand mehrerer Datensätze (CiteULike-a, MovieLens 1M, Pinterest, Amazon Music, Amazon Movies, MovieLens 20M, Netflix, Yelp und Gowalla). Sie zeigten, dass jeder dieser Algorithmen die anderen beiden bei mindestens einem Datensatz übertraf. Einige ihrer Ergebnisse wurden von Anelli et al. **[30] anhand der Datensätze MovieLens 1M und Pinterest** weiter bestätigt. **Zhu et al. [31] verglichen BPRMF, SLIM und Prod2Vec anhand dreier Datensätze (Amazon Books, Yelp und Gowalla) und zeigten die Überlegenheit von SLIM. Dong et al. [32] analysierten die Genauigkeit von RP3Beta, SLIM und ALS anhand von neun Datensätzen, wobei SLIM bei acht Datensätzen die besten Ergebnisse lieferte und ALS**

Machine Translated by Google

Im Gegensatz zu unserer Studie verglichen diese Arbeiten die ausgewählten Methoden weder in Bezug auf Diversität oder Effizienz und führte auch keine Online-Evaluation durch.

In letzter Zeit wurden mehrere Empfehlungsmethoden auf Basis neuronaler Netze entwickelt. vorgeschlagen. Es hat sich gezeigt, dass ausreichend bestimmte und abgestimmte einfache Modelle oft übertreffen sie [28,31–33]. Aufgrund der Komplexität der Berechnungen sind Modelle, die auf auf neuronalen Netzen werden seltener an Datensätzen getestet, die mehrere Millionen Interaktionen [28]. Daher überlassen wir es zukünftigen Forschungen, fortgeschrittenere und skalierbare Methoden, die für unseren Datensatz bessere Ergebnisse liefern.

2.3.1. Matrixfaktorisierungsmodelle

Matrixfaktorisierungstechniken werden seit dem Netflix-Preiswettbewerb [34] erfolgreich in Empfehlungssystemen eingesetzt. Die Hauptidee besteht darin, die sp Benutzer-Element-Interaktionsmatrix als Produkt zweier kleinerer, dichter Matrizen. Wir evaluieren zwei Matrixfaktorisierungsansätze: LightFM und ALS.

Das LightFM -Modell [35] wurde vorgeschlagen, um das Kaltstartproblem von Matrixfaktorisierungsmethoden zu überwinden . In unserem Fall reduziert es sich auf die klassische Matrixfaktor Informationen über Benutzer und Artikel. Trotzdem haben wir uns für diesen Ansatz entschieden, weil seine effiziente offizielle Umsetzung bei der Unterstützung mehrerer Verlustfunktionen: Logistik, BPR [36], WARP [37] und -OS WARP [38].

Sei U eine Menge von Benutzern und I eine Menge von Artikeln. Für einen gegebenen Benutzer ÿ U und Artikel ÿ I define = 1, wenn der Benutzer mit dem Element , und = 0, andernfalls. Das LightFM-Modell interagiert, wird die Punktzahl wie folgt vorhergesagt:

$$\hat{} = \cdot + + \quad ,$$

wobei und die -dimensionalen latenten Repräsentationen des Benutzers und des Elements bezeichnen und und die Vorurteile des Benutzers und des Elements sind. Alle diese Parameter werden wä Das Training.

Die logistische Verlustfunktion sollte verwendet werden, wenn sowohl positives als auch negatives Feedback Typen sind verfügbar, daher ist es für unseren Datensatz ungeeignet, was wir während der Hyperparameter-Optimierung experimentell bestätigt haben.

Die Verlustfunktionen BPR, WARP und -OS WARP sind paarweise Lern-zu-Rang-Ansätze , die für unsere Top-Empfehlungen-Aufgabe relevanter sind. Sie versuchen zu modifizieren Modellparameter zur Maximierung der Differenz zwischen den Punktzahlen der Items ÿ D mit mit denen der Benutzer interagiert und mit den Elementen ÿ D, mit denen der Benutzer nicht interagiert.

Konkret versuchen wir in der BPR-Verlustfunktion [36] den Ausdruck zu maximieren, wie folgt:

$$\sum_{(,,)\ \ddot{y}S} \ln(\hat{} \quad \hat{}) \ \ddot{y} \ \ddot{y} \ ||\ddot{y}||2 \ ,$$

Wo

$$S = \{(,,) | \ \ddot{y} \ U, \ \ddot{y} \ D \text{ und } \ddot{y} \ D\},$$

$$() = \frac{1}{1 + \ddot{y}},$$

ÿ ist ein modellspezifischer Regularisierungsparameter, ÿ ist ein Parametervektor und || · || ist der Euklidische Norm.

Wir präsentieren WARP [37] als einen Fall von -OS WARP [38]. Die -OS WARP-Verlustfunktion ist eine Summe der Verluste für jeden Benutzer:

$$\sum_{\ddot{y}U} (\hat{} \ , D),$$

wobei ^ ein Vektor der geschätzten Punktzahlen^ für jedes Element ÿ I ist.

Machine Translated by Google

Für einen gegebenen , Wir ordnen die interagierenden Elemente D in absteigender Reihenfolge der Benutzer werden die Punkte geschätzt. Wir ordnen die Elemente nämlich als 1, 2, . . . , | D | , wobei die Indizes sind so gewählt, dass wir folgendes haben:

$$\hat{} \quad \hat{} \quad \hat{} \quad \hat{}$$
$$1 \ \ddot{y} \ 2 \ \ddot{y} \ldots \ddot{y} \ |D| .$$

Dann haben wir Folgendes:

$$(\hat{}, D) = \frac{1}{\ddot{y}\ddot{y}} \sum_{=1}^{|\quad|} (\ )\ddot{y}(\text{Rang}(\hat{})),$$

wobei Rang (ˆ) = ÿl\D (1 + ( ) ˆ ÿ ˆ), ist eine Indikatorfunktion, ÿ( ) = $\sum_{=1} \frac{1}{}$, und = $\sum_{=1}^{|D|}$ normalisiert die durch induzierten Gewichte.

In der Praxis wählen wir ein positives Item bezüglich der Gewichtungsfunktion aus und Wählen Sie nach dem Zufallsprinzip unbekannte Elemente aus, bis nach den Proben 1 + ÿ . Dann führen wir eine Gradientenschritt zur Minimierung von ÿ(ÿ ÿ) (1 ÿ + ˆ).

Schließlich wird der -OS WARP-Verlust für ( ) = 1 und ( ) = 0 erreicht, wenn ÿ . Der WARP Verlust wird erreicht, wenn wir ( ) = 1 für ÿ N setzen. Wir können die Bestellung positiver Elemente aus D um die Berechnungen zu vereinfachen.

In unserer Forschung wurden alle Verlustfunktionen (Logistik, BPR, WARP und -OS WARP) behandelt als Hyperparameter des LightFM-Modells und während des Trainings optimiert.

Das ALS-Modell (alternating least squares), auch bekannt als WRMF – weighted regularized matrix factorization [39], ist ein Beispiel für eine Matrixfaktorisierungsmethode, die für implizite Feedback-Datensätze. Wir haben diese Methode aufgrund ihrer nachgewiesenen Leistung gewählt, Skalierbarkeit und Popularität in Forschung und Industrie. Darüber hinaus hat das ALS-Modell wurde bereits in OLX implementiert und verwendet.

Wir definieren die Konfidenz einer Beobachtung als = 1 + , wobei ein Hyperparameter ist . Wir definieren die Präferenz auch als = (> 0). Dann optimieren wir in ALS die folgender Ausdruck:

$$\text{Mindest}_{\ddot{y},\ddot{y}} \ddot{y}\ddot{y} \quad (\ddot{y} \quad \ddot{y})^2 + (\ddot{y}\ddot{y} \,||\,||2 + \ddot{y}\ddot{y} \,||\,||2),$$

wobei die -dimensionale Einbettung des Benutzers ist, die -dimensionale Einbettung des Elements ist, und ist ein Regularisierungsparameter. Der Name des Modells bezieht sich auf die optimierte Methode für diese Funktion. Das Verfahren der alternierenden kleinsten Quadrate geht davon aus, dass Benutzer oder Artikel latent Faktoren sind festgelegt und bestimmen die anderen analytisch. Das Verfahren wird effizient wiederholt einige Male [39,40].

Bei Matrixfaktorisierungsansätzen wird die Einbettung für jeden Benutzer separat gelernt, Dies kann bei einer großen Benutzeranzahl zu einem hohen Zeit- und Speicheraufwand führen. Darüber hinaus könnten die Empfehlungen nicht erklärbar sein, da zwei Benutzer mit dem Die gleichen Interaktionen können unterschiedliche Empfehlungen erhalten. Andere betrachtete Modelle (SLIM, RP3Beta, Prod2Vec) bieten die gleichen Empfehlungen für Benutzer mit dem dieselben Interaktionen.

2.3.2. Nachbarschaftsbasierte Modelle

Das SLIM-Modell (sparse linear methods for Top-N recommender systems [41]) ist basierend auf der itembasierten Nearest Neighbors Regression [8]. Wir haben diese Methode gewählt, weil sie Vorteile gegenüber itembasierten Nearest-Neighbors- und Matrixfaktorisierungsmodellen [41] und gute Leistung gegenüber nicht-neuralen und neuronalen Ansätzen [29].

In diesem Modell approximieren wir die Benutzer-Artikel-Matrix durch das Lernen der Artikel-Artikel-spärliche Ähnlichkeitsmatrix . Explizit wird durch Minimieren des folgenden Ausdrucks gelernt:

$$\frac{1}{} || \ddot{y} \ 2 \quad ||2 + \frac{}{|| \ 2} \quad ||2 + || \ ||1,$$

Machine Translated by Google

wobei ÿ 0 und = 0 für alle , , 1 ÿ , ÿ ||, || · || ist die Matrix-Frobenius-Norm:

$$\| \| = \ddot{y}\ddot{y} \ddot{y}\ddot{y} \sum_{=1}^{\|ch|} \sum_{=1}^{\|ch|} {}^{2},$$

|| · ||1 ist die eintragsweise ÿ1-Norm:

$$\| \|1 = \ddot{y}\ddot{y} \ddot{y}\ddot{y} \sum_{=1}^{\|ch|} \sum_{=1}^{\|ch|} | |$$

und und sind Regularisierungsparameter.

Wir können sehen, dass jede Spalte unabhängig voneinander gelernt werden kann, was die Skalierbarkeit dieses Modells [41].

Eine Einschränkung dieser Methode besteht darin, dass sie keine Benutzer- und Artikeleinbettungen erzeugt, die für andere Aufgaben hilfreich sein könnten (z. B. Clustering von Benutzern oder Artikeln). Wir können die Spalten der Matrix als Artikeleinbettungen behandeln, aber es handelt sich um spärliche ||-dimensionale Vektoren, sodass sie wahrscheinlich nicht direkt für andere Aufgaben verwendet werden können.

2.3.3. Graphenbasierte Ansätze

RP3Beta [42] ist ein graphenbasierter Ansatz für kollaboratives Filtern. Ähnlich wie Dacrema et al. [29] präsentieren wir dieses Modell als Erweiterung des P3alpha-Modells [43]. Wir betrachten einen bipartiten und ungerichteten Graphen = (V, E), wobei jeder Knoten ÿ V entweder einen Benutzer oder ein Element darstellt und jede Kante { , } ÿ E die Interaktion zwischen einem gegebenen Benutzer und einem gegebenen Element darstellt. Lassen Sie die Adjazenzmatrix und d

Anschließend werden die Ergebnisse von P3alpha in der Matrix gespeichert.

$$((\quad)^{\ddot{y}1} \quad)^{\ddot{y}} \quad)^{3},$$

wobei ÿ die elementweise Potenz einer Matrix bezeichnet (Hadamard-Potenz). Genauer gesagt wird **der Score ˆ am Schnittpunkt der Zeile, die den Benutzer darstellt, und der Spalte,** die das **Element darstellt, gespeichert. Wir können sehen, dass dieser Score eine Summe der Scores ist, die den Pfaden der Länge 3 zugewiesen sind, die Benutzer und Element im Graphen verbinden.**

In RP3Beta wird die Punktzahl mit P3alpha berechnet und durch die Popularität der Elemente hoch 1 geteilt, um den Popularitätsfehler zu mildern.

Wir haben uns aufgrund seiner Einfachheit, Skalierbarkeit und nachgewiesenen Leistung für dieses Modell entschieden **[29,30], insbesondere angesichts der hohen Spärlichkeit unseres Datensatzes, die den Speicherbedarf beim Trainieren des Modells erheblich verringert. Dies ermöglicht es uns, die Artikel-Artikel-Ähnlichkeitsmatrix direkt zu berechnen und im Speicher zu speichern, anstatt die im Originalpapier vorgeschlagene Random-Walk-Approximation zu verwenden.**

RP3Beta generiert, ähnlich wie SLIM, keine Benutzer- und Artikeleinbettungen. Eine weitere damit verbundene Einschränkung ist die Unfähigkeit, Artikel zu empfehlen, deren Distanz auf dem bipartiten Benutzer-Artikel- Diagramm größer als drei von einem bestimmten Benutzer ist. Insbesondere kann dem Benutzer ein Artikel nur empfohlen werden, wenn mindestens ein Benutzer ÿ den Artikel besucht hat und mindestens einer der Artikel vom Benutzer besucht wurde. Darüber hinaus verfügt das RP3Beta-Modell nicht über Parameter, die während des Trainings optimiert werden; daher ist es möglicherweise nicht in der Lage, die komplexen Beziehungen i

2.3.4. Prod2Vec

Das Prod2Vec-Modell [44] (oder Item2vec [45]) basiert auf dem Word2Vec-Modell [46], das in der Verarbeitung natürlicher Sprache weit verbreitet ist. Das Modell konzentriert sich auf das Lernen von Vektordarstellungen von Elementen unter Verwendung einer Sequenz von Elementen als „Satz“ und Elementen innerhalb dieser Sequenz als „Wörter“. Diese Methode wurde gewählt, weil sie skalierbar ist und sich stark von den verglichenen Methoden unterscheidet. Darüber hinaus basiert dieser Ansatz auf dem aktuellen hochmodernen hybriden Deep Neural Network-Empfehlungssystem, das in anderen Kategori

Machine Translated by Google

Prod2Vec verwendet das Skip-Gram-Modell durch Maximierung der Zielfunktion über den Satz S von Itemsequenzen, wie folgt definiert:

$$= \ddot{y}\ddot{y} \quad \ddot{y}\ddot{y} \quad \ddot{y}\ddot{y} \quad \log P(+ |),$$

wobei die Kontextlänge für Elementsequenzen ist und Elemente aus derselben Sequenz beliebig angeordnet. Die Wahrscheinlichkeit P(+ |) , ein benachbartes Element + zu beobachten , **wenn Das aktuelle Element wird mithilfe der Soft-Max-Funktion definiert:**

$$P(+ |) = \frac{\exp(\ddot{y} \; \ddot{y} \;)}{\sum_{=1}^{|ch|} \exp(\ddot{y} \; \ddot{y} \;)},$$

wobei und die Eingabe- und Ausgabevektordarstellungen von item sind und || der , Anzahl eindeutiger Elemente im Vokabular. Prod2Vec modelliert den Kontext einer Elementsequenz, wo Elemente mit ähnlichem Kontext (d. h. mit ähnlichen benachbarten Interaktionen) ähnliche Vektordarstellungen.

Bevor wir Empfehlungen generieren, normalisieren wir alle Eingabevektordarstellungen. was die Effizienz verbessert. Um Empfehlungen zu generieren, stellen wir den Benutzer als Durchschnitt der Eingabevektordarstellungen der interagierenden Elemente. Dann suchen wir nach Elementen deren Eingabedarstellungen die höchste Kosinusähnlichkeit mit der Benutzerdarstellung aufweisen (-nächste Nachbarn im latenten Raum).

Eine mögliche Einschränkung dieser Methode ist eine strikte Abhängigkeit zwischen Benutzer und Artikel Darstellungen. Einige Interaktionen sollten möglicherweise einen größeren Einfluss auf Benutzerdarstellungen haben als andere (was in Matrixfaktorisierungsmodellen möglich ist).

2.4. Forschungslücke

Ziel dieses Artikels ist es, verschiedene Empfehlungsmethoden im Bereich Stellenanzeigen zu vergleichen und eine spezifische Forschungslücke zu schließen. Nach bestem Wiss Die angesprochenen Empfehlungsmethoden wurden noch nie in großem Maßstab getestet oder Maßstab bei der Online-Arbeit im Jobbereich. In dieser Arbeit vergleichen wir diese Methoden in einem Laborumgebung unter Verwendung des von uns veröffentlichten Datensatzes und in Online-A/B-Tests, die in einem Produktionsumfeld.

3. OLX-Datensatz zu Jobinteraktionen

Diese Forschung liefert einen kommerziellen Datensatz, der 65.502.201 durchgeführte Veranstaltungen enthält auf OLX Jobs von 3.295.942 Benutzern, die im Jahr 2020 in zwei Wochen mit 185.395 Stellenanzeigen interagierten.

Wir haben den Datensatz auf Kaggle veröffentlicht (https://www.kaggle.com/olxdatascience/olx-jobs-interactions, abgerufen am 26. Juni 2024).

Jede Zeile des Datensatzes stellt das Ereignis dar, das ein bestimmter Benutzer in Bezug auf ein bestimmtes Anzeige zu einem bestimmten Zeitpunkt (Tabelle 1). Die folgenden Ereignisse im Zusammenhang mit dem Nutzerverhalten wurden berücksichtigt:

- **Klick: Der Benutzer hat die Artikeldetailseite besucht;**
- **Lesezeichen: Der Benutzer hat das Element zu den Lesezeichen hinzugefügt;**
- **chat_click: Der Benutzer hat den Chat geöffnet, um den Besitzer des Artikels zu kontaktieren;**
- **contact_phone_click_1: Der Benutzer hat die dem Artikel beigefügte Telefonnummer angezeigt;**
- **contact_phone_click_2: Der Benutzer hat geklickt, um den Besitzer des Artikels anzurufen;**
- **contact_phone_click_3: Der Benutzer hat geklickt, um dem Besitzer des Artikels eine SMS-Nachricht zu senden;**
- **contact_partner_click: Der Benutzer hat geklickt, um auf die externe Seite des Artikelbesitzers zuzugreifen;**
- **contact_chat: Der Benutzer hat dem Besitzer des Artikels eine Nachricht gesendet.**

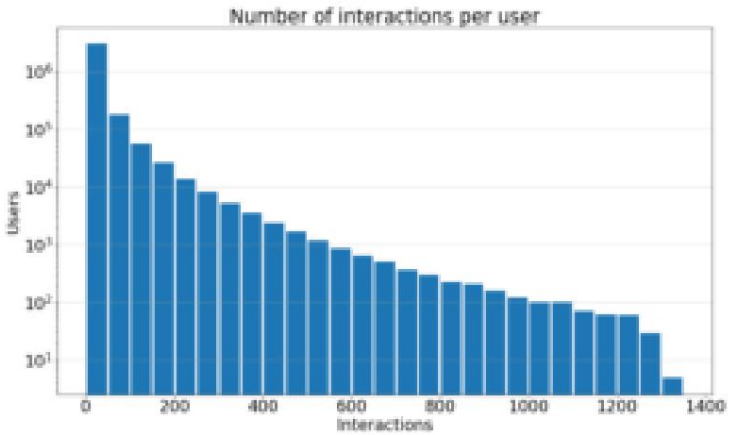
Machine Translated by Google

Tabelle 1. Beispielzeilen aus dem Datensatz.

Benutzer	Artikel	Ereignis	Zeitstempel
1745587	168661		1582216025
843008	62838		1582485868
14285	30469	klicken	1582247367
1142944	80122		1581805847
2835659	23728	klicken Lesezeichen klicken chat_click	1582247836

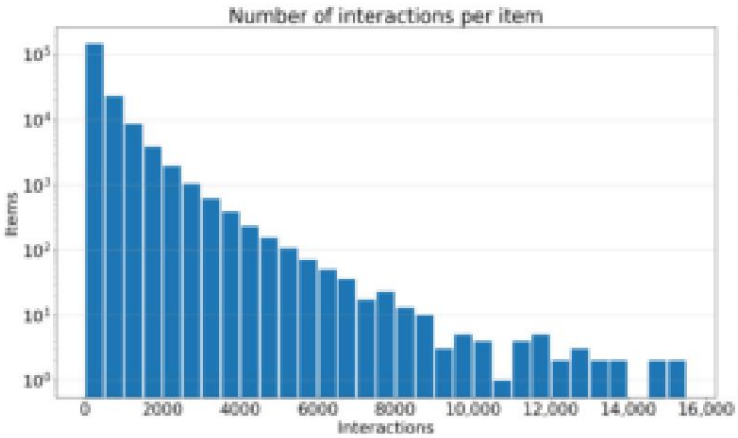
Quelle: eigene Studie.

Die Datensparsity (% fehlender Einträge in der Benutzer-Element-Matrix) beträgt 99,9923 %, Das ist zwar sehr niedrig, aber typisch für Kleinanzeigen (zum Vergleich: Die Datenknappheit auf der Der MovieLens 20M-Datensatz ist zu 99,47 % dicht und damit fast 70-mal dichter. Verteilungen von Interaktionen pro Benutzer, Interaktionen pro Artikel und Ereignisse sind in den Abbildungen 1 und 2 aufgeführt. und Tabelle 2. Wir haben eine logarithmische Skala verwendet, da die Verteilungen stark rechtsschief.



Quantile	Value
0.01	1
0.10	1
0.25	2
0.50	6
0.75	18
0.90	46
0.99	221

Abbildung 1. Anzahl der Interaktionen pro Benutzer. Quelle: eigene Studie.



Quantile	Value
0.01	1
0.10	1
0.25	3
0.50	137
0.75	428
0.90	950
0.99	2944

Abbildung 2. Anzahl der Interaktionen pro Artikel. Quelle: eigene Studie.

Laut De Ruijt und Bhulai [47] sind die beliebtesten Datensätze für die Bewertung von Job-Empfehlungssystemen CareerBuilder12 (https://www.kaggle.com/Wettbewerbe/Job-Empfehlungen, abgerufen am 16. Juni 2024), RecSys16 [12] und Rec-Sys17 [13]. Tabelle 3 zeigt die Statistiken der betrachteten Datensätze. Die Anzahl der eindeutigen Interaktionen ist die Anzahl der eindeutigen Benutzer-Element-Paare, bei denen der Benutzer mit dem Obwohl die Forscher die Datensätze RecSys16 und RecSys17 verwenden, waren sie im Rahmen von Data Science-Wettbewerben unter streng begrenzten Bedingungen veröffentlicht und sind keine nicht mehr verfügbar. Daher konnten wir die Anzahl der einzigartigen Interaktionen für diese

Machine Translated by Google

Datensätze. Wir haben festgestellt, dass der OLX-Datensatz für Jobinteraktionen die höchste Anzahl von Interaktionen und Typen zwischen den betrachteten Datensätzen.

Tabelle 2. Verteilung der Ereignisse.

Aktion	Häufigkeit (%)
	89.794
Klicken Sie auf	2.628
contact_phone_click_1 Lesezeichen	2.511
chat_click	2.136
contact_chat	1.448
contact_partner_click	0,701
contact_phone_click_2	0,679
contact_phone_click_3 Quelle:	0,103

eigene Untersuchung.

Tabelle 3. Vergleich zwischen Datensätzen mit Stellenempfehlungen. Der Buchstabe gibt die Standardabweichung an.

	Jobs bei OLX Interaktionen	CareerBuilder12	RecSys16	RecSys17
Interaktionen	65,50 M	1,60 M	8,83 M	8,27 M
Einzigartige Interaktionen	47,17 M	1,60 M	-	-
Benutzer	3,30 M	0,3 2 Mio.	1,37 Mio.	1,50 Mio.
Durchschnittliche Anzahl einzigartiger Interaktionen pro Benutzer	14.31 (= 29,23)	4,99 (= 11,42)	<6,46	<5,51
Artikel	0,19 M	0,37 Mio.	1,36 Mio.	1,31 Mio.
Durchschnittliche Anzahl einzigartiger Interaktionen pro Artikel	254,42 (= 426,02)	4.38 (= 8,19)	<6,50	<6,31
Dichte der Interaktionsmatrix (%)	0,0077 %	0,0014 %	< 0,0005 %	< 0,0004 %
Arten von Interaktionen	8	1	4	5
Zeitstempel	ÿ	ÿ	ÿ	ÿ
Benutzerfunktionen	ÿ	ÿ	ÿ	ÿ
Artikelmerkmale	ÿ	ÿ	ÿ	ÿ

Quelle: eigene Studie.

Es ist zu beachten, dass die Wahrung der Vertraulichkeit von Anzeigen und Benutzern Priorität hatte bei der Erstellung dieses Datensatzes. Zu den Maßnahmen zum Schutz der Privatsphäre gehörten die folgenden:

- **Ursprüngliche Benutzer- und Artikelkennungen wurden durch eindeutige, zufällige Ganzzahlen ersetzt;**
- **Jedem Zeitstempel wurde eine nicht offengelegte konstante Ganzzahl hinzugefügt;**
- **Einige Teile der Interaktionen wurden herausgefiltert;**
- **Einige zusätzliche künstliche Interaktionen wurden hinzugefügt.**

4. Versuchsaufbau

4.1. Umsetzung

Die oben genannten Methoden haben wir auf Basis früherer Entwicklungen implementiert. Details dazu, wie sich unsere Implementierungen von den zuvor getesteten Ansätzen unterscheiden, finden Sie Die Umsetzung der Methoden zur Reproduzierbarkeit der Ergebnisse ist verfügbar auf GitHub (https://github.com/rob-kwiec/olx-jobs-recommendations, aufgerufen am 29. Mai 2024).

Machine Translated by Google

Tabelle 4. Für die Forschung neu implementierte Methoden.

Methode	Methodenfamilie	Quelle	Unterschied zum Ursprüngliche Methode
LightFM	Matrixfaktorisierung	Implementierung basierend auf: [48]. Begleitpapier: [35].	Keiner
ALS	Matrixfaktorisierung	Implementierung basierend auf: [49]. Unterstützende Papiere: [39,40].	Keiner
SCHLANK	Nachbarschaftsbezogen	Quellen, die unsere Implementierung inspiriert haben: [50,51]. Begleitpapier: [41].	Keiner
RP3Beta	Graphenbasiert	Quelle, die unsere Implementierung inspiriert hat: [52]. Unterstützende Papiere: [29,53].	Direkt durchgeführt Berechnungen auf dünn besetzten Matrizen statt auf zufälligen Spaziergänge Näherung.
Prod2Vec	Word2Vec	Die Implementierung basiert auf: [54]. Unterstützende Papiere: [44,45].	Sequenzen von Interaktionen, sortiert nach einem Zeitstempel. Darstellung des Benutzers als Durchschnitt der Darstellungen interagierter Elemente. Wir haben auch mit CBOW (Continuous Bag of Words) experimentiert.

Quelle: eigene Studie.

4.2. Ziele der Laborforschung und des Online-Tests **Unser oberstes**

Geschäftsziel war es, die Zahl der Benutzer zu erhöhen, die sich über OLX Jobs auf Stellen bewerben. Um dies zu erreichen, versuchen wir, ein Ranking-Problem zu lösen. Genauer gesagt empfehlen wir zehn Elemente mit der höchsten Wahrscheinlichkeit, eine Interaktion eines bestimmten Benutzers zu erhalten. Das Problem, eine Offline-Metrik zu finden, die am stärksten mit der Bewerbungsrate im Jobbereich korreliert, wurde von Mogenet et al. [27] untersucht. Da unser Ziel ähnlich ist, folgten wir ihren Ergebnissen, indem wir auf Präzision@10 optimierten (Präzision für die ersten zehn empfohlenen Elemente [23]). Wir berichten und diskutieren jedoch auch über andere Genauigkeits-, Diversitäts- und Effizienzmetriken.

Wir konzentrieren uns auf Benutzer, die aktiv nach Jobs suchen, indem sie mit Stellenanzeigen interagieren. Daher berücksichtigen wir keine Benutzer ohne Interaktionen (Kaltstart-Benutzer).

4.3. Datenaufteilung

Der in Abschnitt **3 beschriebene Datensatz** wurde **gemäß der Forschungsmethodik in Trainings-Validierungs- und Testdatensätze aufgeteilt .** Der Testsatz enthält 20 % der neuesten Interaktio **Dies bedeutet, dass von 14 Tagen etwa 3 Tage in den Testdatensatz einbezogen wurden (siehe Abbildung** 3). Diese Art der Aufteilung des Datensatzes in Trainings- und Testdatensätze ist **laut Meng et al. [55] die realistischste** Vorgehensweise, **die verschiedene Aufteilungsstrategien zur Bewertung von Empfehlungssystemen untersucht hat.**

In beiden Datensätzen haben wir die Anzahl der Interaktionen auf einmalige Interaktionen zwischen einem Benutzer und einem Element beschränkt. Wenn ein Benutzer mehr als einmal mit einem Element interagierte, wurde nur die erste Interaktion gezählt und der Zeitstempel dieser Interaktion dem Element zugeordnet. Wir haben auch nicht zwischen den Interaktionsarten unterschieden. Darüber hinaus wurden zwei weitere Änderungen vorgenommen: Nur Benutzer, die im Trainings- und Validierungsdatensatz erschienen, wurden in den Testsatz aufgenommen (weil wir anderen Benutzern ohne zusätzliches Wissen über sie keine personalisierten Empfehlungen geben konnten), und wir haben alle Benutzer-Element-Paare herausgefiltert, die im Trainingssatz vorhanden waren (um zu vermeiden, Elemente zu empfehlen, die bereits gesehen wurden). Als Ergebnis hatten wir ungefähr 38 Millionen Zeilen im Trainings- und Validierungsdatensatz und ungefähr 6 Millionen Zeilen im Testsatz. Weitere Statistiken zu diesen beiden Datensätzen finden Sie in Tabelle 5.

Machine Translated by Google

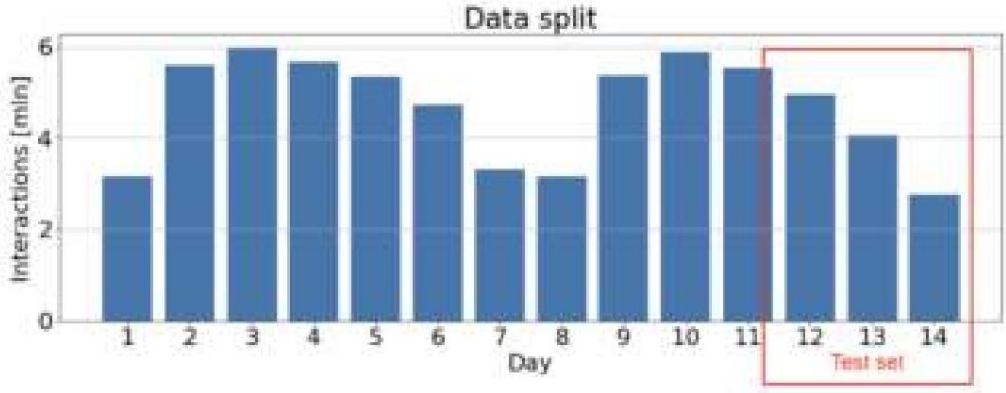


Abbildung 3. Vorbereitung des Datensatzes (aufgeteilt in Training–Validierung und Testen, abhängig von der Anzahl der Tage). Quelle: eigene Studie.

Tabelle 5. Vergleich zwischen Trainings-/Validierungs- und Testsätzen. Der Buchstabe bezeichnet Standard Abweichung.

	OLX-Jobinteraktionen Interaktionen	Trainings- und Validierungsset	Test-Set
Interaktionen	65,50 M	52,40 Mio.	6,11 Mio.
Einzigartige Interaktionen	47,17 M	38,25 Mio.	6,11 Mio.
Benutzer	3,30 M	2,83 Mio.	0,62 Mio.
Durchschnittliche Anzahl einzigartiger Interaktionen pro Benutzer	14.31 (= 29,23)	13,50 (= 26,11)	9,87 (= 14,38)
Artikel	0,19 M	0,18 Mio.	0,13 Mio.
Durchschnittliche Anzahl einzigartiger Interaktionen pro Artikel	254,42 (= 426,02)	214,05 (= 365,93)	47,51 (= 83,84)
Dichte der Interaktionsmatrix (%)	0,0077 %	0,0076 %	0,0077 %

Quelle: eigene Studie.

4.4. Modelloptimierung

Aus Kosten- und Zeitgründen sollten Sie bei der Optimierung der Parameter
Wir haben eine Beschränkung auf 20 % der Benutzer und 20 % der Elemente der Schulung und Validierung vorgenommen
Dann teilen wir den eingeschränkten Datensatz in den Trainings-Validierungs-Datensatz mit dem Regeln, die im obigen Abschnitt beschrieben wurden. Für jedes Modell wurden 100 Iterationen ausgewählt
durch Bayessche Optimierung unter Verwendung von Gaußschen Prozessen, implementiert in scikit-learn (https://scikit-optimize.github.io/stable/modules/generated/skopt.gp_minimize.html, aufgerufen am 29. Mai 2024). Die Konfigurationsdatei definiert Bereiche für Hyperparameter und ist verfügbar in unserem Code-Repository (https://github.com/rob-kwiec/olx-jobs-recommendations/blob/main/src/tuning/config.py, abgerufen am 29. Mai 2024). Für RP3Beta haben wir beispielsweise beschränkten uns auf die Suche nach und im Intervall [0, 2].

Empfehlungen im Optimierungsprozess wurden für 30.000 Benutzer erstellt und Wie oben erwähnt, wurde als Optimierungswert precision@10 verwendet. Abbildung 4 zeigt der Präzisionswert in Abhängigkeit von den Parametern. Es ist ersichtlich, dass ALS nicht sehr optimiert, aber für andere Methoden war diese Optimierung entscheidend für Verbesserung der Präzision. Hyperparameter für alle Modelle sind in Tabelle 6 dargestellt.

Machine Translated by Google

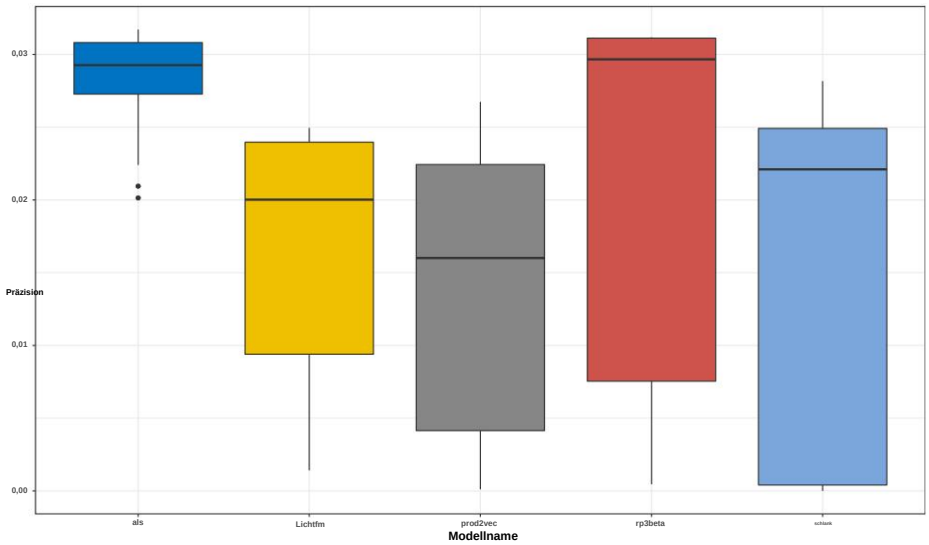


Abbildung 4. Präzision für jedes Modell, abhängig von den Parametern. Quelle: eigene Studie.

Tabelle 6. Modell-Hyperparameter.

Modell	Modellieren von Hyperparametern
als	{'Faktoren': 357, 'Regularisierung': 0,001, 'Iterationen': 20, 'Ereignisgewichtsmultiplikator': 63}
Lichtfm	{'no_components': 512, 'k': 3, 'n': 20, 'learning_schedule': 'adadelta', 'Verlust': 'Warp', 'max_sampled': 61, 'Epochen': 11}
prod2vec	{'Vektorgröße': 168, 'Alpha': 0,028728, 'Fenster': 20, 'Mindestanzahl': 16, „Beispiel“: 0,002690026, „min_alpha“: 0,0, „sg“: 1, „hs“: 1, „negativ “: 200, „ns_exponent“: ÿ0,16447846705441527, „cbow_mean“: 0, 'Epochen': 22}
RP3Beta	{'alpha': 0,61447198, 'beta': 0,1443548}
schlank	{'Alpha': 0,00181289, 'l1_ratio': 0,0, 'Iterationen': 3}

Quelle: eigene Studie.

4.5. Modellbewertung

Wir bewerten unsere Modelle hinsichtlich Genauigkeit, Vielfalt und Effizienz.

Die Genauigkeit der Empfehlungsmethode bei impliziten Feedback-Datensätzen ist typischerweise ausgewertet, indem untersucht wird, welche Elemente aus der Liste der Top-Empfehlungen der Benutzer im Testset interagiert [8]. Viele Bewertungsmetriken bestimmen die Metriken für jedes Benutzer separat und geben den Mittelwert als endgültige Bewertungsnote der Methode an [23]. Die Empfehlungssystem kann als Klassifikator betrachtet werden, bei dem die Interaktionen des Benutzers in Die Testreihe enthält positive Beispiele und die Top-Empfehlungen sind Elemente, die als positiv durch das Modell. Daher können die zur Bewertung von Klassifikatoren verwendeten Methoden direkt auf Empfehlungssysteme angewendet. Beispielsweise ist der Wert von precision@ für einen bestimmten Benutzer die Anzahl der empfohlenen Elemente, mit denen der Benutzer im Testset interagiert hat, geteilt durch die Anzahl der diesem Benutzer empfohlenen Artikel. Weitere in dieser Arbeit diskutierte Metriken waren beschrieben von Tamm et al. [23].

Im Kleinanzeigenbereich, vor allem im Stellenbereich, wird in der Regel nur ein Benutzer benötigt, um eine Transaktion bezüglich einer bestimmten Anzeige abschließen. In diesem Fall die Vielfalt der Empfehlungen ist noch kritischer als in anderen Bereichen (wie Musik, Filme oder andere E-Commerce, ohne Kleinanzeigen). Um diesen Aspekt zu bewerten, verwenden wir drei Maßnahmen: Testabdeckung, Shannon Entropie [56] und der Gini-Index [56]. Die Testabdeckung bezieht sich auf den Anteil der Testitems mindestens einem Benutzer empfohlen. Mit Shannon-Entropie und dem Gini-Index ignorieren wir Elemente außerhalb des Testsatzes. Größere Testabdeckung und Shannon-Entropiewerte deuten auf höhere Vielfalt, während höhere Werte des Gini-Index eine geringere Vielfalt anzeigen.

Machine Translated by Google

Der letzte Schritt der Laborbewertung konzentriert sich auf die Effizienz. Zu diesem Zweck verwendeten wir eine AWS SageMaker ml.m5.4xlarge-Instanz mit 64 GB RAM und 16 vCPUs, mit einer Intel(R) Xeon(R) Platinum 8259CL CPU @ 2,50 GHz, läuft auf Amazon Linux AMI 2018.03. Alle Methoden wurden hinsichtlich dreier Funktionen evaluiert: Datenvorverarbeitung, Modellanpassung und Artikelempfehlung. Die Ausführungszeit wurde berechnet für die Gesamtzeit, die zur Berechnung des Modells benötigt wurde (in Stunden). Der Speicherpeak wurde berechnet in GB und die maximale Speichernutzung wurde beim Ausführen einer bestimmten Funktion berücksichtigt. Diese Auswertung wurde mit der Bibliothek tracemalloc durchgeführt (https://docs.python.org/3/bibliothek/tracemalloc.html, abgerufen am 29. Mai 2024).

5. Laborauswertung

In diesem Abschnitt wird die Laborauswertung der im entwickelten Datensatz diskutierten Methoden beschrieben. Wir haben die im vorherigen Abschnitt ermittelten Hyperparameter mit der besten Leistung verwendet. Jedes Modell wurde mit dem gesamten Trainings- und Validierungsdatensatz trainiert und Empfehlungen für alle 619.389 Benutzer aus dem Testdatensatz erstellt. Um die Werte von Für jede Offline-Metrik verglichen wir die beschriebenen Methoden mit einfachen, nicht personalisierten Baselines: der beliebteste Produktansatz und zufällige Empfehlungen.

5.1. Richtigkeit

Alle Metriken in diesem Abschnitt berücksichtigen die ersten zehn Empfehlungen, z. B. precision@10, recall@10. Höhere Metrikwerte weisen auf eine höhere Qualität der Empfehlungen hin. Die Werte sollte nicht direkt mit den Ergebnissen aus anderen Datensätzen verglichen werden, da Die Metriken hängen stark von der Verteilung des Datensatzes ab (z. B. hohe Spärlichkeit) und die Train/Test-Splitting-Strategie [55].

5.1.1. Bewertung der Genauigkeit

Tabelle 7 stellt die Ergebnisse der Evaluierung dar. Die meisten der beobachteten Unterschiede sind statistisch signifikant aufgrund der großen Anzahl ausgewerteter Benutzer. Wir gehen näher darauf ein die statistische Signifikanz der Ergebnisse in Abschnitt 5.1.2.

Tabelle 7. Genauigkeit: Ergebnisse der Evaluierung. Alle vorgestellten Metriken wurden von Tamm et al. [23] beschrieben. Für Für die mAP-Metrik haben wir die Variante mit = min(,) verwendet.

Metric	RP3Beta	SLIM	ALS	Prod2Vec	LightFM	Beliebteste	Zufällig
Präzision	0,0484	0,0472	0,0434	0,0368	0,0736	0,0359	0,0012	0,00006
Rückruf	0,0783	NDC 0,0657	0,0580	0,0721	0,0657	0,0564	0,0012	0,00005
0,0759	MAP 0,0393	0,0567	0,0365	0,0329	0,0282	0,0545	0,0016	0,00007
MRR	0,1365	LAUC 0,1314	0,1230	0,1065	0,5368	0,0264	0,0006	0,00002
0,5391	HR 0,3131	0,5328	0. 5289	0,3066	0,2878	0,1034	0,0038	0,00019
		0,2537				0,5281	0,5006	0,49999
						0,2547	0,0112	0,00059

Quelle: eigene Studie.

Unsere Ansätze übertreffen die zufälligen und beliebtesten Artikelempfehlungen deutlich . Die schlechte Leistung des beliebtesten Ansatzes resultiert aus der Natur von Kleinanzeigen, vor allem die Domäne des Jobs, wo die Benutzer meist an einem bestimmten Standort und Kategorie. Aufbau ausgefeilterer, personalisierter Empfehlungssysteme ist in einem solchen Szenario sinnvoll.

Es wird gezeigt, dass RP3Beta andere Ansätze in Bezug auf Genauigkeit übertrifft Metriken. Wir glauben, dass dies mit der hohen Spärlichkeit unseres Datensatzes zusammenhängt. RP3Beta berechnet die Empfehlungen deterministisch (durch Nutzung von Pfaden der Länge drei auf einer Benutzer-Element-bipartiter Graph). Alle anderen Ansätze nutzen maschinelle Lerntechniken, um Benutzer zu finden Darstellungen (ALS, LightFM), Artikeldarstellungen (ALS, LightFM, Prod2Vec) oder direkte Artikel-Artikel-Ähnlichkeiten (SLIM), die in Fällen auftreten können, in denen Benutzer oder Artikel eine geringe Anzahl der Interaktionen.

Machine Translated by Google

5.1.2. Statistischer Vergleich für Präzision@10

Wir präsentieren einen detaillierten statistischen Vergleich für unsere Hauptmetrik, precision@10, um Unterschiede zwischen den Methoden zu identifizieren. Zunächst testen wir die Nullhypothese, dass alle Methoden gleich abschneiden und die beobachteten Unterschiede lediglich zufällig sind (Omnibus- Test). Der Friedman-Test [57,58] mit der **Erweiterung von Iman und Davenport [59]** ist wahrscheinlich **der beliebteste Omnibus-Test und ist normalerweise eine gute Wahl, wenn mehr als fünf verschiedene Algorithmen** verglichen werden **[60,61]. Sei der Rang der -ten Methode auf der -ten Stichprobe und**

$$= \frac{1}{} \ddot{y}\ddot{y}_{=1} \quad .$$

In unserem Fall ist die Anzahl der Stichproben gleich der Anzahl der Benutzer im Testsatz = 619.389.

Der Test vergleicht die mittleren Ränge der Methoden und basiert auf den folgenden Statistiken:

$$= \frac{(\ddot{y} 1)^{2}}{(\ddot{y} 1) \ddot{y}^{2}},$$

Wo

$$^{2} = \frac{12}{(+ 1)} \ddot{y}\ddot{y}_{=1} {}^{2} \ddot{y} 3 (+ 1)$$

ist die Friedman-Statistik, die die Verteilung mit ÿ 1 und (ÿ 1) (ÿ 1) Freiheitsgraden hat . **Der -Wert aus diesem Test ist gleich 0. Der erhaltene -Wert zeigt an, dass wir die Nullhypothese, dass alle Algorithmen gleich abschneiden, sicher ablehnen können. Daher** können wir mit den Post-hoc-Tests fortfahren, um signifikante paarweise Unterschiede zwischen allen **Methoden festzustellen .** Demšar **[62] schlägt die Verwendung des Nemenyi-Tests [63] vor, der alle Algorithmen paarweise vergleicht . Für ein Signifikanzniveau bestimmt der Test den kritischen Unterschied (CD). Wenn der Unterschied zwischen den durchschnittlichen Ranglisten zweier Algorithmen größer ist als**

$$\mathbf{CD} = \ddot{y}\ddot{y} \frac{(+ 1)}{6},$$

Die Nullhypothese, dass die Algorithmen die gleiche Leistung aufweisen, wird abgelehnt (auf der basiert studentisierten Spannweitenstatistik geteilt durch ÿ 2). Demšar [62] schlug ein Diagramm vor, um die Unterschiede zu überprüfen (das CD-Diagramm visuell). Algorithmen, die nicht durch eine Linie verbunden sind, können im Diagramm als unterschiedlich betrachtet werden.

In unserem Fall, mit einer Signifikanz von = 0,05, sind zwei beliebige Algorithmen mit einem Unterschied in der Ein Durchschnittsrang über 0,0114 wird als ungleich angesehen (Abbildung 5).

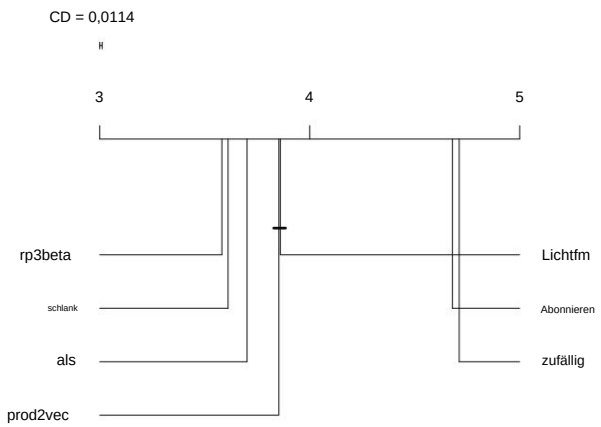


Abbildung 5. Diagramm der kritischen Differenz (für Präzision). Quelle: eigene Studie.

Machine Translated by Google

5.1.3. Genauigkeit für Benutzer mit unterschiedlicher Anzahl von Interaktionen

Die einzigen Informationen, die wir über unsere Benutzer haben, sind ihre Interaktionen. In Abbildung 6 untersuchen wir, wie die Anzahl der Elemente, mit denen ein Benutzer interagiert hat, unsere Hauptmetrik, precision@10, beeinflusst. Wir teilen unsere Benutzer basierend auf der Anzahl der Interaktionen in zehn ähnlich große Gruppen ein .

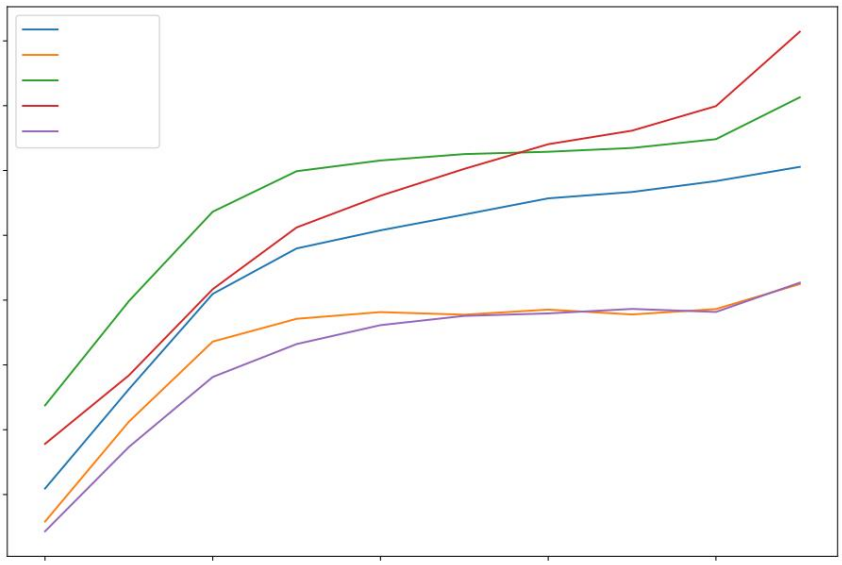


Abbildung 6. Präzision@10 für jedes Modell, abhängig von der Anzahl der Elemente, mit denen der Benutzer interagiert hat. Am besten lässt es sich in Farbe betrachten. Quelle: eigene Studie.

Wir können sehen, dass die Reihenfolge der nach precision@10 sortierten Modelle nicht von der Gruppe abhängt, außer dass SLIM bei Benutzern mit mindestens 22 Interaktionen besser abschneidet als RP3Beta. Der Unterschied ist für jede Benutzergruppe statistisch signifikant (siehe Abschnitt 5.1.4). Das RP3Beta-Modell kann als Sonderfall des itembasierten kollaborativen Filteransatzes angesehen werden, der ein deterministisches Ähnlichkeitsmaß zwischen Elementen verwendet [64]. Andererseits verwendet SLIM Methoden des maschinellen Lernens, um Ähnlichkeiten zu lernen.
Wir gehen davon aus, dass die so erlernten Ähnlichkeitswerte weniger verzerrt sind, aber eine höhere Varianz aufweisen, was zu einer besseren Leistung für Benutzer mit vielen Interaktionen führt, da die Mittelwertbildung über eine größere Anzahl von Werten erfolgt.

5.1.4. Statistische Auswertung der Genauigkeit für Benutzer mit unterschiedlicher Anzahl von Interaktionen

In diesem Unterabschnitt diskutieren wir den Unterschied zwischen den Modellen RP3Beta und SLIM abhängig von der Anzahl der Benutzerinteraktionen, wie in Abbildung 6 zu sehen ist.

Um zwei Methoden über mehrere Datensätze hinweg statistisch zu vergleichen, empfiehlt Demšar [62] den Wilcoxon-Vorzeichen-Rang-Test [65]. Der Wilcoxon-Vorzeichen-Rang-Test ist eine nicht-parametrische Alternative zum Paartest, der die Unterschiede in der Leistung zweier Algorithmen für jeden Datensatz unter Missachtung der Vorzeichen bewertet. Er vergleicht die Ränge der positiven und negativen Unterschiede.

Tabelle 8 zeigt, dass der Unterschied zwischen den Modellen RP3Beta und SLIM statistisch für jede Benutzergruppe von Bedeutung.

Machine Translated by Google

Tabelle 8. Ergebnisse des Wilcoxon-Vorzeichen-Rang-Tests zwischen RP3Beta- und SLIM-Modellen in Abhängigkeit von Anzahl der Elemente, mit denen der Benutzer interagiert hat.

Behälter	p-Wert
[1,0, 3,0)	0
[3,0, 5,0)	0
[5,0, 8,0)	0
[8,0, 11,0)	0
[11,0, 16,0)	0
[16,0, 22,0)	1,25 × 10ÿ6
[22,0, 31,0)	9,27 × 10ÿ7
[31,0, 45,0)	0
[45,0, 74,0)	0
[74,0, 852,0]	0

Quelle: eigene Studie.

5.2. Vielfalt

In Tabelle 9 können wir beobachten, dass Prod2Vec und LightFM die größte Vielfalt bieten nach Ausschluss zufälliger Empfehlungen (die die größte Vielfalt aufweisen sollten). Die Diversitäten der genauesten Modelle, RP3Beta, SLIM und ALS, sind ähnlich, mit Ausnahme der Fall der geringeren Testabdeckung von ALS. Die Beziehung zwischen Testabdeckung und Präzision@10 ist in Abbildung 7 dargestellt .

Tabelle 9. Vielfalt: Bewertungsergebnisse.

Metrisch	RP3Beta	SLIM	ALS	Prod2Vec	LightFM	Beliebteste	Zufällig
Testabdeckung	0,5725	0,5171	0,3038	0,7400	0,7031	0,0002	0,9778
Shannon	9,5271	9,6728	9,6270	10,4031	10,1385	2,3296	11.7267
Gini	0,9083	0,9029	0,9120	0,7956	0,8397	0,9999	0,1159

Quelle: eigene Studie.

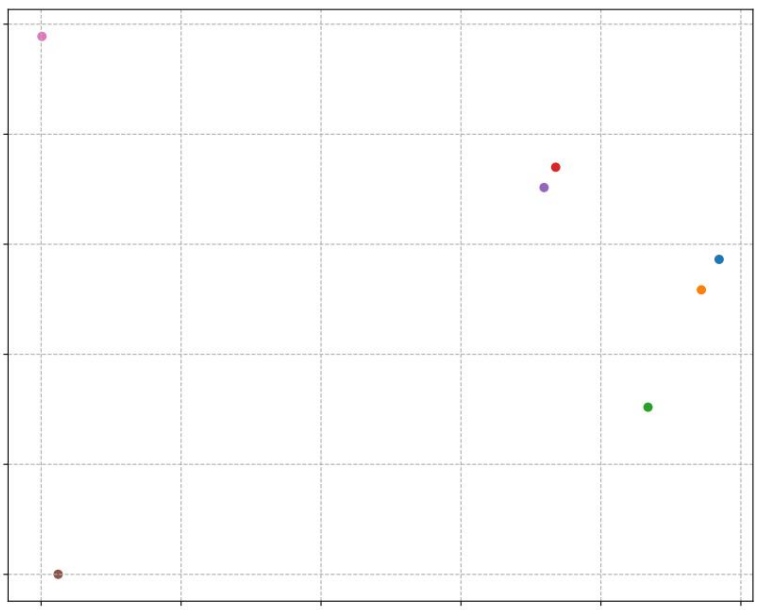


Abbildung 7. Testabdeckung im Vergleich zur Präzision@10 für die betrachteten Modelle. Quelle: eigene Studie.

Machine Translated by Google

5.3. Überschneidungen der Methoden

In unserer Forschung konzentrieren wir uns auf Methoden, die es wert sind, mit echten Benutzern getestet zu werden. Wir untersuchten die Überschneidungen zwischen den Empfehlungen der implementierten Methoden bezüglich der Benutzer-Item-Paare (Overlap-Koeffizient [66]). Wir erwarten signifikantere Unterschiede in den Empfehlungen beim Vergleich von Methoden mit geringer Überlappung.

Die Ergebnisse dieses Experiments sind in Tabelle 10 aufgeführt. RP3Beta und SLIM bieten ähnliche Empfehlungen. Obwohl sie unterschiedlich berechnet werden, verwenden sie eine spärliche Artikel Ähnlichkeitsmatrix. Die Vergleichsgruppe mit ALS weist mittlere Ähnlichkeiten mit anderen Methoden. Am unterschiedlichsten aus der Überlappungsperspektive sind auch die unterschiedlichsten Modelle, nämlich Prod2Vec und LightFM. Die geringe Überlappung von ALS und LightFM unterstreicht die Bedeutung der Wahl der Verlustfunktion bei Matrixfaktorisierungsansätzen.

Tabelle 10. Überschneidung der Modelle.

Modell	RP3Beta	SCHLANK	ALS	Prod2Vec	LichtFM
RP3Beta	100 %	73 %	53 %	37 %	38 %
SLIM	73 %	100 %	50 %	35 %	35 %
ALS	53 %	50 %	100 %	38 %	37 %
Prod2Vec	37 %	35 %	38 %	100 %	28 %
LightFM	38 %	35 %	37 %	28 %	100%

Quelle: eigene Untersuchung.

5.4. Bewertung der Effizienz

Die Ergebnisse der Effizienzbewertung sind in Abbildung 8 dargestellt. RP3Beta, ALS und LightFM übertrifft SLIM und Prod2Vec in Bezug auf die Ausführungszeit, wobei Prod2Vec den geringsten Speicherbedarf hat. Um das Modell häufig neu zu trainieren, haben wir die maximal akzeptable Ausführungszeit von 3 h, was von ALS, LightFM und RP3Beta erfüllt wurde. Aus Kostensicht Perspektive wurde die akzeptable Speicherauslastung auf 110 GB festgelegt, was alle Modelle erfüllt. Wir können beobachten, dass unsere Modelle sogar auf einem einzelnen Laptop trainiert werden konnten.

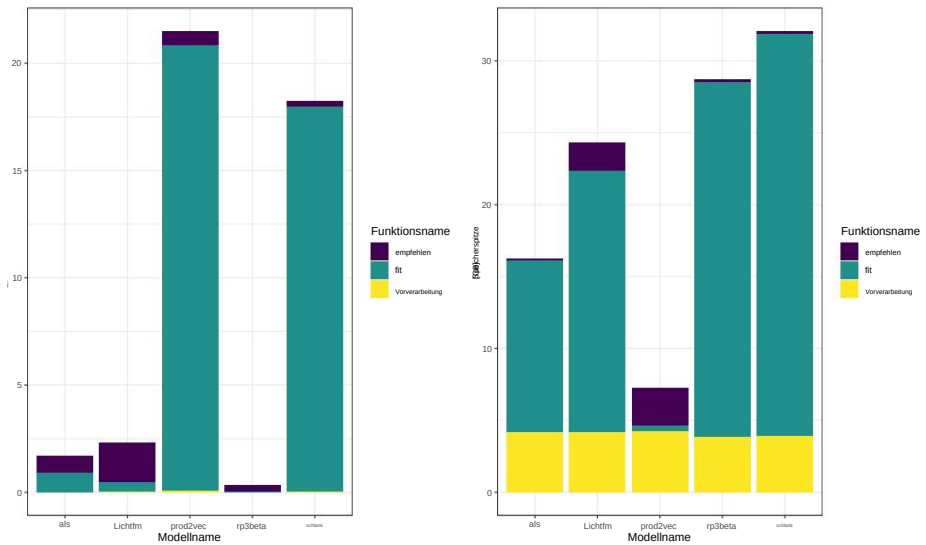


Abbildung 8. Ausführungszeit und Speicherauslastung. Am besten in Farbe betrachten. Quelle: eigene Studie.

5.5. Zusammenfassung

Durch die Laborforschung konnten wir Modelle auswählen, die online mit OLX getestet werden sollten Benutzer. Erstens haben wir das RP3Beta-Modell für Online-Tests gewählt, weil es der beste Ansatz ist hinsichtlich Genauigkeit und Ausführungszeit. Der relativ hohe Speicherverbrauch ist immer noch niedrig genug, um das Modell in der Produktion zu testen. Die Untersuchung der Auswirkungen dieser Die relativ geringe Diversität dieser Methode könnte ein spannender Forschungszweig sein.

Machine Translated by Google

SLIM ist das zweitbeste Modell in Bezug auf Präzision@10. Dennoch haben wir diese Methode aufgrund ihrer geringen Effizienz und der hohen Überschneidung mit dem bereits ausgewählten RP3Beta-Modell nicht für die Online-Evaluierung ausgewählt. Eine Verbesserung der Effizienz der SLIM-Implementierung könnte diese Entscheidung ändern.

Unter anderen Modellen erreicht ALS die höchste Präzision von 10. Darüber hinaus wurde ALS bereits bei OLX implementiert und für Stellenempfehlungen verwendet. Daher haben wir uns entschieden, es in den Online-Vergleich aufzunehmen.

Obwohl Prod2Vec und LightFM die vielfältigsten Empfehlungen liefern, haben wir uns aufgrund der deutlich schlechteren Präzision@10 und Effizienz von Prod2Vec dazu entschlossen, sie nicht online zu testen.

6. Online-Evaluation

Um die Wirksamkeit der ausgewählten Methoden zu bewerten, führten wir zwei Online-A/B- Tests mit OLX Jobs-Benutzern durch. Wir konzentrierten uns auf Benutzer, die kürzlich Stellenanzeigen auf unserer Plattform besucht hatten. Wir schickten ihnen E-Mails mit zehn empfohlenen Stellenanzeigen, und wenn ein Benutzer eine OLX-Anwendung installiert hatte, wurde zusätzlich eine Push-Benachrichtigung mit einer empfohlenen Stellenanzeige gesendet. Dann beobachteten wir, ob sich der Benutzer innerhalb der nächsten 48 Stunden auf eine Stelle bewarb (konvertierter Benutzer).

6.1. Auswirkungen des ALS-Modells

Das erste Experiment zielte darauf ab, die Frage zu beantworten: Erhöht die Einführung der Empfehlungsmethode die Zahl der Benutzer, die sich auf Stellen bewerben? In diesem Test haben wir die Benutzer in zwei Gruppen aufgeteilt: eine Kontrollgruppe ohne Empfehlungen und eine Gruppe mit Empfehlungen, die mithilfe des ALS-Modells generiert wurden. Nur ÿ 10 % der Benutzer wurden der Kontrollgruppe zugewiesen, da das ALS-Modell bereits vor dem Test implementiert und für Stellenempfehlungen bei OLX verwendet wurde. Wir zogen es vor, eine minimale Stichprobe zu sammeln, indem wir einen längeren Test durchführten, um eine längerfristige Auswirkung auf unsere Benutzer zu beobachten. Der Test wurde im März 2021 25 Tage lang durchgeführt. Die Ergebnisse sind in Tabelle 11 aufgeführt. **Die statistische Signifikanz des Vorteils von ALS gegenüber der Kontrollgruppe wurde mithilfe des Chi-Quadrat-Tests untersucht, und der -Wert ergab 0. Dieses Experiment bestätigte unsere Hypothese, dass Empfehlungen die Aktivität beeinflussen, da der Prozentsatz der konvertierten Benutzer nach Erhalt einer Mitteilung signifikant um (16,83 % ÿ 15,98 %)/15,98 % ÿ 5,3 % anstieg.**
Der beobachtete Anstieg bedeutet, dass sich täglich Hunderte zusätzlicher Benutzer über OLX Jobs auf Stellen bewerben.

Tabelle 11. Vergleich der Aktivität mit oder ohne Empfehlungen.

Variante	Benutzer	% Konvertierte Benutzer
Kontrolle	129.308	15,98 %
ALS	1.170.262	16,83 %

Quelle: eigene Studie.

6.2. Vergleich der ALS- und RP3Beta-Modelle

Im zweiten Experiment testeten wir, ob das beste Modell aus Labortests ALS übertraf. Wir teilten unsere Benutzer in drei Gruppen auf, die Empfehlungen von verschiedenen Empfehlungssystemen erhielten: ALS, RP3Beta und eine gemischte Variante. Wir generierten die Hälfte der Empfehlungen mit ALS und die andere Hälfte mit RP3Beta. Das zweite Experiment wurde 28 Tage lang im März und April 2021 durchgeführt. Die E
Der Prozentsatz der konvertierten Benutzer wird hier relativ zu allen Benutzern berechnet. Die Änderung des Empfehlungssystems hatte jedoch keine Auswirkungen auf Benutzer, die die Nachricht nicht geöffnet haben. Wir haben uns entschieden, diese herauszufiltern, da ihr Verhalten unseren Daten unnötiges Rauschen hinzufügte. Tabelle 13 zeigt die Auswirkungen auf Benutzer, die die Nachricht geöffnet haben. Der Vorteil von RP3Beta gegenüber ALS ist statistisch signifikant (p-Wert ÿ 10ÿ6). Der Unterschied zwischen den Varianten RP3Beta und ALS+RP3Beta ist statistisch nicht signifikant (der p-Wert beträgt 0,19).

Machine Translated by Google

Tabelle 12. Vergleich verschiedener Empfehlungsmethoden.

Variante	Benutzer	% Konvertierte Benutzer
ALS	343.892	15,25 %
RP3Beta	345.273	15,40 %
ALS+RP3Beta	343.896	15,30 %

Quelle: eigene Studie.

Tabelle 13. Vergleich verschiedener Empfehlungsmethoden für Benutzer, die die Nachricht geöffnet haben.

Variante	Benutzer	% Konvertierte Benutzer
ALS	44.775	19,66 %
RP3Beta	46.097	20,94 %
ALS+RP3Beta	45.469	20,59 %

Quelle: eigene Studie.

Von den Benutzern, die die Nachricht geöffnet haben, haben wir 1,28 Prozentpunkte beobachtet. mehr konvertierte Benutzer von RP3Beta als von ALS. Da die Nachrichtenöffnungsrate 13,2 % betrug, Dies entspricht einem Anstieg von 1,28 × 0,132 ÿ 0,17 Prozentpunkten aller Zielnutzer. Basierend auf der Ergebnisse des ersten Experiments, ALS bringt eine Steigerung von 0,85 Prozentpunkten gegenüber der Kontrollgruppe. Das bedeutet dass der Ersatz von ALS durch RP3Beta unsere Wirkung auf alle Zielbenutzer um ca. 0,17/0,85 = 20 %.

6.3. Diskussion

Wir haben nachgewiesen, dass das Senden von Jobempfehlungen aus dem ALS-Modell die Anzahl der Nutzer, die auf Stellenanzeigen antworten, um mehr als 5%. Darüber hinaus hat RP3Beta übertrifft ALS.

Zusätzlich führten wir einen weiteren Online-A/B-Test durch, bei dem wir das RP3Beta-Modell mit eine Variante, bei der wir die Hälfte der RP3Beta-Empfehlungen durch Empfehlungen ersetzt haben generiert aus dem Deep-Learning-Hybrid-Empfehlungssystem, das bei OLX verwendet wird (https://tech.olx.com/item2vec-neural-item-embeddings-to-enhance-recommendations-1fd948a6f293, abgerufen am 29. Mai 2024). Wir haben uns entschieden, diesen Test in diesem Artikel nicht zu beschreiben, da diese Arbeit konzentriert sich auf kollaborative Filteransätze. Wir wollten auch, dass unsere Ergebnisse auf dem veröffentlichten Datensatz reproduzierbar bleiben. Leider konnten wir nicht alle Funktionen veröffentlic von unserem internen Modell genutzt. Wir konnten keine Verbesserungen gegenüber RP3Beta feststellen, obwohl das zweite Modell zusätzliches Wissen über Stellenanzeigen nutzt. Daher bestätigte, dass ein richtig bestimmtes und abgestimmtes einfaches Modell mehr fortgeschrittene Tiefenlösungen. Dies steht im Einklang mit den Ergebnissen von Dacrema et al. [64] und Anelli et al. [30] verglichen mehrere nicht-neuronale Ansätze, darunter RP3Beta, mit fortgeschrittenere neuronale Modelle.

Die Effizienz des RP3Beta-Modells ermöglicht es OLX Jobs, es von Grund auf zu trainieren, mehrere mal pro Tag mit sehr geringen Kosten (auf einer CPU). Die hochmoderne kollaborative Filterung Das Modell bei OLX Jobs wird derzeit verwendet, um E-Mail- und Push-Empfehlungen für unsere Nutzer. Wir haben dieses Modell kürzlich angepasst, um Empfehlungen vor Ort in Echtzeit bereitzustellen, unter Berücksichtigung der jüngsten Benutzerinteraktionen mit Elementen [67]. Wir schlugen auch das P3-Lernen vor, um Rangmodell (P3LTR) [68], eine Verallgemeinerung des RP3Beta, bei der wir die Bedeutung von Benutzer-Element-Beziehungen basierend auf den Typen vergangener Benutzer-Element-Interaktionen und Zeitstempeln.

In dieser Arbeit haben wir uns mit der Forschungslücke des Mangels an Vergleichen von klassische kollaborative Filtering-Ansätze im Jobbereich durch umfassende Laborforschung an einem realen Datensatz und eine Online-Auswertung mit Millionen von Benutzern. Darüber hinaus haben wir den Quellcode für die Offline-Bewertung und den Datensatz öffentlich gemacht verfügbar. Wir ermutigen andere Forscher, neue Empfehlungsmethoden zu entwickeln, um helfen Benutzern, schneller ihren Traumjob zu finden.

Die in dieser Arbeit betrachteten Methoden erfordern nur (Benutzer-, Artikel- und Zeitstempel-)Tripletts als Eingangsdaten für Training und Evaluation; daher können sie in mehreren Bereichen eingesetzt werden,

Machine Translated by Google

sogar für Datensätze mit mehreren zehn Millionen Interaktionen. Aufgrund unserer ersten Ergebnisse in anderen OLX-Kategorien erwarten wir gute Ergebnisse im Bereich Kleinanz

7. Zusammenfassung

In diesem Artikel wurde das Thema Stellenempfehlungen in Online-Kleinanzeigen am Beispiel von OLX Jobs behandelt. Wir haben eine umfassende Bewertung von Ansätzen durchgeführt, die verschiedene Familien von Empfehlungsmethoden repräsentieren.

Während der Laborforschung haben wir gezeigt, dass RP3Beta, SLIM und ALS deutlich bessere Ergebnisse erzielen als LightFM und Prod2Vec. Wir haben auch festgestellt, dass Prod2Vec die außergewöhnlichste Vielfalt bietet, die bei der Arbeit an der Neuartigkeit von Empfehlungen erforderlich sein kann . Darüber hinaus haben wir untersucht, wie ähnlich sich die **von verschiedenen Modellen generierten Empfehlungen** sind.

Wir haben außerdem eine Auswertung durchgeführt, indem wir Millionen von Nachrichten an Online-Benutzer gesendet haben.

Es wurden A/B-Tests mit ALS und RP3Beta durchgeführt. Die Ergebnisse dieser Tests zeigten, dass das Senden von Jobempfehlungen, die von diesen Modellen generiert wurden, die Anzahl der Benutzer erhöht, die Werbetreibende von OLX Jobs kontaktieren. Mit RP3Beta in der Produktionsumgebung konnten wir unsere Effektivität um 20 % steigern (unter Berücksichtigung der Benutzerreaktionen).

Bei der Bewertung der Methoden aus der Rubrikenperspektive erwies sich ein Modell, das **die beliebtesten (noch nicht gesehenen) Artikel empfiehlt, für Stellenempfehlungen als ungeeignet.** Um **den Vorteil der untersuchten Methoden aufzuzeigen, wurde zum Vergleich zusätzlich ein Empfehlungsmodell auf Basis zufälliger Empfehlungen herangezogen .**

Ein weiteres wesentliches Ergebnis dieser Untersuchung ist der Datensatz, der weitere Untersuchungen zu Empfehlungen für klassifizierte Bedürfnisse ermöglicht.

Eine der Schwächen unserer Arbeit ist das Fehlen eines Vergleichs mit tiefen neuronalen Netzwerken (DNNs) oder graphisch neuronalen Netzwerken (GNNs). Bei solchen Methoden ist es eine Herausforderung, unsere Zeitbeschränkungen ohne kostspieligere Infrastruktur (z. B. mit GPU) einzuhalten. Für künftige Arbeiten wäre es interessant, unsere Beschränkungen zu lockern oder die effizientesten DNNs oder GNNs zu vergleichen . Eine weitere Schwäche besteht darin, dass der Vergleich nur auf einem einzigen Datensatz basiert. Die Größe des Datensatzes hat einen erheblichen Einfluss auf die Methodenauswahl. Es wäre interessant, dieselben Methoden an Datensätzen unterschiedlicher Größe zu vergleichen, entweder an Teilmengen des ursprünglichen Datensatzes oder an anderen Datensätzen aus dem Berufsbereich (z. B. aus verschiedenen Märkten, in denen OLX tätig ist). Darüber hinaus wäre es interessant, den Zusammenhang zwischen Diversität und Online-Ergebnissen zu untersuchen , um eine Offline-Kennzahl zu entwicke

Beiträge der Autoren: Konzeptualisierung, RK; Methodik, RK und TG; Software, RK und VD; Validierung, RK, VD, TG und AF; formale Analyse, RK und TG; Untersuchung, RK und TG; Ressourcen, RK; Datenkuratierung, RK und VD; Schreiben – Vorbereitung des Originalentwurfs, RK, AF und TG; Schreiben – Überprüfung und Bearbeitung, RK, AF und TG; Visualisierung, RK und TG; Überwachung, TG und AF; Projektverwaltung, RK. Alle Autoren haben die veröffentlichte Version **des Manuskripts gelesen und stimmen ihr zu .**

Finanzierung: Diese Forschung erhielt keine externe Finanzierung.

Datenverfügbarkeitserklärung: Datensatz: https://www.kaggle.com/datasets/olxdatascience/olx- jobs-interactions, abgerufen am 29. Mai 2024; Quellcode: https://github.com/rob-kwiec/olx-jobs-recommendations , abgerufen am 29. Mai 2024

Interessenkonflikte: Die Autoren Robert Kwieci ´nski, Agata Filipowska und Viacheslav Dubrov waren bei der Firma OLX Group beschäftigt. Der andere Autor erklärt, dass die Forschung ohne kommerzielle oder finanzielle Beziehungen durchgeführt wurde, die als potenzieller Interessenkonflikt ausgelegt werden könnten.

Verweise

1. Davis, M. Was sind Online-Kleinanzeigen? Big Easy Magazine, **4. Januar 2021. 2.**
Jones, S. Online-Kleinanzeigen; **Pew Internet & American Life Project: Washington, DC, USA, 2009.**
3. Gourville, J.; Soman, D. Überauswahl und Sortimentstyp: Wann und warum Vielfalt nach hinten losgeht. Mark. Sci. **2005,** 24, **382–395.**
[Querverweis]

Machine Translated by Google

4. Chernev, A. Wenn mehr weniger ist und weniger mehr ist: Die Rolle der idealen Punktverfügbarkeit und des Sortiments bei der Wahl des Verbrauchers. J. Consum. Res. **2003,** 30, **170–183. [CrossRef]**

5. Boatwright, P.; Nunes, J. Reducing Assortment: Ein attributbasierter Ansatz. J. Mark. **2001,** 65, **50–63. [CrossRef]**

6. Gómez-Uribe, C.; Hunt, N. Das Netflix-Empfehlungssystem. ACM Trans. Manag. Inf. Syst. **2015, 6, 1–19. [CrossRef]**

7. Smith, B.; Linden, G. Zwei Jahrzehnte Empfehlungssysteme bei Amazon.com. IEEE Internet Comput. **2017,** 21, **12–18. [CrossRef]**

8. Aggarwal, CC Recommender Systems: The Textbook, **1. Auflage; Springer: Berlin/Heidelberg, Deutschland, 2016. 9.**

Li, S.; Karahanna, E. Online-Empfehlungssysteme im B2C-E-Commerce-Kontext: Ein Überblick und zukünftige Richtungen. J. Assoc. Inf. Syst. **2015,** 16, **72–107. [CrossRef]**

10. Twardowski, B. Modellierung kontextbezogener Informationen in sitzungsabhängigen Empfehlungssystemen mit neuronalen Netzwerken. In Proceedings der 10. ACM-Konferenz zu Empfehlungssystemen, Boston, MA, USA, 15.–19. September 2016; S. 273–276. [Querverweis]

11. Alotaibi, S. Eine Untersuchung von Job-Empfehlungssystemen. Int. J. Phys. Sci. **2012, 7, 5127–5142. [CrossRef]**

12. Abel, F.; Benczúr, A.; Kohlsdorf, D.; Larson, M.; Pálovics, R. RecSys Challenge'16: Proceedings of the Recommender Systems Challenge; **Verband für Rechenmaschinen: New York, NY, USA, 2016.**

13. Abel, F.; Deldjoo, Y.; Elahi, M.; Kohlsdorf, D. RecSys Challenge'17: Proceedings der Recommender Systems Challenge 2017; **Association for Computing Machinery: New York, NY, USA, 2017.**

14. Li, S.; Shi, B.; Yang, J.; Yan, J.; Wang, S.; Chen, F.; He, Q. Tiefes Jobverständnis bei LinkedIn. In Proceedings der 43. Internationalen ACM SIGIR-Konferenz zu Forschung und Entwicklung im Bereich Informationsabruf, virtuell, 25.–30. Juli 2020; S. 2145–2148. [CrossRef]

15. Zhu, G.; Chen, Y.; Wang, S. Graph-Community-gestützte personalisierte Kurs-Job-Empfehlungen mit domänenübergreifenden Daten Integration. Sustainability **2022,** 14, **7439. [CrossRef]**

16. Lacic, E.; Reiter-Haas, M.; Kowald, D.; Reddy Dareddy, M.; Cho, J.; Lex, E. Verwendung von Autoencodern für sitzungsbasierte Jobempfehlungen dations. Benutzermodell. Benutzer-Adapt. Interagieren. **2020,** 30, **617–658. [CrossRef]**

17. Pazzani, MJ; Billsus, D. Inhaltsbasierte Empfehlungssysteme. In The Adaptive Web: Methoden und Strategien des Web Personalisierung; **Springer: Berlin/Heidelberg, Deutschland, 2007; S. 325–341. [CrossRef]**

18. Chen, R.; Hua, Q.; Chang, YS; Wang, B.; Zhang, L.; Kong, X. Eine Untersuchung kollaborativer, filterbasierter Empfehlungssysteme: Von traditionellen Methoden zu hybriden Methoden auf Basis sozialer Netzwerke. IEEE Access **2018, 6, 64301–64320. [CrossRef]**

19. Wang, X.; He, X.; Wang, M.; Feng, F.; Chua, TS Neural Graph Collaborative Filtering. In Proceedings der 42. Internationalen ACM SIGIR-Konferenz zu Forschung und Entwicklung im Bereich Informationsabruf, Paris, Frankreich, 21.–25. Juli 2019; S. 165–174. [Querverweis]

20. Schafer, JB; Frankowski, D.; Herlocker, J.; Sen, S. Collaborative filtering recommender systems. In The Adaptive Web; **Springer: Berlin/Heidelberg, Deutschland, 2007; S. 291–324. [CrossRef]**

21. Takács, G.; Pilászy, I.; Németh, B.; Tikk, D. Skalierbare kollaborative Filteransätze für große Empfehlungssysteme. J. Mach. Lernen. Res. **2009,** 10, **623–656. [CrossRef]**

22. Jawaheer, G.; Szomszor, M.; Kostkova, P. Vergleich von implizitem und explizitem Feedback eines Online-Musikempfehlungsdienstes . In Proceedings des 1. Internationalen Workshops zu Informationsheterogenität und -fusion in Empfehlungssystemen, Barcelona, Spanien, 26. September 2010; S. 47–51. [CrossRef]

23. Tamm, YM; Damdinov, R.; Vasilev, A. Qualitätsmetriken in Empfehlungssystemen: Berechnen wir Metriken konsistent? In Proceedings der fünfzehnten ACM-Konferenz zu Empfehlungssystemen, Amsterdam, Niederlande, 27. September–1. Oktober 2021; S. 708–713. [CrossRef]

24. Kaminskas, M.; Bridge, D. Diversität, Serendipität, Neuheit und Reichweite: Eine Umfrage und empirische Analyse von über die Genauigkeit hinausgehenden Zielen in Empfehlungssystemen. ACM Trans. Interact. Intell. Syst. **2016, 7, 1–42. [CrossRef]**

25. Kunaver, M.; Pozrl, T. Vielfalt in Empfehlungssystemen, eine Untersuchung. Knowl.-Based Syst. **2017,** 123, **154–162. [CrossRef]**

26. Ge, M.; Delgado-Battenfeld, C.; Jannach, D. Mehr als Genauigkeit: Bewertung von Empfehlungssystemen anhand von Abdeckung und Serendipität. In Proceedings der vierten ACM-Konferenz zu Empfehlungssystemen, Barcelona, Spanien, 26.–30. September 2010; S. 257–260. [Querverweis]

27. Mogenet, A.; Pham, TAN; Kazama, M.; Kong, J. Vorhersage der Online-Leistung von Job-Empfehlungssystemen mit Offline- Bewertung. In Proceedings der 13. ACM-Konferenz zu Empfehlungssystemen, Kopenhagen, Dänemark, 16.–20. September 2019; S. 477–480. [CrossRef]

28. Anelli, VW; Bellogín, A.; Di Noia, T.; Jannach, D.; Pomo, C. Top-N-Empfehlungsalgorithmen: Eine Suche nach dem neuesten Stand der Technik. In Proceedings der 30. ACM-Konferenz zu Benutzermodellierung, Anpassung und Personalisierung, Barcelona, Spanien, 4.–7. Juli 2022; S. 121–131. [CrossRef]

29. Dacrema, MF; Boglio, S.; Cremonesi, P.; Jannach, D. Eine beunruhigende Analyse der Reproduzierbarkeit und des Fortschritts bei Empfehlungs Systemforschung. ACM Trans. Inf. Syst. **2021,** 39, **1–49. [CrossRef]**

30. Anelli, VW; Bellogín, A.; Di Noia, T.; Pomo, C. Neugestaltung des Vergleichs zwischen neuronalem kollaborativem Filtern und Matrixfaktorisierung. In Proceedings der fünfzehnten ACM-Konferenz zu Empfehlungssystemen, Amsterdam, Niederlande, 27. September–1. Oktober 2021; S. 521–529. [CrossRef]

Machine Translated by Google

31. Zhu, J.; Dai, Q.; Su, L.; Ma, R.; Liu, J.; Cai, G.; Xiao, X.; Zhang, R. BARS: Auf dem Weg zum offenen Benchmarking für Empfehlungssysteme. In Proceedings der 45. Internationalen ACM SIGIR-Konferenz zu Forschung und Entwicklung im Bereich Informationsabruf, Madrid, Spanien, 11.–15. Juli 2022; S. 2912–2923. [CrossRef]
32. Dong, Y.; Li, J.; Schnabel, T. Wenn neuer nicht besser ist: Profitiert Deep Learning wirklich von Empfehlungen aus implizitem Feedback? In Proceedings der 46. Internationalen ACM SIGIR-Konferenz zu Forschung und Entwicklung im Bereich Informationsabruf , Taipeh, Taiwan, 23.–27. Juli 2023; S. 942–952. [CrossRef]
33. Anelli, VW; Malitesta, D.; Pomo, C.; Bellogin, A.; Di Sciascio, E.; Di Noia, T. Den Mythos des Graph Collaborative Filtering in Frage stellen: Eine begründete und reproduzierbarkeitsorientierte Analyse. In Proceedings der 17. ACM-Konferenz zu Empfehlungssystemen , Singapur, 18.–22. September 2023; S. 350–361. [CrossRef]
34. Koren, Y.; Bell, R.; Volinsky, C. Matrixfaktorisierungstechniken für Empfehlungssysteme. Computer **2009**, 42, **30–37.** [CrossRef]
35. Kula, M. Metadateneinbettungen für Benutzer- und Artikel-Kaltstart-Empfehlungen. arXiv **2015, arXiv:1507.08439.**
36. Rendle, S.; Freudenthaler, C.; Gantner, Z.; Schmidt-Thieme, L. BPR: Bayesian Personalized Ranking from Implicit Feedback. arXiv **2012, arXiv:1205.2618 37. Weston,**
J.; Bengio, S.; Usunier, N. Wsabie: Scaling up to large vocabulary image annotation. In Proceedings der 22. Internationalen Gemeinsamen Konferenz für Künstliche Intelligenz, Barcelona, Spanien, 16.–22. Juli 2011; S. 2764–2770.
38. Weston, J.; Yee, H.; Weiss, RJ Lernen, Empfehlungen mit dem K-Order-Statistikverlust zu bewerten. In Proceedings der 7. ACM-Konferenz über Empfehlungssysteme, Hongkong, China, 12.–16. Oktober 2013; S. 245–248. [CrossRef]
39. Hu, Y.; Koren, Y.; Volinsky, C. Kollaboratives Filtern für implizite Feedback-Datensätze. In Proceedings der 2008 8. IEEE Internationale Konferenz zum Thema Data Mining, Pisa, Italien, 15.–19. Dezember 2008; S. 263–272. [CrossRef]
40. Takács, G.; Pilászy, I.; Tikk, D. Anwendungen der konjugierten Gradientenmethode für kollaboratives Filtern mit implizitem Feedback. In Proceedings der fünften ACM-Konferenz zu Empfehlungssystemen, Chicago, IL, USA, 23.–27. Oktober 2011; S. 297–300. [Querverweis]
41. Ning, X.; Karypis, G. SLIM: Sparse Linear Methods for Top-N Recommender Systems. In Proceedings der 2011 IEEE 11th International Conference on Data Mining, Vancouver, BC, Kanada, 11.–14. Dezember 2011; S. 497–506. [CrossRef]
42. Christoffel, F.; Paudel, B.; Newell, C.; Bernstein, A. Blockbusters and Wallflowers: Accurate, Diverse, and Scalable Recommen- dations with Random Walks. In Proceedings der 9. ACM-Konferenz zu Empfehlungssystemen, Wien, Österreich, 16.–20. September 2015; S. 163–170. [CrossRef]
43. Cooper, C.; Lee, SH; Radzik, T.; Siantos, Y. Zufallswanderungen in Empfehlungssystemen: Genaue Berechnung und Simulationen. In Proceedings der 23. Internationalen Konferenz zum World Wide Web, New York, NY, USA, 7.–11. April 2014; S. 811–816. [Querverweis]
44. Grbovic, M.; Radosavljevic, V.; Djuric, N.; Bhamidipati, N.; Savla, J.; Bhagwan, V.; Sharp, D. E-Commerce in Ihrem Posteingang: Produktempfehlungen im großen Maßstab. In Proceedings der 21. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australien, 10.–13. August 2015; S. 1809–1818. [CrossRef]
45. Barkan, O.; Koenigstein, N. ITEM2VEC: Neurale Item-Einbettung für kollaboratives Filtern. In Proceedings des 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP), Vietri sul Mare, Italien, 13.–16. September 2016; S. 1–6. [Querverweis]
46. Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.; Dean, J. Verteilte Darstellungen von Wörtern und Phrasen und ihre Kompositionalität. In Proceedings der 26. Internationalen Konferenz über neuronale Informationsverarbeitungssysteme – Band 2, Red Hook, NY, USA, 5.–10. Dezember 2013; S. 3111–3119.
47. De Ruijt, C.; Bhulai, S. Job-Empfehlungssysteme: Eine Überprüfung. arXiv **2021, arXiv:2111.13576.**
48. Kula, M. LightFM—Github Repository. 2023. Online verfügbar: https://github.com/lyst/lightfm (abgerufen am 29. Mai 2024).
49. Frederickson, B. ALS—Github Repository. 2022. Online verfügbar: https://github.com/benfred/implicit/blob/main/implicit/als.py (abgerufen am 29. Mai 2024).
50. Levy, M.; Grisel, O. SLIM—Github Repository. 2014. Online verfügbar: https://github.com/Mendeley/mrec/blob/master/ mrec/item_similarity/slim.py (abgerufen am 29. Mai 2024).
51. Quadrana, M. SLIM—Github Repository. 2021. Online verfügbar: https://github.com/MaurizioFD/RecSys2019_DeepLearning_ Evaluation/blob/master/SLIM_ElasticNet/SLIMElasticNetRecommender.py (abgerufen am 29. Mai 2024).
52. Bernardis, C. RP3Beta—Github Repository. 2021. Online verfügbar: https://github.com/MaurizioFD/RecSys2019_DeepLearning_ Auswertung/blob/master/GraphBased/RP3betaRecommender.py (abgerufen am 29. Mai 2024).
53. Paudel, B.; Christoffel, F.; Newell, C.; Bernstein, A. Aktualisierbare, genaue, vielfältige und skalierbare Empfehlungen für interaktive Anwendungen. ACM Trans. Interact. Intell. Syst. **2016, 7, 1–34.** [CrossRef]
54. Reh ůuřek, R.; Sojka, P. Software Framework for Topic Modelling with Large Corpora. In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, Valletta, Malta, 22. Mai 2010; S. 45–50, ISBN 2-9517408-6-7. Online verfügbar: http://is.muni.cz/publication/884893/en (abgerufen am 29. Mai 2012).
55. Meng, Z.; McCreadie, R.; Macdonald, C.; Ounis, I. Erforschung von Datenaufteilungsstrategien zur Bewertung von Empfehlungsmodellen . In Proceedings der 14. ACM-Konferenz zu Empfehlungssystemen, virtuell, 22.–26. September 2020; S. 681–686. [Querverweis]
56. Shani, G.; Gunawardana, A., Evaluating Recommendation Systems. In Recommender Systems Handbook; **Springer: Boston, MA, USA, 2011; S. 257–297.** [CrossRef]

Machine Translated by Google

57. Friedman, M. Die Verwendung von Rängen zur Vermeidung der impliziten Annahme der Normalität in der Varianzanalyse. J. Am. Stat. Assoc. **1937,** 32, **675–701.** [CrossRef]

58. Friedman, M. Ein Vergleich alternativer Signifikanztests für das Problem der Rangfolge. Ann. Math. Stat. **1940,** 11, **86–92.** [Querverweis]

59. Iman, RL; Davenport, JM Näherungen des kritischen Bereichs der Friedman-Statistik. Commun. Stat.-Theory Methods **1980,** 9, 571–595. [CrossRef]

60. García, S.; Herrera, F. Eine Erweiterung zum „Statistischen Vergleich von Klassifikatoren über mehrere Datensätze“ für alle paarweisen Vergleiche. J. Mach. Learn. Res. **2008, 9, 2677–2694.**

61. García, S.; Fernández, A.; Luengo, J.; Herrera, F. Erweiterte nichtparametrische Tests für Mehrfachvergleiche bei der Versuchsplanung in Computational Intelligence und Data Mining: Experimentelle Leistungsanalyse. Inf. Sci. **2010,** 180, **2044-** [Querverweis]

62. Demšar, J. Statistische Vergleiche von Klassifikatoren über mehrere Datensätze. J. Mach. Learn. Res. **2006, 7, 1–30.**

63. Nemenyi, P. Verteilungsfreie Mehrfachvergleiche. Dissertation, Princeton University, Princeton, NJ, USA, 1963.

64. Dacrema, MF; Cremonesi, P.; Jannach, D. Machen wir wirklich große Fortschritte? Eine beunruhigende Analyse der jüngsten neuronalen Empfehlungsansätze. In Proceedings der 13. ACM-Konferenz zu Empfehlungssystemen, Kopenhagen, Dänemark, 16.–20. September 2019; S. 101–109. [CrossRef]

65. Wilcoxon, F. Individuelle Vergleiche durch Rankingmethoden. Biom. Bull. **1945, 1, 80–83.** [CrossRef]

66. Vijaymeena, M.; Kavitha, K. Eine Untersuchung zu Ähnlichkeitsmaßen im Text Mining. Mach. Learn. Appl. Int. J. **2016, 3, 19–28.** [CrossRef]

67. Kwieci ´nski, R.; Melniczak, G.; Górecki, T. Vergleich von Empfehlungen für Echtzeit- und Batch-Jobs. IEEE Access **2023,** 11, **20553–20559.** [CrossRef]

68. Kwieci ´nski, R.; Górecki, T.; Filipowska, A. Lernen der Kantenwichtigkeit in bipartiten, graphenbasierten Empfehlungen. In Proceedings der 17. Konferenz für Informatik und Intelligenzsysteme (FedCSIS) 2022, Sofia, Bulgarien, 4.–7. September 2022; S. 227–233. [CrossRef]

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Effiziente Sprachsignal-Dimensionsreduktion durch Komplexwertige Techniken

Sungkyun Ko [1] **und Minho Park 2,***

[1] Abteilung für KI-IT-Konvergenz, Soongsil University, Seoul 06978, Republik Korea; kosungkyun@naver.com
[2] Fakultät für Elektrotechnik, Soongsil University, Seoul 06978, Republik Korea
* Korrespondenz: mhp@ssu.ac.kr

Zusammenfassung: In dieser Studie schlagen wir den CVMFCC-DR- Algorithmus (Complex-Valued Mel-Frequency Cepstral Co-efficients Dimensionality Reduction) als effiziente Methode zur Reduzierung der Dimensionalität von Sprachsignalen vor. Durch die Verwendung der komplexwertigen MFCC-Technik, die sowohl reale als auch imaginäre Komponenten berücksichtigt, ermöglicht unser Algorithmus eine Dimensionsreduzierung ohne Informationsverlust bei gleichzeitiger Senkung des Rechenaufwands. Die Wirksamkeit des vorgeschlagenen Algorithmus wird durch Experimente bestätigt, die seine Effektivität beim Aufbau eines Spracherkennungsmodells unter Verwendung eines komplexwertigen neuronalen Netzwerks demonstrieren. Zusätzlich wird eine komplexwertige Softmax-Interpretationsmethode für komplexe Zahlen eingeführt. Die experimentellen Ergebnisse zeigen, dass der Ansatz im Vergleich zu herkömmlichen MFCC-basierten Techniken eine verbesserte Leistung bietet, was sein Potenzial im Bereich der Spra

Schlüsselwörter: Mel-Frequenz-Cepstrum-Koeffizienten; Fehlerkorrektur bei komplexwertigen neuronalen Netzwerken; Backpropagation; komplexwertiger Softmax

check for updates

Zitat: Ko, S.; Park, M. Effiziente Reduzierung der Sprachsignaldimensionalität mithilfe komplexwertiger Techniken. Electronics **2024,** 13, **3046. https://doi.org/10.3390/Elektronik13153046**

Wissenschaftliche Herausgeber: Xin Geng, Padma Iyenghar und Elke Pulvermüller

Erhalten: 3. Juni 2024
Überarbeitet: 31. Juli 2024
Akzeptiert: 31. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Einleitung

Heutzutage spielen Sprachsignalverarbeitung und -erkennung in zahlreichen Anwendungsbereichen eine entscheidende Rolle . Diese Technologien werden für unterschiedliche Zwecke eingesetzt, beispielsweise um Sprachbefehle in Text umzuwandeln oder den Besitzer einer Stimme zu identifizieren. Die Weiterentwicklung der Spracherkennungstechnologie kann die Mensch-Maschine-Interaktion verbessern und so die Produktivität und Effizienz in zahlreichen Branchen steigern. Sprachsignale weisen jedoch sehr unterschiedliche Formen und hochdimensionale Eigenschaften auf, was eine effektive Verarbeitung und Erkennung zu einer ständigen Herausforderung macht. Die Vielfalt und hohe Dimensionalität von Sprachsignalen erschweren die Dimensionsreduzierung und Mustererkennung, was zu erhöhten Rechenkosten führt. Daher ist die effiziente Reduzierung der Dimensionalität von Sprachsignalen und die Verbesserung der Leistung neuronaler netzbasierter S

Technologien zur Sprachverarbeitung und -erkennung sind für zahlreiche Anwendungen wie Sprachschnittstellen, virtuelle Assistenten und automatische Übersetzungen unverzichtbar. Genauere und effizientere Ergebnisse bei der Sprachsignalverarbeitung und -erkennung können das Benutzererlebnis bei entsprechenden Diensten erheblich verbessern. Aus diesem Grund wird im Bereich der Sprachsignalverarbeitung und -erkennung kontinuierlich geforscht. Sprachsignale weisen hochdimensionale Eigenschaften auf, was eine effektive Dimensionsreduzierung Darüber hinaus erfordert die genaue Erkennung unterschiedlicher Sprachmuster komplexe neuronale Netzwerkmodelle , was wiederum den Rechenaufwand erhöht [1]. Daher ist es eine entscheidende Aufgabe, die Dimensionalität von Sprachsignalen ohne Informationsverlust effektiv zu reduzieren und gleichzeitig die Genauigkeit und Effizienz neuronaler Netzwerk-basierter Spracherkennungsmodelle zu verbessern .

Trotz der umfangreichen Literatur zu komplexwertigen Netzwerken gibt es nur wenige Beispiele für deren Anwendung in der Audio- oder Sprachverarbeitung. Da das menschliche Hörsys

Machine Translated by Google

Da neuronale Netzwerke Schall im Frequenzbereich verarbeiten, ist es auch sinnvoll, komplexwertige neuronale Netzwerke für die Audioverarbeitung zu verwenden, anstatt sich ausschließlich auf realwertige neuronale Netzwerke im Zeitbereich zu verlassen **[2].**

Durch die Kombination **traditioneller Methoden wie der Mel-Frequency-Cepstral-Koeffizienten (MFCCs) mit neuronalen Netzwerken wurden in der jüngsten Vergangenheit erhebliche Fortschritte bei der Spracherkennung erzi Obwohl die MFCC-Technik häufig zur Extraktion von Sprachmerkmalen verwendet wird, hat sie den Nachteil, dass sie keine verlustfreie Dimensionsreduzierung durchführen kann, da sie nur reale Werte verwendet. Mit anderen Worten: Die Reduzierung der Anzahl der MFCCs führt zu Informationsverlusten, was die Genauigkeit der Spracherkennung verringert. Diese Methoden sind immer noch nur begrenzt effektiv und es sind verbesserte Ansätze erforderlich, die über traditionelle Lernmethoden hinausgehen.**

Die Verwendung der komplexen Darstellung von MFCC ermöglicht die Berücksichtigung sowohl des Real- als auch des Imaginärteils und ermöglicht so eine verlustfreie Dimensionsreduktion. Darüber hinaus können durch den Einsatz komplexwertiger neuronaler Netze Phasenänderungen durch komplexe Multiplikationsoperationen abgebildet werden [4], wodurch die Eigenschaften von Sprachsignalen besser dargestellt werden. Dieser Ansatz dürfte im Vergleich zu bestehenden Methoden eine effizientere Dimensionsreduktion und eine auf neuronalen Netzen basierende Spracherkennung ermögliche

In dieser Studie schlagen wir den CVMFCC-DR-Algorithmus zur effizienten und verlustfreien Reduzierung der Sprachsignaldimensionalität vor. Durch die Kombination von CVMFCC-DR mit komplexwertigen neuronalen Netzwerken kann diese Technik **im Vergleich zu bestehenden Methoden eine bessere Spracherkennungsleistung aufweisen .**

Die wichtigsten Beiträge dieser Studie sind wie folgt:

- **Vorschlag einer effizienten Dimensionsreduktionstechnik (CVMFCC-DR) unter Verwendung der komplexen Darstellung von MFCCs.**
- **Verbesserung der Spracherkennungsleistung durch Kombination von komplexwertigen MFCCs und komplexwertigen neuronalen Netzwerken.**
- **Vorschlag einer neuen Interpretationsmethode für eine komplexwertige Softmax-Funktion mit komplexe Eingaben.**

Dieser Artikel beginnt mit einer Einführung in die bestehende MFCC-Technik und komplexwertige neuronale Netze. Anschließend werden der vorgeschlagene CVMFCC-DR- Algorithmus und die Lernmethode für komplexwertige neuronale Netze [5] ausführlich erläutert. Anschließend wird die Wirksamkeit der vorgeschlagenen Methode experimentell validiert und die Ergebnisse analysiert . Abschließend schließt der Artikel mit einer Zusammenfassung der Ergebnisse der Studie und Vorschlägen für zukünftige Forschungen.

2. Verwandte Werke

Im Bereich der Spracherkennung werden verschiedene Ansätze und Techniken auf Basis neuronaler Netze eingesetzt. Dabei kommen unterschiedliche Arten neuronaler Netze zum Einsatz, wie CNNs, RNNs, LSTM und DNNs, die jeweils ihre eigenen Vor- und Nachteile haben [3]. Darüber hinaus werden auch Hybridmodelle erforscht, die traditionelle Methoden mit neuronalen Netzstrukturen kombinieren . Hybridmodelle können die Leistung optimieren, indem sie die Stärken jeder Struktur kombinieren, haben aber auch den Nachteil einer erhöhten Komplexität bei der Implementierung und Interpretation [6].

Inmitten dieser Entwicklungen werden traditionelle Techniken zur Merkmalsextraktion, wie die MFCC- Technik, noch immer häufig als Eingaben für neuronale Netzwerke verwendet [3]. Darüber hinaus könnte die Erforschung neuartiger neuronaler Architekturen, wie etwa komplexwertiger neuronaler Netzwerke, neue Perspektiven und potenzielle Durchbrüche bei der Spracherkennung eröffnen. Jüngste Studien haben gezeigt, dass sich die Berechnungsgeschwindigkeit komplexer Zahlen mithilfe von FPGAs deutlich verbessern lässt [7]. Diese Forschungsergebnisse zeigen empirisch, dass die Hardwareimplementierung von Operationen mit komplexen Zahlen praktische Leistungsverbesserungen mit sich bringen kann . Darüber hinaus wurden in einer kürzlich durchgeführten Studie komplexwertige neuronale Netzwerke zur Klassifizierung auf FPGAs implementiert. Dabei wurde eine höhere Genauigkeit als realwertige neuronale Netzwerke nachgewiesen und im Vergleich zu CPU-basierten **Systemen eine deutliche Geschwindigkeitsverbesserung erreicht [8]. Dies lässt erwarten, dass CPUs, die heute nur Operationen mit reellen Zahlen unterstützen, in Zukunft vielle**

Machine Translated by Google

Basierend auf diesen Studien ermöglicht die Verwendung komplexer Zahlendarstellungen die Kodierung von reellen Zahlenpaaren als einzelne komplexe Zahlen, wodurch die Eingabedimension effektiv um die Hälfte reduziert wird. Diese Reduzierung der Dimensionalität verringert die Anzahl der Modellparameter und verbessert die Rechenleistung. Folglich zeigen experimentelle Beweise, dass komplexwertige neuronale Netzwerke im Vergleich zu herkömmlichen realwertigen neuronalen Netzwerken mit weniger Parametern eine gleiche oder bessere Leistung erzielen können.

Während die Forschung immer neue Grenzen erweitert, könnte eine synergetische Kombination aus bewährten Methoden und innovativen Ansätzen den Weg für die nächste Generation robuster und leistungsstarker Spracherkennungssysteme ebnen.

3. Hintergrundwissen

Es werden derzeit verschiedene Studien im Bereich der Spracherkennung und -klassifizierung durchgeführt. In diesem Abschnitt stellen wir MFCCs (Mel-Frequency Cepstral Coefficients) vor [3]. Die MFCC-Technik ist eine der in den Bereichen Sprachsignalverarbeitung und Spracherkennung häufig verwendeten Methoden zur Merkmalsextraktion. Die MFCC-Technik stellt die Frequenzeigenschaften von Sprachsignalen dar und wird hauptsächlich zur effektiven Darstellung und Erkennung von Sprachsignalen eingesetzt. Das Grundprinzip der MFCC-Technik ist in Abbildung 1 dargestellt.

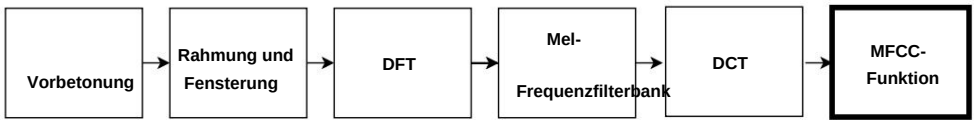


Abbildung 1. Ein Überblick über den Prozess der MFCC-Technik.

Vorverzerrung: Bei der Vorverzerrung handelt es sich im Wesentlichen um einen Hochpassfilter. Aufgrund der Struktur des menschlichen Körpers neigen hochfrequente Komponenten dazu, bei der Sprachproduktion im tatsächlich abgegebenen Ton abzunehmen. Daher handelt es sich bei der Vorverzerrung um den Prozess, bei dem zunächst ein Hochpassfilter angewendet wird, um die gedämpften hochfrequenten Komponenten zu verstärken.

Framing und Windowing: Sprachsignale werden zur Analyse in kurze Segmente, sogenannte Frames, unterteilt. Dies liegt daran, dass die Spracheigenschaften über kurze Zeiträume relativ stabil bleiben. Diese Frames überlappen sich leicht, um Sprachänderungen effektiver zu erfassen.
Jeder Frame wird mithilfe einer Fenstertechnik verarbeitet, die die Analyse verbessert. Zu den gängigen Fenstertypen gehören Hamming- und Hanning-Fenster. Diese Framing-Methode ermöglicht eine genauere Analyse komplexer Sprachdaten, was für die Entwicklung von Spracherkennungs- und -synthesetechnologien von entscheidender Bedeutung ist. Während die spezifischen Implementierungsdetails variieren können, werden diese grundlegenden Prinzipien in Sprachverarbeitungssystemen häufig verwendet.

Diskrete Fourier-Transformation: Zur Berechnung des Leistungsspektrums wird die DFT (Diskrete Fourier-Transformation) verwendet, um das Leistungsspektrum für jeden zuvor erhaltenen Frame zu berechnen. Dieser Prozess ermöglicht die Unterscheidung der frequenzspezifischen Energieverteilung des Sprachsignals über jeden Frame hinweg.

Mel-Frequenz-Filterbank: Der Mel-Bandpassfilter ist eine Reihe von Filtern, die dazu dienen, Sprachsignale auf nichtlineare Weise zu extrahieren und so das menschliche Hörsystem nachzuahmen.
Es besteht normalerweise aus 40 Dreieckfiltern und analysiert die Frequenzeigenschaften von Sprachsignalen.

Diskrete Kosinustransformation: DCT (Diskrete Kosinustransformation) drückt eine endliche Folge von Datenpunkten als Summe von Kosinusfunktionen aus, die mit unterschiedlichen Frequenzen schwingen. Im MFCC-Prozess wird DCT auf die Mel-Filterbank angewendet, um Beschleunigungskoeffizienten auszuwählen oder die Beziehung in log-spektralen Größen von der Filterbank zu trennen.

Spracherkennung mit MFCCs kann in verschiedenen Umgebungen effektiv eingesetzt werden [9], und diese Methode kann in Kombination mit Deep-Learning-Modellen eine hohe Genauigkeit erreichen . Daher dient MFCC als entscheidender Input für Deep-Learning-Modelle zur Spracherkennung [10].

Machine Translated by Google

4. Vorgeschlagene

Methode In diesem Abschnitt schlagen wir eine Methode zur effizienten Durchführung einer Dimensionsreduzierung von Sprachsignalen unter Verwendung komplexwertiger MFCCs und komplexwertiger neuro Abbildung 2 bietet einen Überblick über den vorgeschlagenen Ansatz. Diese Abbildung veranschaulicht den vorgeschlagenen Ansatz zur Sprachklassifizierung mithilfe komplexwertiger neuronaler Netzwerke. Der Prozess beginnt mit der Spracheingabe, gefolgt von der Merkmalsextraktion mithilfe von Mel-Frequency- Cepstrum-Koeffizienten (MFCCs). Diese Merkmale werden dann einer Dimensionsreduzierung unterzogen und in komplexwertige Darstellungen u Die verarbeiteten Merkmale werden anschließend in ein komplexwertiges neuronales Netzwerk eingegeben. Die Leistung des Netzwerks wird durch die Auswahl geeigneter Aktivierungsfunktionen und Lernalgorithmen verbessert. Schließlich wird die Ausgabe des neuronalen Netzwerks für Klassifizierungszwecke verwendet.

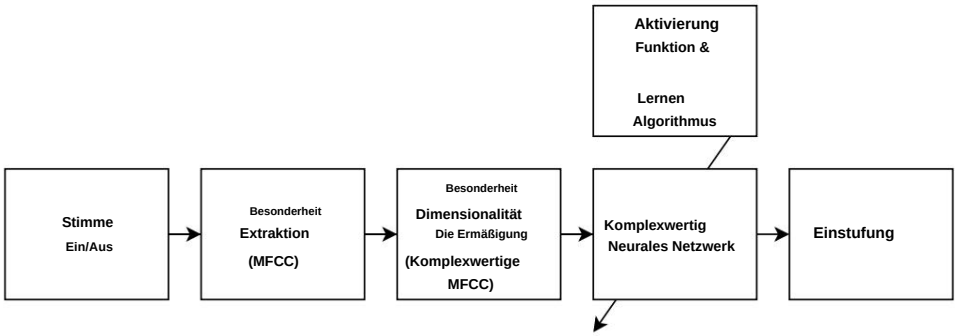


Abbildung 2. Übersicht über den vorgeschlagenen Ansatz.

Komplexwertige MFCCs: Die Anzahl der bei der Spracherkennung verwendeten MFCCs kann je nach Anwendung und Modell variieren. Normalerweise werden jedoch MFCCs im Bereich von 13 bis 40 verwendet. Dieser Bereich funktioniert nachweislich in verschiedenen Spracherkennungssystemen effektiv . Im Allgemeinen kann eine höhere Anzahl von MFCCs dem Modell ermöglichen, mehr Frequenz- und Spektraldetails zu erfassen, aber dies erhöht auch den Re Daher wird die ausgewählte Anzahl von MFCCs unter Berücksichtigung eines Gleichgewichts zwischen Rechenleistung und Modellleistung bestimmt.

Bei der Darstellung physikalischer Probleme mit komplexen Zahlen gibt es zwei Hauptmethoden . Eine davon ist eine einfache Substitutionsmethode und die andere ist ein eleganterer Ansatz, der als mathematische Äquivalenz bekannt ist, wie die Fourier-Transformation. Bei der Substitution werden zwei reale physikalische Parameter verwendet und einer davon in den Realteil der komplexen Zahl und der andere in den Imaginärteil gesetzt. Dadurch können die beiden Werte als eine einzige Entität, d. h. eine einzige komplexe Zahl, verarbeitet werden **[11]. Die in diesem Dokument zur Dimensionsreduzierung vorgeschlagene Methode umfasst die Umwandlung von n MFCC-Werten in komplexe Zahlen. Der erste MFCC wird als Re Anschließend wechselt jeder MFCC zwischen dem Real- und Imaginärteil innerhalb der komplexen Zahl. Dieses Muster setzt sich bis zum n-ten MFCC fort, was zu n/2 komplexen Zahlen führt. Wenn wir beispielsweise** n **MFCC-Werte haben, könnten diese wie folgt lauten:**

[x1, x2, x3, x4, . . . , xnÿ1, xn]

In diesem Fall wird der erste MFCC x1 als Realteil der ersten komplexen Zahl verwendet. Der zweite MFCC x2 wird als Imaginärteil der ersten komplexen Zahl verwendet. Der dritte MFCC x3 wird als Realteil der zweiten komplexen Zahl verwendet. Der vierte MFCC x4 wird als Imaginärteil der zweiten komplexen Zahl verwendet. Dieses Muster wechselt zwischen Real- und Imaginärteilen bis zum n-ten MFCC xn, **der zum Imaginärteil der** (n/2)-ten **komplexen Zahl wird. Daher werden mit n MFCC-Werten n/2 komplexe Zahlen erzeugt. Die erste komplexe Zahl ist x1** + **jx2, die zweite komplexe**

Machine Translated by Google

Zahl ist x3 + **jx4, und so weiter bis zur letzten komplexen Zahl, die** xnÿ1 + jxn ist. **Wir haben diese Erklärung in Abbildung 3 illustriert.**

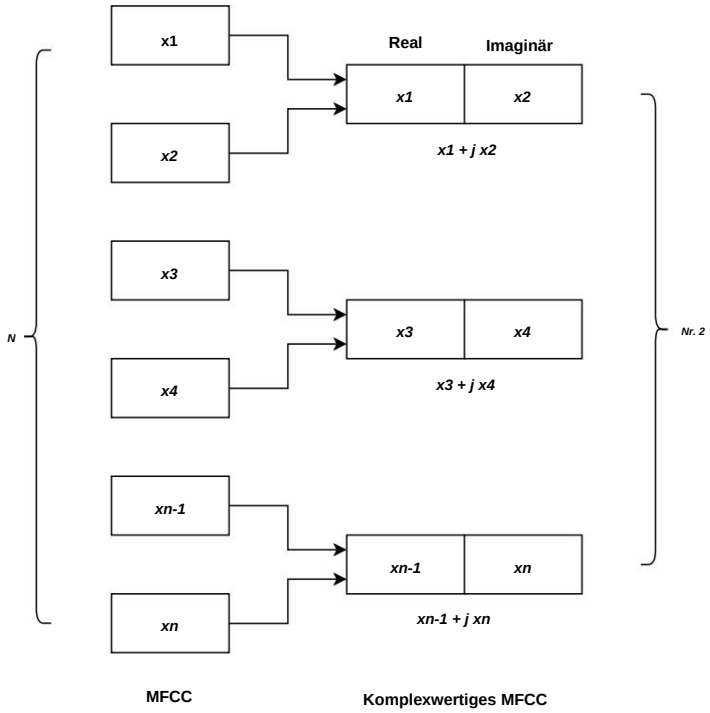


Abbildung 3. Prozess der Generierung komplexwertiger MFCCs.

Komplexwertiges neuronales Netzwerk: Komplexwertige neuronale Netzwerke sind neuronale Netzwerkmodelle, die komplexe Zahlen als Eingabewerte oder Gewichte verwenden, und sie haben den Vorteil, dass sie im Vergleich zu realwertigen neuronalen Netzwerken umfangreichere Informationen verarbeiten können. Darüber hinaus wurden in verschiedenen Bereichen verschiedene Forschungsstudien durchgeführt, und komplexwertige Methoden zur Dimensionsreduzierung oder Klassifizierung und Erkennung mithilfe komplexer Zahlen werden effektiv eingesetzt. Zu diesem Zweck wird die Multiplikation komplexer Eingaben und komplexer Gewichte durchgeführt, was eine Rotation in der komplexen Ebene impliziert und somit Phasenänderungen einführt [5]. Die Summe der Multiplikation von Eingabe X und Gewicht W kann als Gleic

$$Z = \mathbf{W0} + W1X1 + W2X2 + \ldots + WnXn \quad \textbf{(1)}$$

$$Z = \ddot{y}n\ Wn \cdot Xn \quad \textbf{(2)}$$

Aktivierungsfunktion und Lernalgorithmus: In komplexwertigen neuronalen Netzwerken wird eine phasenabhängige Aktivierungsfunktion verwendet, die eine Normalisierung der Betragsinformation auf einen Einheitskreis von 1 und nur die Extraktion der Phaseninformation erzwingt, wie in Gleichung (3) gezeigt.

$$P(Z) = \frac{Z}{|Z|} = e^{j \cdot \arg(Z)} \quad \textbf{(3)}$$

Der Lernalgorithmus komplexwertiger neuronaler Netze funktioniert wie folgt: Wenn ein gewünschtes Ziel t und ein Ausgabewert z vorliegen, wie in Abbildung 4 dargestellt, werden die Gewichte gemäß Gleichung (4) angepasst , **um sich schrittweise in Richtung des gewünschten Ziels zu be**

Machine Translated by Google

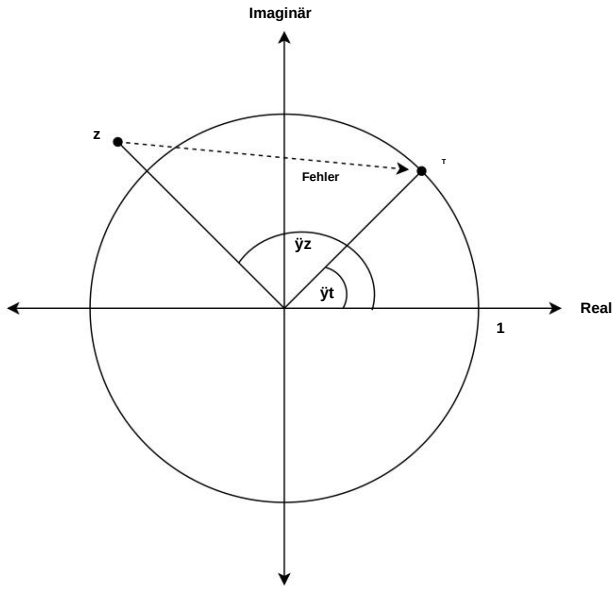


Abbildung 4. Lernalgorithmus für CVNN.

$$t = ÿn\ (Wn + ÿWn)\cdot Xn \tag{4}$$

Der Fehler in Bezug auf das Trainingsziel t **beträgt** e = t ÿ **z wie in Gleichung (5) gezeigt.**

$$e = ÿn\ (Wn + ÿWn)\cdot Xn\ ÿ\ ÿn\ (Wn\ )\cdot Xn = ÿn\ (ÿWn)\cdot Xn \tag{5}$$

Um die Berechnung des Fehlers in Bezug auf Gewichtsänderungen zu vereinfachen und zu beschleunigen, wird unter der Annahme, dass jeder Knoten gleichermaßen zum Fehler beiträgt, das Summationssymbol *ÿ* **aus der Gleichung entfernt und beide Seiten werden zur Vereinfachung der Berechnung durch die Anzahl der beitragenden Knoten N im komplexwertigen neuronalen Netzwerk geteilt. Die resultierende Gleichung ist Gleichung (6).**

$$ÿWn \cdot Xn = \frac{t}{N} \tag{6}$$

Um ÿWn **zu erhalten und damit die Gewichte zu ändern, multiplizieren Sie beide Seiten der** $X_N^{ÿ1}$ $X_N^{ÿ1}$ **Gleichung (6) mit Xn , wenn Sie es mit der komplex konjugierten** Zahl Xn multiplizieren, **wobei n sowohl Zähler als auch Nenner ist. Somit ist** $X_N^{ÿ1} = \frac{\overline{Xn}}{|Xn|}$ **wobei** |Xn| = **1 für** Xn **innerhalb der Einheit Kreis. Daher ist** $X_N^{ÿ1} = \overline{Xn}$, **und Gleichung (6)** wird erneut als Gleichungen **(7) und (8) hergeleitet .**

$$ÿWn \cdot Xn \cdot X_N^{ÿ1} = \frac{t}{N} \cdot X_N^{ÿ1} = \frac{t}{N} \cdot \overline{Xn} \tag{7}$$

$$ÿWn = \frac{t}{N} \cdot \overline{Xn} \tag{8}$$

Auf diese Weise kann die erzielte Änderung der Gewichte in Gleichung (8) beobachtet werden. Diese Änderung stellt die Anpassung der komplexen Gewichte für jeden Knoten dar, der zum Fehler in Bezug auf das Trainingsziel beiträgt, und wird einheitlich auf das gesamte Netzwerk angewendet.

Durch diesen Prozess passen sich komplexwertige neuronale Netzwerke unter Verwendung phasenabhängiger Aktivierungsfunktionen und effizienter Lernalgorithmen näher an das Ziel an, während sie reichhaltige Informationen durch komplexe Multiplikationsoperationen zwischen Eingaben und G

Komplexe Domänen-Backpropagation: Bei der Verwendung von versteckten Schichten wird der Fehler von der Ausgabeschicht an die versteckten Schichten zurückpropagiert, um zu berechnen, wie viel jeder

Machine Translated by Google

in den verborgenen Schichten haben zu diesem Fehler beigetragen. Darauf basierend müssen die Gewichte zwischen den verborgenen und den Ausgabeschichten entsprechend angepasst werden. Dieser Prozess wird als Fehlerrückausbreitung bezeichnet , durch den das neuronale Netzwerk die Gewichte anpassen kann, um den Fehler in den Ausgabewerten während des Lernens zu minimieren . Die Struktur eines mehrschichtigen komplexwertigen neuronalen Netzwerks ist in Abbildung 5 dargestellt.

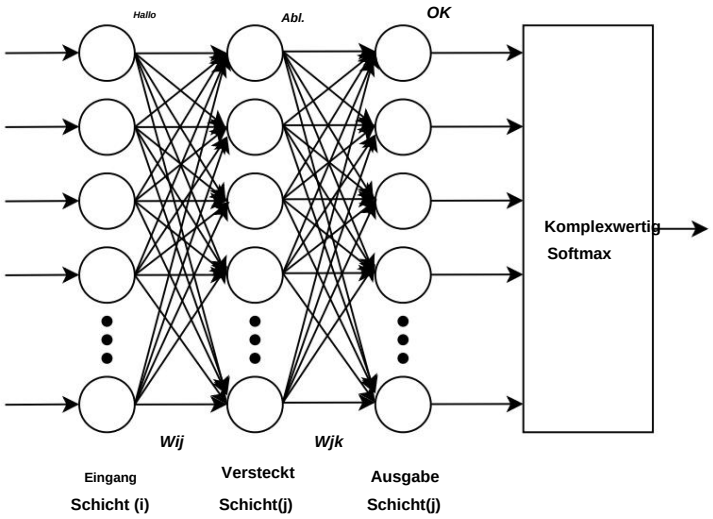


Abbildung 5. Mehrschichtiges CVNN.

In diesem Artikel wird eine Fehlerkorrekturmethode für die Ausgabeschicht k verwendet, um die Gewichte zu aktualisieren, während eine Fehlerrückpropagationsmethode für die verborgene Schicht j verwendet wird, um ihren Einfluss auf die Ausgabe zu untersuchen. Mit anderen Worten: Die Ausgabeschicht k verwendet die Fehlerkorrekturmethode, um das Erreichen lokaler Minima zu vermeiden, während die Fehlerrückpropagationsmethode auf die verborgene Schicht j angewendet wird, um eine bessere Verteilung der Fehler von der Ausgabeschicht auf die verborgene Schicht zu gewährleisten. Diese Kombination aus Fehlerrückpropagation in der verborgenen Schicht und Fehlerkorrektur in der Ausgabeschicht dürfte das Problem lösen, dass das gesamte neuronale Netzwerk in lokalen Minima hängen bleibt, Die Fehler-Backpropagation-Methode für komplexwertige neuronale Netzwerke ist wie folgt.

Bei der Verwendung des Gradientenabstiegs zur Gewichtsaktualisierung in komplexwertigen neuronalen Netzen wird die Wirtinger-Rechnung zur Differenzierung komplexer Funktionen verwendet. Da komplexe Funktionen aus Real- und Imaginärteilen bestehen, führt die Wirtinger-Rechnung die Differenzierung getrennt für Real- und Imaginärteile durch. Für eine komplexe Zahl z = x + jy , **die mit Real- und Imaginärteilen** x **bzw. y definiert ist, und eine komplexe Funktion** f(z) = u(x, y) + jv(x, y) werden **die Wirtinger-Ableitungen wie folgt definiert [14].**

$$\frac{\partial f}{\partial z} = \frac{1}{2}\left(\frac{\partial f}{\partial x} - j\frac{\partial f}{\partial y}\right) = \frac{1}{2}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + j\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)\right) \quad (9)$$

$$\frac{\partial f}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y}\right) = \frac{1}{2}\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} + j\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)\right) \quad (10)$$

$$+ \partial z \, \partial \bar{z} \, \frac{\partial f}{\partial \bar{z}} \partial f = \quad (11)$$

In diesem Zusammenhang bezeichnet z **die komplex konjugierte Funktion von z. Bei der Wirtinger-Ableitungsmethode können komplexe Funktionen als Erweiterungen reeller Funktionen behandelt werden, indem man sie zur Differenzierung in Real- und Imaginärteile aufteilt. Im Folgenden wird der Aktualisierungsprozess komplexer Variablen unter Verwendung der Schrittweite ÿ/2 und der Verlustfunktion L** beso

Machine Translated by Google

$$xn+1 = xn \ddot{y} \frac{\ddot{y}}{2} \frac{\ddot{y}L}{\ddot{y}x} \quad (12)$$

$$yn+1 = yn \ddot{y} \frac{\ddot{y}}{2} \frac{\ddot{y}L}{\ddot{y}y} \quad (13)$$

Durch Einsetzen von zn+1 = xn+1 + j yn+1 **in die Gleichungen (12) und (13) und anschließendes Einsetzen von die Wirtinger-Ableitung am Ende des Ableitungsprozesses, sieht sie wie folgt aus.**

$$zn+1 = xn \ddot{y} \frac{\ddot{y}}{2} \frac{\ddot{y}L}{} + j\, yn \ddot{y} \ddot{y}x \frac{\ddot{y}}{2} \frac{\ddot{y}L}{\ddot{y}y}$$

$$= xn \ddot{y} \frac{\ddot{y}}{2} \frac{\ddot{y}L}{} + jyn \ddot{y} j \ddot{y}x \frac{\ddot{y}}{2} \frac{\ddot{y}L}{\ddot{y}y}$$

$$= xn + jyn \ddot{y} \frac{\ddot{y}}{2} \frac{\ddot{y}L \ddot{y}L + j \ddot{y}x}{\ddot{y}y}$$

$$\mathbf{1} = zn \ddot{y} \ddot{y} \mathbf{2} \frac{\ddot{y}L \ddot{y}L + j \ddot{y}x}{\ddot{y}y}$$

$$= zn \ddot{y} \ddot{y} \ddot{y}z \frac{\ddot{y}L}{} \quad (14)$$

Die partielle Ableitung von L in Bezug auf die komplex konjugierte Wirtinger-Ableitung. $\frac{\ddot{y}L}{\ddot{y}z}$ **bezeichnet Dies zeigt, wie die konjugierte Wirtinger- Ableitung verwendet werden kann, um die Aktualisierungsformel für komplexe Variablen zu vereinfachen. Bei Verwendung der obigen Methode ergibt das Umschreiben von Gleichung (5) als Ausgabeschichtfehler Gleichung (15), wobei** tk **den Zielwert und** ok **den tatsächlichen Ausgabewert der Ausgabeschicht k** darstellt .

$$ek = tk \ddot{y} ok \quad (15)$$

Und die Verlustfunktion wird als Gleichung (16) definiert.

$$L = ek^{2} = |tk \ddot{y} ok| \overline{|tk \ddot{y} ok|} \quad (16)$$

Bei großen Gewichten ist es wahrscheinlicher, dass sich der Fehler auf die verborgenen Schichten ausbreitet. Der verborgene Schichtfehler ej **kann durch die folgende Gleichung ausgedrückt werden.**

$$ej = WTjk \cdot ek \quad (17)$$

Mit dem Ergebnis der Gleichung (14) lässt sich **der Gradient der Fehlerfunktion in Bezug auf das Gewicht** Wij **zwischen der Eingabeschicht** i **und der verborgenen Schicht** j **wie folgt ausdrücken. Dabei ist** oi **der Ausgabewert der Eingabeschicht i und** oj **der Ausgabewert der verborgenen Schicht j.**

$$\frac{\ddot{y}L}{\ddot{y}Wij} = \frac{\ddot{y}L}{\ddot{y}oj} \cdot \frac{\ddot{y}oj}{\ddot{y}Wij} \quad (18)$$

$$\frac{\ddot{y}L}{\ddot{y}Wij} \ddot{y}ej \cdot \frac{\ddot{y}}{\ddot{y}} P \ddot{y} Wij \cdot oi = \quad (19)$$

Machine Translated by Google

$$\frac{\ddot{y}L}{} = \ddot{y}ej \cdot P\,\dot{\ddot{y}}Wij \cdot \bar{o}h \quad (20)$$

In den Gleichungen (19) und (20) stellt P **die phasenabhängige Aktivierungsfunktion dar, die in der Ausgabe der verborgenen Schicht verwendet wird,** $\dot{\ddot{y}}$ **ist die Ableitung der abhängigen Aktivierungsfunktion.** und P. **Die Aktualisierung der Gewichte der verborgenen Schicht kann wie folgt ausgedrückt werden.**

$$\text{neues Wij} = \text{altes Wij } \ddot{y}\ lr \cdot \ddot{y}Wij \frac{\ddot{y}L}{} \quad (21)$$

Beim Umgang mit komplexen Variablen muss anstelle von $\ddot{y}Wij$ $\ddot{y}Wij$ **zum Aktualisieren verwendet werden. Dabei** spielt lr $\frac{\ddot{y}L}{}$ **eine Rolle bei der Anpassung der Intensität der Änderung, um ein Überschwingen zu verhindern, und wird als Lernrate bezeichnet. Darüber hinaus ist die Ableitung der Aktivierungsfunktion P** eine **komplexe Funktion, und um die Ableitung einer komplexen Funktion zu finden, wird die Wirtinger-Ableitung verwendet [15].**

Komplexwertiges Softmax: Die Softmax-Funktion, die als Aktivierungsfunktion in der Ausgabeschicht verwendet wird, nimmt Eingabewerte und normalisiert jeden Wert auf einen Wert zwischen 0 und 1. Sie stellt sicher, dass die Summe der normalisierten Werte immer 1 ist. Diese Funktion ermöglicht die Interpretation der Ergebnisse als Wahrscheinlichkeiten. Sie kann als Gleichung (22) ausgedrückt werden.

$$\text{Also f tmax(Z)i} = e\ Zi\ \ddot{y}\ i=1 \frac{e^{Zi}}{N} \quad (22)$$

Die Softmax-Funktion kann Ergebnisse jedoch nur als Wahrscheinlichkeiten für reelle Zahleneingaben interpretieren, und für komplexe Zahleneingaben ist eine andere Interpretationsmethode erforderlich. In diesem Artikel schlagen wir eine komplexwertige Softmax-Interpretationsmethode für komplexe Zahleneingaben vor. Das Ergebnis der Softmax-Funktion für komplexe Zahleneingaben ergibt komplexe Ausgabewerte, und da die Summe aller Ausgabewerte 1+0j ist, schlagen wir einen Algorithmus vor , der den Abstand zwischen den Ausgabewerten und 1 + 0j **in der komplexen Ebene berechnet und den Index des nächstgelegenen Werts auswählt. Dies kann als Gleichung (24) ausgedrückt werden.**

$$\text{distancei} = |\text{So f tmax(Z)i } \ddot{y}\ \ddot{y}| \quad (23)$$

$$\text{nächstgelegener_Index} = \text{argmin(Entfernung i)} \quad (24)$$

Hier ist $\ddot{y}$ = **1** + 0j **definiert. Folglich entscheiden wir uns, die komplexen Zahlen entspricht closest_index als Klasse 1.**

5. Leistungsanalyse In unseren

Experimenten zur Sprachsignalerkennung wurde ein detaillierter Vergleich zwischen den Ergebnissen von realwertigen neuronalen Netzwerken und komplexwertigen neuronalen Netzwerken durchgeführt.
Die Leistungsanalyse wurde parallel zur Vorverarbeitung der komplexwertigen Mel- Frequency-Cepstrum-Koeffizienten (MFCCs) zur Dimensionsreduzierung durchgeführt. Zusätzlich wurde der Einfluss der komplexwertigen MFCCs auf die Ergebnisse der Spracherkennung beobachtet. Insbesondere bei der Erkennung vorgegebener Sprachsignale mithilfe neuronaler Netzwerke wurde der Einfluss der MFCC-Werte und ihr Komplexifizierungseffekt auf die Modellleistung ausgewertet und verglichen, um zu ermitteln, ob eine Leistungssteigerung bei gleichzeitiger Reduzierung der Parameteranzahl möglich ist .

Experimentelle Methode: Abbildung 6 zeigt die Architektur eines mehrschichtigen komplexwertigen neuronalen Netzwerks, das zum Trainieren und Testen mit einem Datensatz aus 8000 kurzen Sprachproben verwendet wird, von denen jede acht unterschiedliche Wörter enthält und weniger als eine Sekunde dauert. Die in dieser Studie verwendeten Sprachdaten sind Teil des von Google gesammelten und unter der CC BY-Lizenz verbreiteten Datensatzes von Sprachbefehlen.

Machine Translated by Google

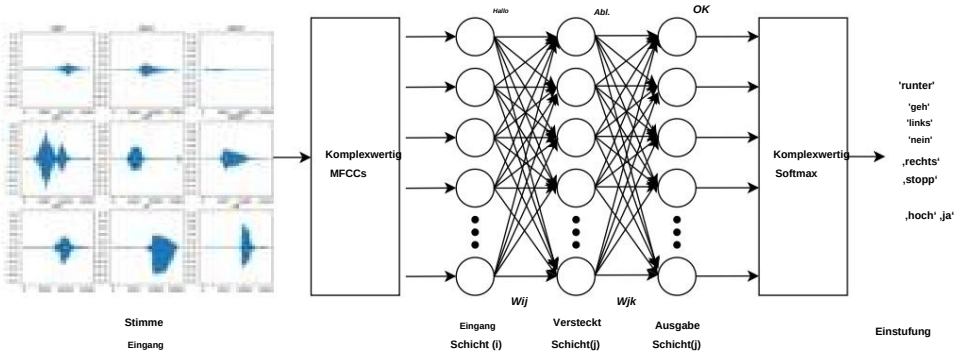


Abbildung 6. Experimenteller Aufbau eines mehrschichtigen CVNN.

Komplexwertige neuronale Netze sind strukturell realwertigen neuronalen Netzen ähnlich, der Unterschied liegt jedoch in der Verwendung komplexer Zahlen, die Phasenänderungen einführen, da die Summe der Produkte zwischen Eingaben und Gewichten in der komplexen Ebene rotiert. Diese Phasenänderungen spielen in komplexwertigen neuronalen Netzen eine wichtige Rolle. Das komplexwertige neuronale Netzwerk besteht aus einer Eingabeschicht mit dimensionsreduzierten komplexwertigen MFCCs, die aus den Sprachdaten durch den in Abbildung 3 gezeigten Komplexifizierungsprozess gewonnen wurden, einer verborgenen Schicht mit 500 Knoten und einer Ausgabeschicht mit 8 Knoten zur Klassifizierung. Alle Knoten und Gewichte verwenden komplexe Zahlen. Die endgültigen Ergebnisse wurden in der Ausgabeschicht klassifiziert. Der Datensatz umfasste 7000 Proben zum Training und 1000 Proben zum Testen. Als Daten wurden Audioclips mit acht Wörtern verwendet und die Experimente wurden durchgeführt, indem die Eingabedaten danach aufgeteilt wurden, ob eine komplexwertige MFCC-Transformation angewendet wurde oder nicht. Im Fall der Verwendung realwertiger MFCCs wurde ein realwertiges neuronales Netzwerk eingesetzt, während ein komplexwertiges neuronales Netzwerk mit komplexwertigen MFCCs verwendet wurde. Sowohl die realwertigen als auch die komplexwertigen neuronalen Netzwerke wurden unter denselben Bedingungen trainiert: lr = 0,01, Batchgröße = 50 und Anzahl der Epochen = 50. Für die Eingabedaten wurden die Ergebnisse von drei verschiedenen Vorverarbeitungsschritten zur Dimensionsreduzierung mit dem realwertigen neuronalen Netzwerk verglichen. Diese Vorverarbeitungsschritte umfassten MFCC- Dimensionen von 20, 10 und 5. Die Länge des Eingangssignals für das neuronale Netzwerk kann wie in Gleichung (25) gezeigt bestimmt werden

$$\text{Die Länge der Eingangssignaldaten} = \text{MFCCs} \times \text{Frames} \quad (25)$$

Beispielsweise wird im Fall des komplexwertigen MFCC 20 zuerst der realwertige MFCC 40 berechnet. Mit Gleichung (25) ergibt sich für den realwertigen MFCC 40 eine Länge des Eingangssignals von 40 × 32 = 1280. Wenn wir diese 40 MFCCs dann in komplexe Zahlen umwandeln, wie in Abbildung 3 dargestellt, wird eine Dimensionsreduzierung auf die Hälfte der Länge erreicht, was zu 16 Frames führt. Folglich wird die Eingangssignallänge 40 × 16 = 640, was **der Eingangssignallänge des realwertigen MFCC 20 entspricht .** Eine ähnliche Dimensionsreduzierung kann **mit derselben Methode für den komplexwertigen MFCC 10 und den komplexwertigen MFCC 5 durchgeführt werden . Die Gesamtzahl der Parameter für das gesamte neuronale Netzwerk, die durch diesen Prozess erhalten wurde, ist in Tabelle 1 aufgeführt. In dieser Tabelle stellt I(n) die Anzahl der Neuronen in der Eingabeschicht dar, H(n) die Anzahl der Neuronen in der verborgenen Schicht und O(n) die Anzahl der Neuronen in der Ausgabeschicht.**

Tabelle 1. Gesamtzahl der Parameter im neuronalen Netzwerk.

MFCC	Realwertiger MFCC	Komplexwertiges MFCC
20	I(20 × 32 = 640) × H(500) + H(500) × O(8) = 324.000	I(40 × 16 = 640) × H(500) + H(500) × O(8) = 324.000
10	I(10 × 32 = 320) × H(500) + H(500) × O(8) = 164.000	I(20 × 16 = 320) × H(500) + H(500) × O(8) = 164.000
5	I(5 × 32 = 160) × H(500) + H(500) × O(8) = 84.000	I(10 × 16 = 160) × H(500) + H(500) × O(8) = 84.000

Machine Translated by Google

Komplexwertiger MFCC 20 vs. realwertiger MFCC 20: Wie in Abbildung 7 dargestellt, erreichte der komplexwertige MFCC 20 eine maximale Trainingsgenauigkeit von 99,40 % und eine Testgenauigkeit von 84,90 %, obwohl er die gleiche Eingabedatenlänge wie der realwertige MFCC 20 hatte, der eine maximale Trainingsgenauigkeit von 98,13 % und eine Testgenauigkeit von 83,10 % erreichte. Die mit komplexwertigen MFCCs erzielten Ergebnisse übertrafen diejenigen mit realwertigen MFCCs sowohl hinsichtlich der Trainings- als auch der Testgenauigkeit.

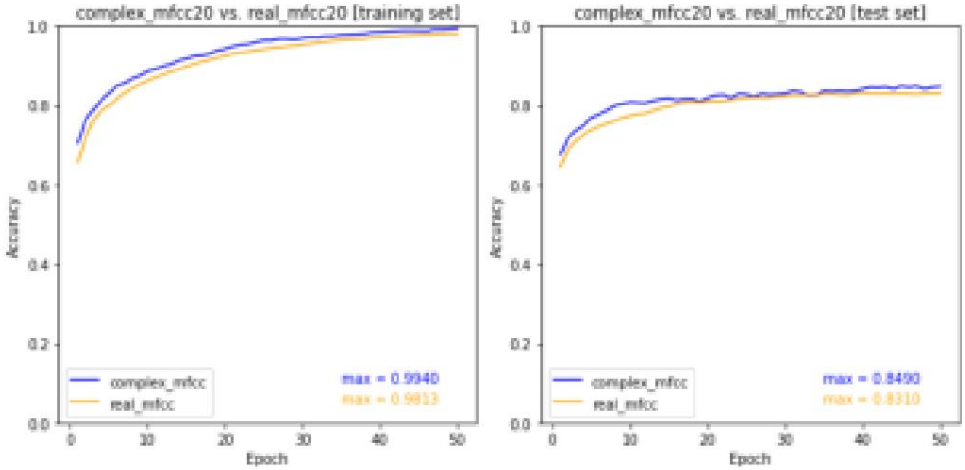


Abbildung 7. Vergleich der Ergebnisse des komplexwertigen MFCC 20 und des realwertigen MFCC 20.

Komplexwertiger MFCC 10 vs. realwertiger MFCC 10: Wie in Abbildung 8 dargestellt, erreichte der komplexwertige MFCC 10 eine maximale Trainingsgenauigkeit von 96,90 % und eine Testgenauigkeit von 83,80 %, obwohl er die gleiche Eingabedatenlänge wie der realwertige MFCC 10 hatte, der eine maximale Trainingsgenauigkeit von 94,24 % und eine Testgenauigkeit von 81,20 % erreichte. Die mit komplexwertigen MFCCs erzielten Ergebnisse übertrafen diejenigen mit realwertigen MFCCs sowohl hinsichtlich der Trainings- als auch der Testgenauigkeit.

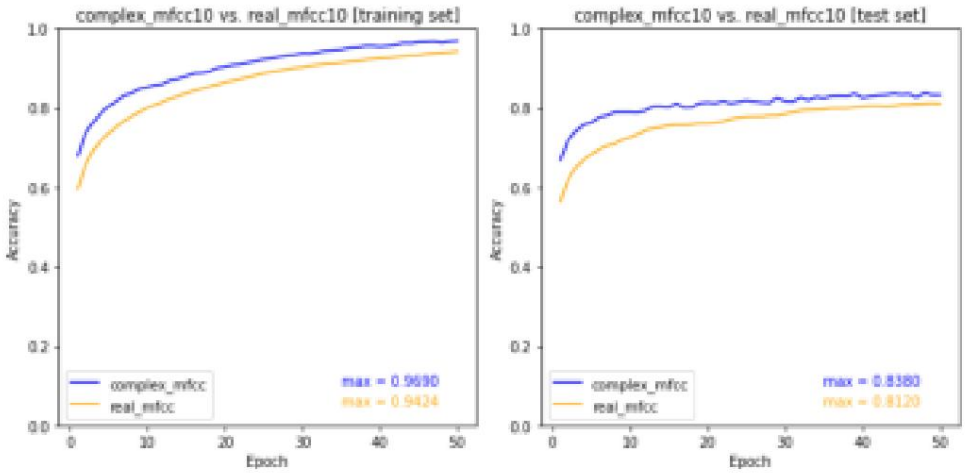


Abbildung 8. Vergleich der Ergebnisse des komplexwertigen MFCC 10 und des realwertigen MFCC 10.

Komplexwertiger MFCC5 vs. realwertiger MFCC5: Wie in Abbildung 9 dargestellt, erreichte der komplexwertige MFCC 5 eine maximale Trainingsgenauigkeit von 92,33 % und eine Testgenauigkeit von 81,40 %, obwohl er die gleiche Eingabedatenlänge hatte wie der realwertige MFCC 5, der eine maximale Trainingsgenauigkeit von 86,34 % und eine Testgenauigkeit von 74,90 % erreichte. Die mit komplexwertigen MFCCs erzielten Ergebnisse übertrafen diejenigen mit realwertigen MFCCs sowohl hinsichtlich der Trainings- als auch der Testger

Machine Translated by Google

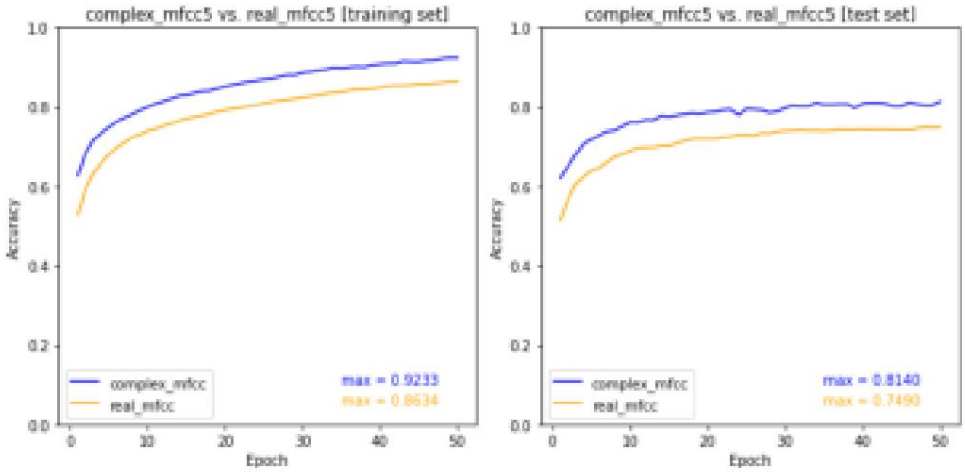


Abbildung 9. Vergleich der Ergebnisse von komplexwertigem MFCC 5 und realwertigem MFCC 5.

Gesamtergebnisse der Analyse: Bei der Überprüfung der experimentellen Ergebnisse wurde festgestellt, dass die Verwendung von MFCCs mit komplexen Werten im Vergleich zur Verwendung von MFCCs mit realen Werten eine höhere Leistung bringt, selbst bei gleicher Eingabedatenlänge. Daher wurde bestätigt, dass die Verwendung von MFCCs mit komplexen Werten und komplexwertigen neuronalen Netzwerken zu besseren Ergebnissen führt . Die Variation der Genauigkeit je nach Anzahl der MFCC-Funktionen ist **in Abbildung 10 dargestellt. Obwohl hier nicht gezeigt, wurden mit dem F1-Score dies**

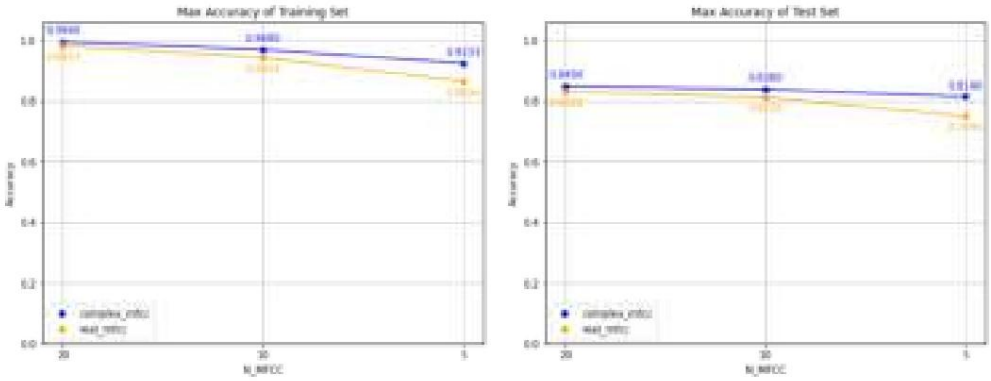


Abbildung 10. Genauigkeitsvariation mit der Anzahl der MFCCs.

Wenn eine ausreichende Anzahl von MFCCs vorhanden ist, scheint sich die Leistung bei Verwendung reeller Zahlen mit zunehmenden Epochen der Leistung bei Verwendung komplexer Zahlen anz Abbildung 9 zeigt jedoch ein anderes Ergebnis. Komplexwertiger MFCC 5 verfügt über 5 dimensionsreduzierte komplexe MFCCs, die durch den in Abbildung 3 gezeigten Komplexifizierungsprozess aus 10 MFCCs gewonnen wurden. Ein realwertiger MFCC 5 mit nur 5 MFCCs kann komplexwertigen MFCC 5 auch mit zunehmenden Epochen nicht einholen. Mit anderen Worten: Wenn die Anzahl der MFCCs gering ist, erreichen die Ergebnisse realwertiger MFCCs selbst mit zunehmenden Epochen nicht die Leistung komplexwertiger MFCCs. Wie in Abbildung 10 gezeigt, sind die Ergebnisse von komplexwertigem MFCC 5 denen von realwertigem MFCC 5 trotz derselben Anzahl von Parametern überlegen. Es wurde also beobachtet, dass bei Verwendung komplexwertiger MFCCs **die Änderung der Genauigkeit mit abnehmender Anzahl von MFCC-Funktionen gleichmäßiger verläuft als bei Verwendung realwertiger MFCCs.**

6. Schlussfolgerungen

In diesem Artikel haben wir komplexwertige neuronale Netzwerke eingeführt, um das Problem der Spracherkennung zu lösen. Unsere Experimente nutzten komplexwertige MFCCs als Eingab

Machine Translated by Google

komplexwertiges neuronales Netzwerk. Die Ergebnisse zeigten eine bessere Leistung beim Einsatz komplexwertiger MFCCs, mit einer allmählicheren Änderung der Genauigkeit bei abnehmender Anzahl der MFCC-Funktionen im Vergleich zur Verwendung realwertiger MFCCs. Dies deutet darauf hin, dass komplexwertige MFCCs dazu beitragen können, die Leistung mit weniger Funktionen aufrechtzuerhalten. Diese Erkenntnisse liefern wertvolle Einblicke in das Potenzial komplexwertiger neuronaler Netzwerke und zeigen die Wirksamkeit der Kombination komplexwertiger MFCCs mit komplexwertigen neuronalen Netzwerken bei der Spracherkennung.

Trotz dieser vielversprechenden Ergebnisse sind wir uns der gewissen Einschränkungen unseres Ansatzes bewusst . Diese Studie wurde mit einem begrenzten Datensatz durchgeführt, der möglicherweise nicht die gesamte Vielfalt realer Sprachmuster abbildet. Darüber hinaus ist die in dieser Studie beschriebene Architektur des komplexwertigen neuronalen Netzwerks relativ einfach und erfasst möglicherweise nicht alle Feinheiten von Sprachsignalen.

Mit Blick auf die Zukunft möchten wir auf dieser Arbeit aufbauen, um die Genauigkeit und Leistung von Sprachsignalverarbeitungs- und -erkennungsmodellen weiter zu verbessern. Unsere zukünftige Forschung wird sich auf die Entwicklung ausgefeilterer komplexwertiger neuronaler Netzwerkmodelle und die Durchführung von Experimenten mit unterschiedlichen Datensätzen zur Verbesserung der Leistung konzentrieren. Darüber hinaus planen wir, die Anwendung komplexwertiger neuronaler Netzwerke in anderen Bereichen als der Spracherkennung zu untersuchen und ihren Einsatz möglicherweise auf verschiedene Be

Beiträge der Autoren: Konzeptualisierung, SK; Projektverwaltung, MP; Überwachung, MP; Schreiben – Originalentwurf, SK. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Arbeit wurde gemeinsam von der National Research Foundation of Korea (NRF) über ein Stipendium der koreanischen Regierung (MSIT) (Stipendium-Nr. NRF-2023R1A2C1005461) und vom MSIT (Ministerium für Wissenschaft und IKT), Korea, im Rahmen des Convergence Security Core Talent Training Business Support Program (IITP-2024-RS-2024-00426853) unterstützt, das vom IITP (Institute of Information & Communications Technology Planning & Evaluation) beaufsichtigt wird .

Datenverfügbarkeitserklärung: Die in der Studie vorgestellten Originalbeiträge sind im Artikel enthalten. Weitere Anfragen können an die entsprechenden Autoren gerichtet werden.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte vorliegen.

Verweise

1. **Tebelskis, J. Spracherkennung mit neuronalen Netzwerken. Dissertation, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA, 1995.**
2. **Sarroff, AM Komplexe neuronale Netzwerke für Audio. Doktorarbeit, Dartmouth College, Hanover, NH, USA, 2018.**
3. **Abdul, ZK; Al-Talabani, AK Mel Frequency Cepstral Coefficient und seine Anwendungen: Eine Übersicht.** IEEE Access **2022,** 10, **122136–122158. [CrossRef]**
4. **Barrachina, JA; Ren, C.; Vieillard, G.; Morisseau, C.; Ovarlez, JP Theorie und Implementierung komplexwertiger neuronaler Netzwerke.** arXiv **2023, arXiv:2302.08286.**
5. **Aizenberg, I.** Komplexwertige neuronale Netze mit mehrwertigen Neuronen; **Springer: Berlin/Heidelberg, Deutschland, 2016.**
6. **Graves, A.; Jaitly, N.; Mohamed, AR Hybride Spracherkennung mit Deep Bidirectional LSTM. In Proceedings des 2013 IEEE Workshop on Automatic Speech Recognition and Understanding, Olomouc, Tschechische Republik, 8.–12. Dezember 2013; S. 273–278. [Querverweis]**
7. **Paz, P.; Garrido, M. Effiziente Implementierung komplexer Multiplikatoren auf FPGAs mit DSP-Slices.** J. Signal Process. Syst. **2023,** 95, **543–550. [CrossRef]**
8. **Ahmad, M.; Zhang, L.; Chowdhury, MEH FPGA-Implementierung eines komplexwertigen neuronalen Netzwerks für polar dargestellte Bildklassifizierung.** Sensors **2024,** 23, **2627. [CrossRef] [PubMed]**
9. **Anggraeni, D.; Sanjaya, WS; Nurasyidiek, MY; Munawwaroh, M. Die Implementierung der Spracherkennung mithilfe von Mel- Frequency-Cepstrum-Koeffizienten (MFCC) und der auf Python basierenden Support Vector Machine (SVM)-Methode zur Steuerung eines Roboterarms.** IOP Conf. Ser. Mater. Sci. Eng. **2018,** 288, **012042. [CrossRef]**
10. **Dhanjal, AS; Singh, W.** Eine umfassende Untersuchung zur automatischen Spracherkennung mithilfe neuronaler Netzwerke; **Springer Nature: Berlin/ Heidelberg, Deutschland, 2023.**
11. **Smith, SW** The Scientist and Engineer's Guide to Digital Signal Processing, **2. Auflage; California Technical Publishing: San Diego, CA,**

^ "Die Macht der Freiheit". In: Die Macht der Freiheit. Berlin 1999, S. 557–558.

12. **Bassey, J.; Li, X.; Qian, L. Eine Untersuchung komplexwertiger neuronaler Netzwerke.** arXiv **2021, arXiv:2101.12249v1.**
13. **MYONeuralNet. Komplexwertige neuronale Netzwerke – Experimente. 2016. Online verfügbar: http://makeyourownneuralnetwork. blogspot.com/2016/05/complex-valued-neural-networks.html (abgerufen am 19. Februar 2021).**

Machine Translated by Google

14. **Fischer, R. Wirtinger Calculus. In** Precoding and Signal Shaping for Digital Transmission; **John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2002; S. 405–413.**
15. **PyTorch. Automatisches Differenzierungspaket – torch.autograd, nd Online verfügbar: https://pytorch.org/docs/stable/notes/autograd.html (abgerufen am 3. November 2023).**
16. **Lee, C.; Hasegawa, H.; Gao, S. Komplexwertige neuronale Netzwerke: Eine umfassende Untersuchung.** IEEE/CAA J. Autom. Sin. **2022, 9, 1433–1454. [CrossRef]**

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Job-Empfehlungen: Benchmarking von Collaborative Filtering Methoden für Kleinanzeigen

Robert Kwieci ´nski 1,2, Tomasz Górecki 2,*, Agata Filipowska 3 und Viacheslav Dubrov 1

1 OLX Group, Królowej Jadwigi 43, 61-872 Posen, Polen; r.kwiecinskipl@gmail.com (RK); slavadubrov@gmail.com (VD)
2 Fakultät für Mathematik und Informatik, Adam-Mickiewicz-Universität, Uniwersytetu Pozna ´nskiego 4, 61-614 Pozna ´n, Polen; Institut
3 für Informationssysteme, Wirtschaftsuniversität Pozna ´n, Al. Niepodlegÿo´sci 10, 61-875 Pozna ´n, Polen; agata.filipowska@ue.poznan.pl
* Korrespondenz: tomasz.gorecki@amu.edu.pl

Zusammenfassung: Kleinanzeigen stellen zahlreiche Herausforderungen für Empfehlungsmethoden dar, darunter die temporäre Sichtbarkeit von Anzeigen, die Anonymität der meisten Benutzer und die Tatsache, dass normalerweise nur ein Benutzer einen beworbenen Artikel konsumieren kann. In dieser Arbeit begegnen wir diesen Herausforderungen, indem wir Modelle und Bewertungsverfahren auswählen, die als genau, vielfältig und effizient gelten (in Bezug auf Speicher und Zeitverbrauch während des Trainings und der Vorhersage). Ziel dieses Artikels ist es, verschiedene Empfehlungsmethoden für Stellenanzeigen am Beispiel von OLX Jobs zu vergleichen, um die Konversionsrate von Anzeigen und die Benutzerzufriedenheit zu verbessern. In unserer Forschung implementieren wir skalierbare Methoden und stellen verschiedene Ansätze für die Empfehlungen dar: Alternating Least Square (ALS), LightFM, Prod2Vec, RP3Beta und Sparse Linear Methods (SLIM). Wir führten A/B-Tests durch, indem wir Millionen von Nachrichten mit Empfehlungen versendeten, um Online-Bewertungen ausgewählter Methoden durchzuführen. Darüber hinaus haben wir den für unsere Forschung erstellten Datensatz veröffentlicht. Nach bestem Wissen ist dies der erste Datensatz dieser Art. Es enthält 65.502.201 Ereignisse, die auf OLX Jobs von 3.295.942 Benutzern durchgeführt wurden, die im Jahr 2020 innerhalb von zwei Wochen mit 185.395 Stellenanzeigen interagierten (sie anzeigten, beantworteten oder mit Lesezeichen versahen). Wir zeigen, dass RP3Beta, SLIM und ALS bei Tests in einer Laborumgebung deutlich besser abschneiden als Prod2Vec und LightFM. Online-A/B-Tests zeigen auch, dass das Senden von Nachrichten mit Empfehlungen, die von den ALS- und RP3Beta-Modellen generiert werden, die A

Schlüsselwörter: Stellenempfehlungen; Kleinanzeigen; kollaboratives Filtern; Skalierbarkeitstests; A/B-Tests; Empfehlungssysteme; Matrixfaktorisierung; graphenbasierte Empfehlungssysteme; OLX- Datensatz zu Stelleninteraktionen; Diversitätsmetriken

check for updates

Zitat: Kwieci ´nski, R.; Górecki, T.; Filipowska, A.; Dubrov, V. Job Empfehlungen: Benchmarking von kollaborativen Filtermethoden für Kleinanzeigen. Elektronik **2024**, 13, **3049**. https://doi.org/10.3390/Elektronik13153049

Wissenschaftliche Herausgeber: Marco Lombardi, Domenico Santaniello, Zilu Liang und Muhammad Khan

Empfangen: 30. Mai 2024
Überarbeitet: 24. Juli 2024
Akzeptiert: 27. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz. Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Einleitung

Online-Kleinanzeigen sind Websites, auf denen Inserenten Anzeigen zum Verkauf oder zur Vermietung von Dienstleistungen oder Produkten veröffentlichen [1]. In diesem Beitrag wird Online-Werbung am Beispiel von OLX Jobs behandelt.

Kleinanzeigen zielen darauf ab, sowohl die Bedürfnisse von Werbetreibenden als auch von Online-Werbungsempfängern zu erfüllen . Um diese Gruppen zu erreichen, werden bessere Empfehlungsmethoden immer wichtiger. Im Vergleich zum traditionellen E-Commerce (z. B. Amazon (https://www.amazon.com/, abgerufen am 26. Juni 2024) oder eBay (https://www.ebay.com/, abgerufen am 26. Juni 2024)) stehen Kleinanzeigen vor verschiedenen Herausforderungen, die wir in unserem Bewertungsverfahren berücksichtigen. Eine davon ist die zeitliche Sichtbarkeit von Anzeigen, die wir durch häufiges Neutrainieren Daher ist die Effizienz des Modells, über die wir berichten, für die Bewertung von entscheidender Bedeutung. Eine weitere Herausforderung besteht darin, dass ein Werbetreibender normalerweise nur einen Benutzer benötigt, um die Transaktion a Diesem Aspekt haben wir Rechnung getragen, indem wir bei unserer Bewertung auch eine Diversitätsanalyse einbezogen haben.

Machine Translated by Google

In unserer Forschung konzentrieren wir uns auf Ansätze, die sich aus der kollaborativen Filterung ergeben. Kollaborative Filtermodelle haben ihre Genauigkeit bewiesen, obwohl sie keine zusätzlichen Informationen über Benutzer oder Artikel verwenden, die in der Praxis manchmal schwer zu erhalten oder zu verarbeiten sind. Ziel dieses Artikels ist es, verschiedene Empfehlungsmethoden für den Einsatz im Kleinanzeigensektor zu bewerten. Die hohe Spärlichkeit unseres Datensatzes sowie die beträchtliche Anzahl von Anzeigen und Benutzern tragen zur Komplexität des Forschungsproblems bei.

Nachfolgend listen wir unsere wichtigsten Beiträge auf.

- Wir bewerten in einer Laborumgebung umfassend verschiedene, skalierbare und kollaborative Filteransätze hinsichtlich Genauigkeit, Vielfalt und Effizienz. Wir bewerten die Genauigkeit für verschiedene Benutzergruppen, abhängig von der Anzahl der Elemente, mit denen sie interagiert haben, zusammen mit einer statistischen Analyse der Unterschiede. Die vorgeschlagene Bewertungsmethode befasst sich mit den wichtigsten Herausforderungen von Kleinanzeigen.
- Wir präsentieren die Ergebnisse zweier Online-A/B-Tests, die die Auswirkungen der ALS- und RP3Beta-Modelle auf OLX-Benutzer demonstrieren. An jedem Test nahmen mehr als 1 Million Benutzer teil. Die berichteten Auswirkungen können verwendet werden, um die Bedeutung von Empfehlungssystemen im Kleinanzeigenbereich abzuschätzen und anderen Unternehmen zu ermöglichen, fundiertere Entscheidungen über die Entwicklung solcher Modelle zu treffen.
- Wir veröffentlichen und präsentieren einen Datensatz zu Jobinterak öffentlich verfügbarer Datensatz zu Interaktionen am Arbeitsplatz.

Dieses Dokument besteht aus sieben Abschnitten. Abschnitt 2 ist ein Literaturüberblick über die zugrunde liegenden Entwicklungen. Abschnitt 3 beschreibt den Datensatz, der entwickelt und auf Kaggle öffentlich zugänglich gemacht wurde. Das Design des Laborexperiments sowie die Implementierung der Methoden werden in Abschnitt 4 beschrieben. Abschnitt 5 konzentriert sich auf die Ergebnisse der Laborauswertung unserer Implementierungen sowie die statistische Bedeutung der Ergebnisse. Abschnitt 6 präsentiert die Ergebnisse der Online-Implementierungen, die während A/B-Tests mit Benutzern durchgeführt wurden, sowie eine Diskussion der Ergebnisse. Schlussfolgerungen werden in Abschnitt 7 präsentiert.

2. Verwandte Arbeiten

In diesem Abschnitt beschreiben wir zunächst die besonderen Merkmale der Bereiche Kleinanzeigen und Stellenangebote . Anschließend geben wir einen Überblick über Empfehlungssysteme. Als Nächstes diskutieren wir die Faktoren, die unsere Modellauswahl beeinflussen, und geben eine detaillierte Beschreibung jedes Modells. Abschließend geben wir die Forschungslücken an.

2.1. Kleinanzeigen-Websites

Kleinanzeigen-Websites (Classifieds) sind Online-Versionen von Zeitungsanzeigenteilen mit Anzeigen zu Fahrzeugen, Immobilien, Stellenangeboten, Haustieren usw. Kleinanzeigen sind eine besondere Art von E-Commerce-Website, da sie nicht den Abschluss von Transaktionen zwischen Parteien unterstützen – sie werden stattdessen verwendet, um über das Produkt oder die Dienstleistung zu kommunizieren und Treffen zu vereinbaren, während die Transaktion offline durchgeführt wird; dies unterscheidet sie von traditionellen E-Commerce-Websites wie eBay oder Amazon [2].

Unsere Forschung konzentriert sich auf OLX, das in über 30 Ländern weltweit vertreten ist. Einer der Anzeigentypen, die in Kleinanzeigen veröffentlicht werden, sind Stellenangebote. In unserer Untersuchung haben wir mehrere Merkmale identifiziert, die Stellenanzeigen in Kleinanzeigen beispielsweise von Auto- oder Immobilienanzeigen unterscheiden:

- Die Anforderungen in einer Anzeige geben den Benutzertyp an, der auf die Anzeige reagieren kann – nicht alle Benutzer verfügen über die erforderlichen Kompetenzen.
- Intensivere Einbindung des Nutzers während des Suchvorgangs, da die Suche nach einem Arbeitsplatz ein vorrangiges Bedürfnis darstellt, verglichen mit beispielsweise dem Kauf eines Autos.
- Der Arbeitsort hat erhebliche Auswirkungen auf den Nutzer – um ein Mobiltelefon oder ein Auto zu kaufen, muss der Nutzer möglicherweise reisen oder einen Paketdienst für die Lieferung des Artikels nutzen. Dennoch erfordert ein Arbeitsplatz einen Umzug, wenn er weit vom Wohnort des Nutzers entfernt ist.
- Aus technischer Sicht weist eine Stellenbeschreibung viele qualitative Aspekte auf und je nach Position können die Unterschiede selbst zwischen Anzeigen desselben Unternehmens erheblich sein.

Machine Translated by Google

Zusätzlich zu diesen Merkmalen haben wir während unserer Untersuchungen auch die folgenden Herausforderungen für Empfehlungsmethoden im Kleinanzeigenbereich festgestellt, die unsere Implementierung der Methoden beeinflusst haben:

- Die Anzahl der Benutzer und Angebote: Bei OLX Jobs hatten wir es mit Millionen von Benutzern und Zehntausenden von Angeboten zu tun, die sich beide ständig ändern.
- Ein Benutzer muss kein Profil erstellen, um mit Anzeigen zu interagieren. Dies bedeutet, dass Informationen aus einem Profil nicht für Empfehlungen verwendet werden können und Empfehlungen durch das Benutzerverhalten beim Anzeigen bestimmt werden.
- Eine Anzeige hat normalerweise eine begrenzte Anzahl von Besuchern (die mit der Anzeige interagieren), da es sich bei einer Anzeige normalerweise um ein einmaliges Angebot handelt. Nachdem der Bedarf gedeckt wurde, wird die Anzeige deaktiviert.
- Es gibt keine Informationen darüber, ob eine Transaktion stattgefunden hat; bei Kleinanzeigen bezieht sich die Konvertierung auf den Erhalt einer Antwort auf eine veröffentlichte Anzeige, da jede Transak

Diese Herausforderungen erschweren nicht nur die Implementierung, sondern auch die Prüfung der Genauigkeit der vorgeschlagenen Methoden. Wir haben diese Herausforderungen bei der Vorbereitung des Datensatzes sowie bei der Implementierung und Bewertung der in diesem Dokument beschriebenen Methoden berücksi

2.2. Empfehlungssysteme

Es hat sich gezeigt, dass das Problem der Überauswahl [3] den Konsum von Produkten verringern [4] und die Einnahmen der Site reduzieren kann [5]. Anbieter können ihre Benutzer nach zusätzlichen Informationen über ihre Erwartungen in Bezug auf Artikel fragen, um die Anzahl der angebotenen Auswahlmöglichkeiten zu reduzieren. Selbst wenn der Benutzer bereit ist, Zeit für die Bereitstellung dieser Informationen aufzuwenden, ist die Anzahl der Möglichkeiten normalerweise zu groß, um sie dem Benutzer anzuzeigen. Um dieses Problem zu lösen, haben viele Unternehmen wie Netflix [6] und Amazon [7] personalisierte Empfehlungssysteme entwickelt, deren Aufgabe es ist, Benutzern relevante Artikel v

Empfehlungen werden in vielen Bereichen verwendet, darunter Musik, Tourismus, Mode und Lebensmittel. Zahlreiche Studien wurden in den Bereichen E-Commerce [9], Kleinanzeigen [10] und Stellenempfehlungen [11–13] durchgeführt. Li et al. [14] verbesserten ihre Empfehlungen auf LinkedIn erheblich, indem sie Deep Transfer Learning nutzten, um domänenspezifische Modelle zum Verständnis von Stellen zu erstellen. Zhu et al. [15] schlugen ein domänenübergreifendes Empfehlungsmodell vor, das einen heterogenen Graphen und einen Community-Erkennungsalgorithmus verwendet, um Kurse und Stellen zu empfehlen. Lacic et al. [16] bewerteten verschiedene Autoencoder-Architekturen für sitzungsbasierte Stellenempfehlungen und präsentierten die Ergebnisse anhand beliebter Datensätze zu Stellene

Typischerweise unterscheiden wir zwei Ansätze zur Bereitstellung von Empfehlungen: inhaltsbasierte Empfehlungen und kollaboratives Filtern [8].

Bei inhaltsbasierten Empfehlungen [17] werden Informationen über Nutzer (z. B. Hobby, Ausbildung, Fähigkeiten) und Artikel (z. B. Preis, Standort, Kategorie) berücksichtigt. Diese Systeme nutzen in der Regel die Interaktionshistorie (Bewertungen, Besuche, Käufe und Antworten) eines einzelnen Nutzers, berücksichtigen aber nicht die Bewertungen aller Nutzer gleichzeitig.

Das Sammeln/Verarbeiten von Informationen über Benutzer und Artikel ist nicht immer machbar. Beim kollaborativen Filteransatz [18,19] verwenden wir nur Informationen über Benutzerinteraktionen zwischen Artikeln. Auch ohne Informationen über Benutzermerkmale und Artikelmerkmale können diese Systeme sehr genaue Empfehlungen liefern. Wir haben uns für diesen Ansatz entschieden, weil Benutzermerkmale fehlen und es schwierig ist, kollaboratives Filtern durch die Verwendung von Artikelmerkmalen in einer Produktionsumgebung zu übertreffen (wir gehen näher darauf in Abschnitt 6.3 ein).

Einer der größten Nachteile des kollaborativen Filterns ist die Unmöglichkeit, Empfehlungen zu Benutzern und Elementen ohne Interaktionen bereitzustellen, das sogenannte Kaltstartproblem [20]. Eine mögliche Lösung zur Reduzierung dieses Problems ist das häufige Neutrainieren des Modells. Bei vielen Benutzerinteraktionen erfordert diese Strategie einen Fokus auf die Skalierbarkeit des Modells [21].

Ein weiterer wichtiger Faktor, der die Wahl des Modells bestimmt, ist die Art der Interaktion zwischen einem Benutzer und einem Artikel [20]. Wir nennen das Feedback explizit, wenn der Benutzer seine Präferenz für den Artikel ausdrücklich angibt (z. B. indem er ihn auf einer Skala von 1 bis 5 bewertet). Ein anderer Typ ist implizites Feedback, wenn wir die Präferenz des Benutzers für den Artikel nur anhand von

Machine Translated by Google

auf das Nutzerverhalten (z. B. beim Kauf eines Artikels oder beim Ansehen eines Films). Implizites Feedback ist weniger genau, häufiger und drückt in der Regel nur positive Präferenzen aus [22].

Laborbewertungen von Empfehlungssystemen basieren üblicherweise auf Genauigkeitsmetriken [23]. In letzter Zeit wurde viel Aufmerksamkeit auf andere Aspekte wie Vielfalt, Unsicherheit, Neuartigkeit und Reichweite gelegt [24,25]. Bei Kleinanzeigen kann ein Artikel normalerweise nur von einem Benutzer konsumiert werden, daher sollten wir es vermeiden, denselben Artikel einer signifikanten Anzahl von Benutzern zu empfehlen. Daher ist es wichtig, sich auf die Katalogabdeckung zu konzentrieren [26] . Eine der schwierigsten Aufgaben bei der Bewertung von Empfehlungssystemen besteht darin, die Offline-Metrik zu finden, die am stärksten mit dem Online-Ziel korreliert [27].

In diesem Artikel vergleichen wir kollaborative Filtermodelle anhand eines umfangreichen impliziten Feedback-Datensatzes. Wir vergleichen die Genauigkeit, Vielfalt und Effizienz ausgewählter Empfehlungsmethoden in einer Laborumgebung und beweisen anschließend die Qualität der Methoden in Onlinetests mit Benutzern.

2.3. Empfehlungsmethoden

Nachfolgend führen wir die wichtigsten Aspekte auf, die die Methodenauswahl für unsere Forschung beeinflusst haben. • Popularität in der Literatur. Wir haben Methoden ausgewählt, deren Leistung gründlich evaluiert wurde. Alle ausgewählten Methoden außer einer wurden mindestens hundertmal zitiert .

• Beliebtheit in der Branche. Die Bewertung dieses Aspekts war schwierig, da Unternehmen die Ergebnisse durchgeführter Experimente selten veröffentlichen. Einige Beliebtheitsindikatoren sind die Verfügbarkeit der Implementierung und die Anzahl der Benutzer (z. B. widergespiegelt in der Anzahl der Sterne und Forks auf GitHub). Wir haben auch die Erfahrung der in unserem Unternehmen tätigen Datenwissenschaftler berücksichtigt , die die meisten der ausgewählten Methoden **in der Vergangenheit für einige andere Anwendungsfälle eingesetzt haben .**

- **Skalierbarkeit. Wir haben nur Lösungen in Betracht gezogen, die Millionen von Benutzern und Elementen verarbeiten können . Darüber hinaus haben wir das Kaltstartproblem durch häufiges Neutrainieren des Modells reduziert. Daher mussten wir Techniken auswählen, die innerhalb weniger Stunden an unserem Datensatz trainiert werden konnten.**

• Vielfalt der Ansätze. Wir wollten Methoden verschiedener „Familien“ testen, um ihre Anwendbarkeit auf Kleinanzeigen zu beurteilen. Wir wollten die vielversprechendste Familie von Ansätzen identifizieren, um die Forschung in dieser Richtung fortzusetzen.

Schließlich wurden ALS, LightFM, Prod2Vec, RP3Beta und SLIM zum Testen und Evaluieren ausgewählt . In den folgenden Unterabschnitten werden die Gründe für die Auswahl der einzelnen Methoden und die Methodendefinition beschrieben. Eine gemeinsame Einschränkung dieser Methoden ist ihre Unfähigkeit, Elemente zu empfehlen, die von keinem Benutzer besucht wurden, und Empfehlungen für Benutzer zu generieren, die kein Element besucht haben (Kaltstartproblem, [21]). Mit unserem Datensatz ohne Benutzer- und Elementmerkmale ist es unmöglich, dieses Problem durch eine personalisierte Methode zu überwinden.

Ausgewählte Methoden wurden bereits in verschiedenen Bereichen verglichen. Anelli et al. [28] verglichen RP3Beta, SLIM, ALS und BPRMF (das bei Verwendung **der BPR-Verlustfunktion LightFM entspricht)** anhand von drei Datensätzen: Amazon Digital Music, Epinions und MovieLens **1M hinsichtlich Genauigkeit, Vielfalt und Effizienz. Die Anzahl der Interaktionen in diesen Datensätzen beträgt 145.523, 300.475 bzw. 571.531. Unsere Studie verwendete einen Datensatz mit mehreren zehn Millionen Interaktionen, ähnlich denen in OLX-Produktionssystemen.**

Zusätzlich stellen wir Ihnen die Ergebnisse einer Online-Auswertung zur Verfügung.

Dacrema et al. [29] verglichen die Genauigkeit von SLIM, RP3Beta und ALS anhand mehrerer Datensätze (CiteULike-a, MovieLens 1M, Pinterest, Amazon Music, Amazon Movies, MovieLens 20M, Netflix, Yelp und Gowalla). Sie zeigten, dass jeder dieser Algorithmen die anderen beiden bei mindestens einem Datensatz übertraf. Einige ihrer Ergebnisse wurden von Anelli et al. **[30] anhand der Datensätze MovieLens 1M und Pinterest** weiter bestätigt. **Zhu et al. [31] verglichen BPRMF, SLIM und Prod2Vec anhand dreier Datensätze (Amazon Books, Yelp und Gowalla) und zeigten die Überlegenheit von SLIM. Dong et al. [32] analysierten die Genauigkeit von RP3Beta, SLIM und ALS anhand von neun Datensätzen, wobei SLIM bei acht Datensätzen die besten Ergebnisse lieferte und ALS**

Machine Translated by Google

Im Gegensatz zu unserer Studie verglichen diese Arbeiten die ausgewählten Methoden weder in Bezug auf Diversität oder Effizienz und führte auch keine Online-Evaluation durch.

In letzter Zeit wurden mehrere Empfehlungsmethoden auf Basis neuronaler Netze entwickelt. vorgeschlagen. Es hat sich gezeigt, dass ausreichend bestimmte und abgestimmte einfache Modelle oft übertreffen sie [28,31–33]. Aufgrund der Komplexität der Berechnungen sind Modelle, die auf auf neuronalen Netzen werden seltener an Datensätzen getestet, die mehrere Millionen Interaktionen [28]. Daher überlassen wir es zukünftigen Forschungen, fortgeschrittenere und skalierbare Methoden, die für unseren Datensatz bessere Ergebnisse liefern.

2.3.1. Matrixfaktorisierungsmodelle

Matrixfaktorisierungstechniken werden seit dem Netflix-Preiswettbewerb [34] erfolgreich in Empfehlungssystemen eingesetzt. Die Hauptidee besteht darin, die sp Benutzer-Element-Interaktionsmatrix als Produkt zweier kleinerer, dichter Matrizen. Wir evaluieren zwei Matrixfaktorisierungsansätze: LightFM und ALS.

Das LightFM -Modell [35] wurde vorgeschlagen, um das Kaltstartproblem von Matrixfaktorisierungsmethoden zu überwinden . In unserem Fall reduziert es sich auf die klassische Matrixfaktor Informationen über Benutzer und Artikel. Trotzdem haben wir uns für diesen Ansatz entschieden, weil seine effiziente offizielle Umsetzung bei der Unterstützung mehrerer Verlustfunktionen: Logistik, BPR [36], WARP [37] und -OS WARP [38].

Sei U eine Menge von Benutzern und I eine Menge von Artikeln. Für einen gegebenen Benutzer ÿ U und Artikel ÿ I define = 1, wenn der Benutzer mit dem Element , und = 0, andernfalls. Das LightFM-Modell interagiert, wird die Punktzahl wie folgt vorhergesagt:

$$\hat{} = \cdot + + \quad ,$$

wobei und die -dimensionalen latenten Repräsentationen des Benutzers und des Elements bezeichnen und und die Vorurteile des Benutzers und des Elements sind. Alle diese Parameter werden wä Das Training.

Die logistische Verlustfunktion sollte verwendet werden, wenn sowohl positives als auch negatives Feedback Typen sind verfügbar, daher ist es für unseren Datensatz ungeeignet, was wir während der Hyperparameter-Optimierung experimentell bestätigt haben.

Die Verlustfunktionen BPR, WARP und -OS WARP sind paarweise Lern-zu-Rang-Ansätze , die für unsere Top-Empfehlungen-Aufgabe relevanter sind. Sie versuchen zu modifizieren Modellparameter zur Maximierung der Differenz zwischen den Punktzahlen der Items ÿ D mit mit denen der Benutzer interagiert und mit den Elementen ÿ D, mit denen der Benutzer nicht interagiert.

Konkret versuchen wir in der BPR-Verlustfunktion [36] den Ausdruck zu maximieren, wie folgt:

$$\sum_{(,,)\ \ddot{y}S} \ln(\hat{} \quad \hat{}) \ \ddot{y} \ \ddot{y} \ ||\ddot{y}||2 \ ,$$

Wo

$$S = \{(,,) | \ \ddot{y} \ U, \ \ddot{y} \ D \text{ und } \ddot{y} \ D\},$$

$$() = \frac{1}{1 + \ddot{y}},$$

ÿ ist ein modellspezifischer Regularisierungsparameter, ÿ ist ein Parametervektor und || · || ist der Euklidische Norm.

Wir präsentieren WARP [37] als einen Fall von -OS WARP [38]. Die -OS WARP-Verlustfunktion ist eine Summe der Verluste für jeden Benutzer:

$$\sum_{\ddot{y}U} (\hat{} \ , D),$$

wobei ^ ein Vektor der geschätzten Punktzahlen^ für jedes Element ÿ I ist.

Machine Translated by Google

Für einen gegebenen , Wir ordnen die interagierenden Elemente D in absteigender Reihenfolge der Benutzer werden die Punkte geschätzt. Wir ordnen die Elemente nämlich als 1, 2, . . . , | D | , wobei die Indizes sind so gewählt, dass wir folgendes haben:

$$\hat{\ddot{y}}_1 \ddot{y} \hat{\ddot{y}}_2 \ddot{y} \ldots \ddot{y} \hat{\ddot{y}}_{|D|}.$$

Dann haben wir Folgendes:

$$(\hat{\ddot{y}}, D) = \frac{1}{\ddot{y}\ddot{y}} \sum_{=1}^{|\quad|} (\quad)\ddot{y}(\text{Rang}(\hat{\ddot{y}})),$$

wobei Rang (ˆ) = ÿl\D (1 + ( ) ˆ ÿ ˆ), ist eine Indikatorfunktion, ÿ( ) = $\sum_{=1} \frac{1}{}$, und = $\sum_{=1}^{|D|}$ normalisiert die durch induzierten Gewichte.

In der Praxis wählen wir ein positives Item bezüglich der Gewichtungsfunktion aus und Wählen Sie nach dem Zufallsprinzip unbekannte Elemente aus, bis nach den Proben 1 + ÿ . Dann führen wir eine Gradientenschritt zur Minimierung von ÿ(ÿ ÿ) (1 ÿ + ˆ).

Schließlich wird der -OS WARP-Verlust für ( ) = 1 und ( ) = 0 erreicht, wenn ÿ . Der WARP Verlust wird erreicht, wenn wir ( ) = 1 für ÿ N setzen. Wir können die Bestellung positiver Elemente aus D um die Berechnungen zu vereinfachen.

In unserer Forschung wurden alle Verlustfunktionen (Logistik, BPR, WARP und -OS WARP) behandelt als Hyperparameter des LightFM-Modells und während des Trainings optimiert.

Das ALS-Modell (alternating least squares), auch bekannt als WRMF – weighted regularized matrix factorization [39], ist ein Beispiel für eine Matrixfaktorisierungsmethode, die für implizite Feedback-Datensätze. Wir haben diese Methode aufgrund ihrer nachgewiesenen Leistung gewählt, Skalierbarkeit und Popularität in Forschung und Industrie. Darüber hinaus hat das ALS-Modell wurde bereits in OLX implementiert und verwendet.

Wir definieren die Konfidenz einer Beobachtung als = 1 + , wobei ein Hyperparameter ist . Wir definieren die Präferenz auch als = (> 0). Dann optimieren wir in ALS die folgender Ausdruck:

$$\underset{\ddot{y},\ddot{y}}{\text{Mindest}} \ddot{y}\ddot{y} \quad (\ddot{y} \quad \ddot{y})^2 + (\ddot{y}\ddot{y} \,||\,||2 + \ddot{y}\ddot{y} \,||\,||2),$$

wobei die -dimensionale Einbettung des Benutzers ist, die -dimensionale Einbettung des Elements ist, und ist ein Regularisierungsparameter. Der Name des Modells bezieht sich auf die optimierte Methode für diese Funktion. Das Verfahren der alternierenden kleinsten Quadrate geht davon aus, dass Benutzer oder Artikel latent Faktoren sind festgelegt und bestimmen die anderen analytisch. Das Verfahren wird effizient wiederholt einige Male [39,40].

Bei Matrixfaktorisierungsansätzen wird die Einbettung für jeden Benutzer separat gelernt, Dies kann bei einer großen Benutzeranzahl zu einem hohen Zeit- und Speicheraufwand führen. Darüber hinaus könnten die Empfehlungen nicht erklärbar sein, da zwei Benutzer mit dem Die gleichen Interaktionen können unterschiedliche Empfehlungen erhalten. Andere betrachtete Modelle (SLIM, RP3Beta, Prod2Vec) bieten die gleichen Empfehlungen für Benutzer mit dem dieselben Interaktionen.

2.3.2. Nachbarschaftsbasierte Modelle

Das SLIM-Modell (sparse linear methods for Top-N recommender systems [41]) ist basierend auf der itembasierten Nearest Neighbors Regression [8]. Wir haben diese Methode gewählt, weil sie Vorteile gegenüber itembasierten Nearest-Neighbors- und Matrixfaktorisierungsmodellen [41] und gute Leistung gegenüber nicht-neuralen und neuronalen Ansätzen [29].

In diesem Modell approximieren wir die Benutzer-Artikel-Matrix durch das Lernen der Artikel-Artikel-spärliche Ähnlichkeitsmatrix . Explizit wird durch Minimieren des folgenden Ausdrucks gelernt:

$$\frac{1}{} || \ddot{y}\,2 \quad ||2 + \frac{}{||\,2} \quad ||2 + ||\,||1,$$

Machine Translated by Google

wobei ÿ 0 und = 0 für alle , , 1 ÿ , ÿ ||, || · || ist die Matrix-Frobenius-Norm:

$$\| \| = \ddot{y}\ddot{y} \ddot{y}\ddot{y}^{\|ch\|} \; {}^{\|ch\|} \; {}^{2}, \quad {}_{=1} \; {}_{=1}$$

|| · ||1 ist die eintragsweise ÿ1-Norm:

$$\| \| 1 = \ddot{y}\ddot{y} \ddot{y}\ddot{y}^{\|ch\|} \; {}^{\|ch\|} \; | \; |, \quad {}_{=1} \; {}_{=1}$$

und und sind Regularisierungsparameter.

Wir können sehen, dass jede Spalte unabhängig voneinander gelernt werden kann, was die Skalierbarkeit dieses Modells [41].

Eine Einschränkung dieser Methode besteht darin, dass sie keine Benutzer- und Artikeleinbettungen erzeugt, die für andere Aufgaben hilfreich sein könnten (z. B. Clustering von Benutzern oder Artikeln). Wir können die Spalten der Matrix als Artikeleinbettungen behandeln, aber es handelt sich um spärliche ||-dimensionale Vektoren, sodass sie wahrscheinlich nicht direkt für andere Aufgaben verwendet werden können.

2.3.3. Graphenbasierte Ansätze

RP3Beta [42] ist ein graphenbasierter Ansatz für kollaboratives Filtern. Ähnlich wie Dacrema et al. [29] präsentieren wir dieses Modell als Erweiterung des P3alpha-Modells [43]. Wir betrachten einen bipartiten und ungerichteten Graphen = (V, E), wobei jeder Knoten ÿ V entweder einen Benutzer oder ein Element darstellt und jede Kante { , } ÿ E die Interaktion zwischen einem gegebenen Benutzer und einem gegebenen Element darstellt. Lassen Sie die Adjazenzmatrix und d

Anschließend werden die Ergebnisse von P3alpha in der Matrix gespeichert.

$$((\quad {}^{\ddot{y}1} \quad)^{\ddot{y}} \quad)^{3},$$

wobei ÿ die elementweise Potenz einer Matrix bezeichnet (Hadamard-Potenz). Genauer gesagt wird **der Score ˆ am Schnittpunkt der Zeile, die den Benutzer darstellt, und der Spalte,** die das **Element darstellt, gespeichert. Wir können sehen, dass dieser Score eine Summe der Scores ist, die den Pfaden der Länge 3 zugewiesen sind, die Benutzer und Element im Graphen verbinden.**

In RP3Beta wird die Punktzahl mit P3alpha berechnet und durch die Popularität der Elemente hoch 1 geteilt, um den Popularitätsfehler zu mildern.

Wir haben uns aufgrund seiner Einfachheit, Skalierbarkeit und nachgewiesenen Leistung für dieses Modell entschieden **[29,30], insbesondere angesichts der hohen Spärlichkeit unseres Datensatzes, die den Speicherbedarf beim Trainieren des Modells erheblich verringert. Dies ermöglicht es uns, die Artikel-Artikel-Ähnlichkeitsmatrix direkt zu berechnen und im Speicher zu speichern, anstatt die im Originalpapier vorgeschlagene Random-Walk-Approximation zu verwenden.**

RP3Beta generiert, ähnlich wie SLIM, keine Benutzer- und Artikeleinbettungen. Eine weitere damit verbundene Einschränkung ist die Unfähigkeit, Artikel zu empfehlen, deren Distanz auf dem bipartiten Benutzer-Artikel- Diagramm größer als drei von einem bestimmten Benutzer ist. Insbesondere kann dem Benutzer ein Artikel nur empfohlen werden, wenn mindestens ein Benutzer ÿ den Artikel besucht hat und mindestens einer der Artikel vom Benutzer besucht wurde. Darüber hinaus verfügt das RP3Beta-Modell nicht über Parameter, die während des Trainings optimiert werden; daher ist es möglicherweise nicht in der Lage, die komplexen Beziehungen i

2.3.4. Prod2Vec

Das Prod2Vec-Modell [44] (oder Item2vec [45]) basiert auf dem Word2Vec-Modell [46], das in der Verarbeitung natürlicher Sprache weit verbreitet ist. Das Modell konzentriert sich auf das Lernen von Vektordarstellungen von Elementen unter Verwendung einer Sequenz von Elementen als „Satz“ und Elementen innerhalb dieser Sequenz als „Wörter“. Diese Methode wurde gewählt, weil sie skalierbar ist und sich stark von den verglichenen Methoden unterscheidet. Darüber hinaus basiert dieser Ansatz auf dem aktuellen hochmodernen hybriden Deep Neural Network-Empfehlungssystem, das in anderen Kategori

Machine Translated by Google

Prod2Vec verwendet das Skip-Gram-Modell durch Maximierung der Zielfunktion über den Satz S von Itemsequenzen, wie folgt definiert:

$$= \ddot{y}\ddot{y} \quad \ddot{y}\ddot{y} \quad \ddot{y}\ddot{y} \quad \log P(+ |),$$

wobei die Kontextlänge für Elementsequenzen ist und Elemente aus derselben Sequenz beliebig angeordnet. Die Wahrscheinlichkeit P(+ |) , ein benachbartes Element + zu beobachten , **wenn Das aktuelle Element wird mithilfe der Soft-Max-Funktion definiert:**

$$P(+ |) = \frac{\exp(\ddot{y} \; \ddot{y} \;)}{\sum_{=1}^{|ch|} \exp(\ddot{y} \; \ddot{y} \;)},$$

wobei und die Eingabe- und Ausgabevektordarstellungen von item sind und || der , Anzahl eindeutiger Elemente im Vokabular. Prod2Vec modelliert den Kontext einer Elementsequenz, wo Elemente mit ähnlichem Kontext (d. h. mit ähnlichen benachbarten Interaktionen) ähnliche Vektordarstellungen.

Bevor wir Empfehlungen generieren, normalisieren wir alle Eingabevektordarstellungen. was die Effizienz verbessert. Um Empfehlungen zu generieren, stellen wir den Benutzer als Durchschnitt der Eingabevektordarstellungen der interagierenden Elemente. Dann suchen wir nach Elementen deren Eingabedarstellungen die höchste Kosinusähnlichkeit mit der Benutzerdarstellung aufweisen (-nächste Nachbarn im latenten Raum).

Eine mögliche Einschränkung dieser Methode ist eine strikte Abhängigkeit zwischen Benutzer und Artikel Darstellungen. Einige Interaktionen sollten möglicherweise einen größeren Einfluss auf Benutzerdarstellungen haben als andere (was in Matrixfaktorisierungsmodellen möglich ist).

2.4. Forschungslücke

Ziel dieses Artikels ist es, verschiedene Empfehlungsmethoden im Bereich Stellenanzeigen zu vergleichen und eine spezifische Forschungslücke zu schließen. Nach bestem Wiss Die angesprochenen Empfehlungsmethoden wurden noch nie in großem Maßstab getestet oder Maßstab bei der Online-Arbeit im Jobbereich. In dieser Arbeit vergleichen wir diese Methoden in einem Laborumgebung unter Verwendung des von uns veröffentlichten Datensatzes und in Online-A/B-Tests, die in einem Produktionsumfeld.

3. OLX-Datensatz zu Jobinteraktionen

Diese Forschung liefert einen kommerziellen Datensatz, der 65.502.201 durchgeführte Veranstaltungen enthält auf OLX Jobs von 3.295.942 Benutzern, die im Jahr 2020 in zwei Wochen mit 185.395 Stellenanzeigen interagierten. Wir haben den Datensatz auf Kaggle veröffentlicht (https://www.kaggle.com/olxdatascience/olx-jobs-interactions, abgerufen am 26. Juni 2024).

Jede Zeile des Datensatzes stellt das Ereignis dar, das ein bestimmter Benutzer in Bezug auf ein bestimmtes Anzeige zu einem bestimmten Zeitpunkt (Tabelle 1). Die folgenden Ereignisse im Zusammenhang mit dem Nutzerverhalten wurden berücksichtigt:

- **Klick: Der Benutzer hat die Artikeldetailseite besucht;**
- **Lesezeichen: Der Benutzer hat das Element zu den Lesezeichen hinzugefügt;**
- **chat_click: Der Benutzer hat den Chat geöffnet, um den Besitzer des Artikels zu kontaktieren;**
- **contact_phone_click_1: Der Benutzer hat die dem Artikel beigefügte Telefonnummer angezeigt;**
- **contact_phone_click_2: Der Benutzer hat geklickt, um den Besitzer des Artikels anzurufen;**
- **contact_phone_click_3: Der Benutzer hat geklickt, um dem Besitzer des Artikels eine SMS-Nachricht zu senden;**
- **contact_partner_click: Der Benutzer hat geklickt, um auf die externe Seite des Artikelbesitzers zuzugreifen;**
- **contact_chat: Der Benutzer hat dem Besitzer des Artikels eine Nachricht gesendet.**

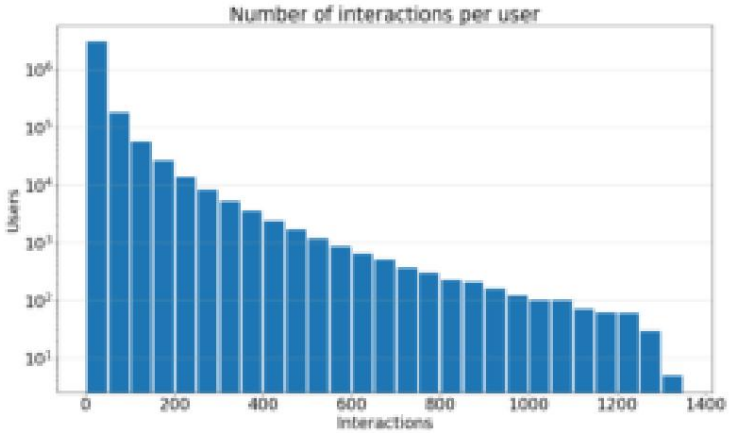
Machine Translated by Google

Tabelle 1. Beispielzeilen aus dem Datensatz.

Benutzer	Artikel	Ereignis	Zeitstempel
1745587	168661		1582216025
843008	62838		1582485868
14285	30469	klicken	1582247367
1142944	80122		1581805847
2835659	23728	klicken Lesezeichen klicken chat_click	1582247836

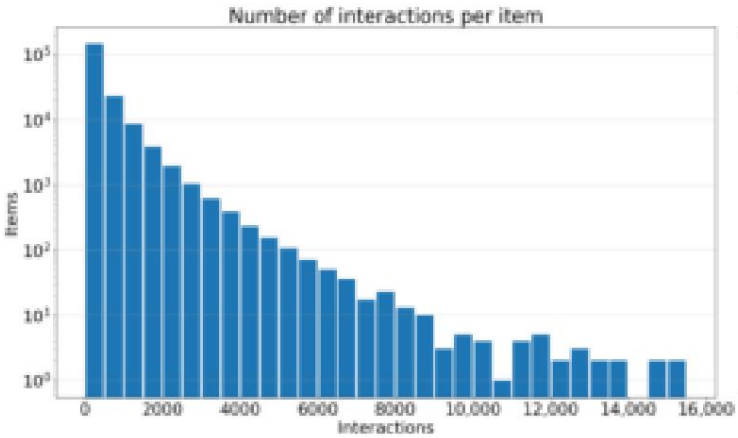
Quelle: eigene Studie.

Die Datensparsity (% fehlender Einträge in der Benutzer-Element-Matrix) beträgt 99,9923 %, Das ist zwar sehr niedrig, aber typisch für Kleinanzeigen (zum Vergleich: Die Datenknappheit auf der Der MovieLens 20M-Datensatz ist zu 99,47 % dicht und damit fast 70-mal dichter. Verteilungen von Interaktionen pro Benutzer, Interaktionen pro Artikel und Ereignisse sind in den Abbildungen 1 und 2 aufgeführt. und Tabelle 2. Wir haben eine logarithmische Skala verwendet, da die Verteilungen stark rechtsschief.



Quantile	Value
0.01	1
0.10	1
0.25	2
0.50	6
0.75	18
0.90	46
0.99	221

Abbildung 1. Anzahl der Interaktionen pro Benutzer. Quelle: eigene Studie.



Quantile	Value
0.01	1
0.10	1
0.25	3
0.50	137
0.75	428
0.90	950
0.99	2944

Abbildung 2. Anzahl der Interaktionen pro Artikel. Quelle: eigene Studie.

Laut De Ruijt und Bhulai [47] sind **die beliebtesten Datensätze für die Bewertung von Job-Empfehlungssystemen CareerBuilder12 (https://www.kaggle.com/ Wettbewerbe/Job-Empfehlungen, abgerufen am 16. Juni 2024), RecSys16 [12] und Rec-Sys17 [13]. Tabelle 3 zeigt die Statistiken der betrachteten Datensätze. Die Anzahl der eindeutigen Interaktionen ist die Anzahl der eindeutigen Benutzer-Element-Paare, bei denen der Benutzer mit dem Obwohl die Forscher die Datensätze RecSys16 und RecSys17 verwenden, waren sie im Rahmen von Data Science-Wettbewerben unter streng begrenzten Bedingungen veröffentlicht und sind keine nicht mehr verfügbar. Daher konnten wir die Anzahl der einzigartigen Interaktionen für diese**

Machine Translated by Google

Datensätze. Wir haben festgestellt, dass der OLX-Datensatz für Jobinteraktionen die höchste Anzahl von Interaktionen und Typen zwischen den betrachteten Datensätzen.

Tabelle 2. Verteilung der Ereignisse.

Aktion	Häufigkeit (%)
	89.794
Klicken Sie auf contact_phone_click_1 Lesezeichen	2.628
chat_click	2.511
contact_chat	2.136
contact_partner_click	1.448
contact_phone_click_2	0,701
contact_phone_click_3 Quelle:	0,679
eigene Untersuchung.	0,103

Tabelle 3. Vergleich zwischen Datensätzen mit Stellenempfehlungen. Der Buchstabe gibt die Standardabweichung an.

	Jobs bei OLX Interaktionen	CareerBuilder12	RecSys16	RecSys17
Interaktionen	65,50 M	1,60 M	8,83 M	8,27 M
Einzigartige Interaktionen	47,17 M	1,60 M	-	-
Benutzer	3,30 M	0,3 2 Mio.	1,37 Mio.	1,50 Mio.
Durchschnittliche Anzahl einzigartiger Interaktionen pro Benutzer	14.31 (= 29,23)	4,99 (= 11,42)	<6,46	<5,51
Artikel	0,19 M	0,37 Mio.	1,36 Mio.	1,31 Mio.
Durchschnittliche Anzahl einzigartiger Interaktionen pro Artikel	254,42 (= 426,02)	4.38 (= 8,19)	<6,50	<6,31
Dichte der Interaktionsmatrix (%)	0,0077 %	0,0014 %	< 0,0005 %	< 0,0004 %
Arten von Interaktionen	8	1	4	5
Zeitstempel	ÿ	ÿ	ÿ	ÿ
Benutzerfunktionen	ÿ	ÿ	ÿ	ÿ
Artikelmerkmale	ÿ	ÿ	ÿ	ÿ

Quelle: eigene Studie.

Es ist zu beachten, dass die Wahrung der Vertraulichkeit von Anzeigen und Benutzern Priorität hatte bei der Erstellung dieses Datensatzes. Zu den Maßnahmen zum Schutz der Privatsphäre gehörten die folgenden:

- **Ursprüngliche Benutzer- und Artikelkennungen wurden durch eindeutige, zufällige Ganzzahlen ersetzt;**
- **Jedem Zeitstempel wurde eine nicht offengelegte konstante Ganzzahl hinzugefügt;**
- **Einige Teile der Interaktionen wurden herausgefiltert;**
- **Einige zusätzliche künstliche Interaktionen wurden hinzugefügt.**

4. Versuchsaufbau

4.1. Umsetzung

Die oben genannten Methoden haben wir auf Basis früherer Entwicklungen implementiert. Details dazu, wie sich unsere Implementierungen von den zuvor getesteten Ansätzen unterscheiden, finden Sie Die Umsetzung der Methoden zur Reproduzierbarkeit der Ergebnisse ist verfügbar auf GitHub (https://github.com/rob-kwiec/olx-jobs-recommendations, aufgerufen am 29. Mai 2024).

Machine Translated by Google

Tabelle 4. Für die Forschung neu implementierte Methoden.

Methode	Methodenfamilie	Quelle	Unterschied zum Ursprüngliche Methode
LightFM	Matrixfaktorisierung	Implementierung basierend auf: [48]. Begleitpapier: [35].	Keiner
ALS	Matrixfaktorisierung	Implementierung basierend auf: [49]. Unterstützende Papiere: [39,40].	Keiner
SCHLANK	Nachbarschaftsbezogen	Quellen, die unsere Implementierung inspiriert haben: [50,51]. Begleitpapier: [41].	Keiner
RP3Beta	Graphenbasiert	Quelle, die unsere Implementierung inspiriert hat: [52]. Unterstützende Papiere: [29,53].	Direkt durchgeführt Berechnungen auf dünn besetzten Matrizen statt auf zufälligen Spaziergänge Näherung.
Prod2Vec	Word2Vec	Die Implementierung basiert auf: [54]. Unterstützende Papiere: [44,45].	Sequenzen von Interaktionen, sortiert nach einem Zeitstempel. Darstellung des Benutzers als Durchschnitt der Darstellungen interagierter Elemente. Wir haben auch mit CBOW (Continuous Bag of Words) experimentiert.

Quelle: eigene Studie.

4.2. Ziele der Laborforschung und des Online-Tests **Unser oberstes**

Geschäftsziel war es, die Zahl der Benutzer zu erhöhen, die sich über OLX Jobs auf Stellen bewerben. Um dies zu erreichen, versuchen wir, ein Ranking-Problem zu lösen. Genauer gesagt empfehlen wir zehn Elemente mit der höchsten Wahrscheinlichkeit, eine Interaktion eines bestimmten Benutzers zu erhalten. Das Problem, eine Offline-Metrik zu finden, die am stärksten mit der Bewerbungsrate im Jobbereich korreliert, wurde von Mogenet et al. [27] untersucht. Da unser Ziel ähnlich ist, folgten wir ihren Ergebnissen, indem wir auf Präzision@10 optimierten (Präzision für die ersten zehn empfohlenen Elemente [23]). Wir berichten und diskutieren jedoch auch über andere Genauigkeits-, Diversitäts- und Effizienzmetriken.

Wir konzentrieren uns auf Benutzer, die aktiv nach Jobs suchen, indem sie mit Stellenanzeigen interagieren. Daher berücksichtigen wir keine Benutzer ohne Interaktionen (Kaltstart-Benutzer).

4.3. Datenaufteilung

Der in Abschnitt **3 beschriebene Datensatz** wurde **gemäß der Forschungsmethodik in Trainings-Validierungs- und Testdatensätze aufgeteilt .** Der Testsatz enthält 20 % der neuesten Interaktio **Dies bedeutet, dass von 14 Tagen etwa 3 Tage in den Testdatensatz einbezogen wurden (siehe Abbildung** 3). Diese Art der Aufteilung des Datensatzes in Trainings- und Testdatensätze ist **laut Meng et al. [55] die realistischste** Vorgehensweise, **die verschiedene Aufteilungsstrategien zur Bewertung von Empfehlungssystemen untersucht hat.**

In beiden Datensätzen haben wir die Anzahl der Interaktionen auf einmalige Interaktionen zwischen einem Benutzer und einem Element beschränkt. Wenn ein Benutzer mehr als einmal mit einem Element interagierte, wurde nur die erste Interaktion gezählt und der Zeitstempel dieser Interaktion dem Element zugeordnet. Wir haben auch nicht zwischen den Interaktionsarten unterschieden. Darüber hinaus wurden zwei weitere Änderungen vorgenommen: Nur Benutzer, die im Trainings- und Validierungsdatensatz erschienen, wurden in den Testsatz aufgenommen (weil wir anderen Benutzern ohne zusätzliches Wissen über sie keine personalisierten Empfehlungen geben konnten), und wir haben alle Benutzer-Element-Paare herausgefiltert, die im Trainingssatz vorhanden waren (um zu vermeiden, Elemente zu empfehlen, die bereits gesehen wurden). Als Ergebnis hatten wir ungefähr 38 Millionen Zeilen im Trainings- und Validierungsdatensatz und ungefähr 6 Millionen Zeilen im Testsatz. Weitere Statistiken zu diesen beiden Datensätzen finden Sie in Tabelle 5.

Machine Translated by Google

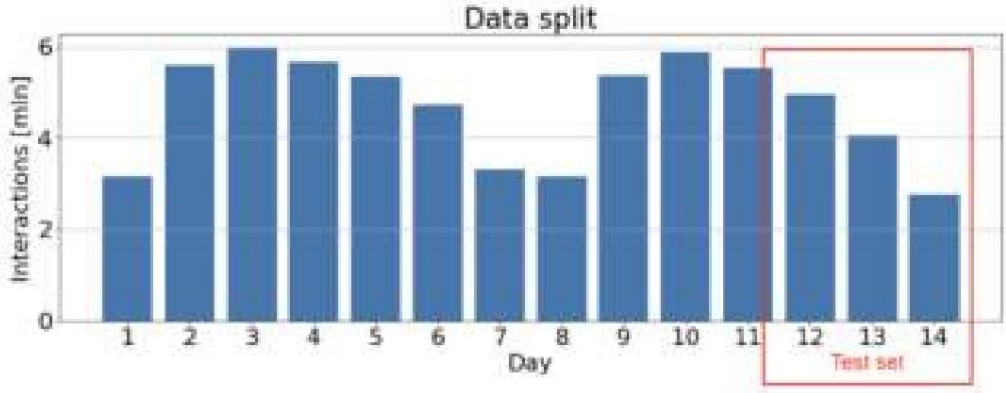


Abbildung 3. Vorbereitung des Datensatzes (aufgeteilt in Training–Validierung und Testen, abhängig von der Anzahl der Tage). Quelle: eigene Studie.

Tabelle 5. Vergleich zwischen Trainings-/Validierungs- und Testsätzen. Der Buchstabe bezeichnet Standard Abweichung.

	OLX-Jobinteraktionen Interaktionen	Trainings- und Validierungsset	Test-Set
Interaktionen	65,50 M	52,40 Mio.	6,11 Mio.
Einzigartige Interaktionen	47,17 M	38,25 Mio.	6,11 Mio.
Benutzer	3,30 M	2,83 Mio.	0,62 Mio.
Durchschnittliche Anzahl einzigartiger Interaktionen pro Benutzer	14.31 (= 29,23)	13,50 (= 26,11)	9,87 (= 14,38)
Artikel	0,19 M	0,18 Mio.	0,13 Mio.
Durchschnittliche Anzahl einzigartiger Interaktionen pro Artikel	254,42 (= 426,02)	214,05 (= 365,93)	47,51 (= 83,84)
Dichte der Interaktionsmatrix (%)	0,0077 %	0,0076 %	0,0077 %

Quelle: eigene Studie.

4.4. Modelloptimierung

Aus Kosten- und Zeitgründen sollten Sie bei der Optimierung der Parameter
Wir haben eine Beschränkung auf 20 % der Benutzer und 20 % der Elemente der Schulung und Validierung vorgenommen
Dann teilen wir den eingeschränkten Datensatz in den Trainings-Validierungs-Datensatz mit dem Regeln, die im obigen Abschnitt beschrieben wurden. Für jedes Modell wurden 100 Iterationen ausgewählt durch Bayessche Optimierung unter Verwendung von Gaußschen Prozessen, implementiert in scikit-learn (https://scikit-optimize.github.io/stable/modules/generated/skopt.gp_minimize.html, aufgerufen **am 29. Mai 2024). Die Konfigurationsdatei definiert Bereiche für Hyperparameter und ist verfügbar** in unserem Code-Repository (https://github.com/rob-kwiec/olx-jobs-recommendations/blob/main/src/tuning/config.py, abgerufen am 29. Mai 2024). Für RP3Beta haben wir beispielsweise beschränkten uns auf die Suche nach und im Intervall [0, 2].

Empfehlungen im Optimierungsprozess wurden für 30.000 Benutzer erstellt und
Wie oben erwähnt, wurde als Optimierungswert precision@10 verwendet. Abbildung 4 zeigt der Präzisionswert in Abhängigkeit von den Parametern. Es ist ersichtlich, dass ALS nicht sehr optimiert, aber für andere Methoden war diese Optimierung entscheidend für Verbesserung der Präzision. Hyperparameter für alle Modelle sind in Tabelle 6 dargestellt.

Machine Translated by Google

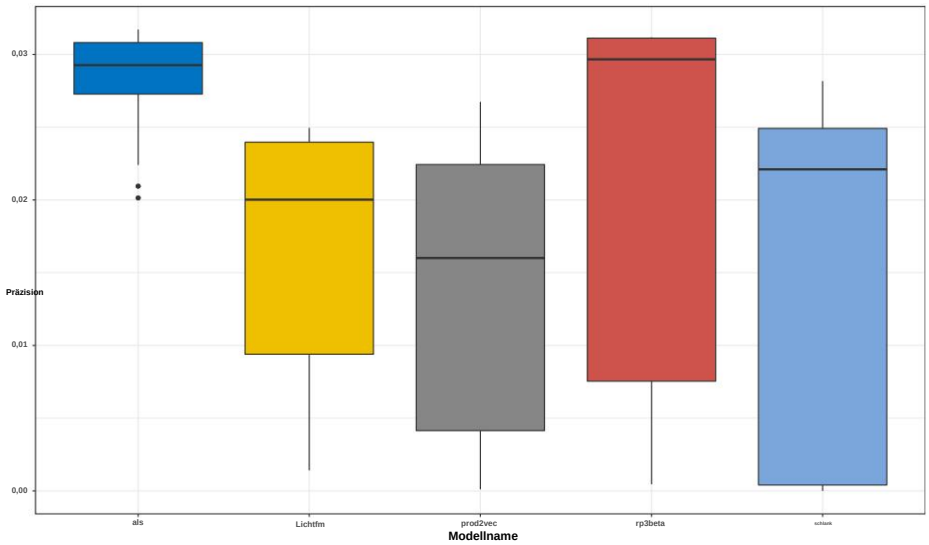


Abbildung 4. Präzision für jedes Modell, abhängig von den Parametern. Quelle: eigene Studie.

Tabelle 6. Modell-Hyperparameter.

Modell	Modellieren von Hyperparametern
als	{'Faktoren': 357, 'Regularisierung': 0,001, 'Iterationen': 20, 'Ereignisgewichtsmultiplikator': 63}
Lichtfm	{'no_components': 512, 'k': 3, 'n': 20, 'learning_schedule': 'adadelta', 'Verlust': 'Warp', 'max_sampled': 61, 'Epochen': 11}
prod2vec	{'Vektorgröße': 168, 'Alpha': 0,028728, 'Fenster': 20, 'Mindestanzahl': 16, „Beispiel“: 0,002690026, „min_alpha“: 0,0, „sg“: 1, „hs“: 1, „negativ “: 200, „ns_exponent“: ÿ0,16447846705441527, „cbow_mean“: 0, 'Epochen': 22}
RP3Beta	{'alpha': 0,61447198, 'beta': 0,1443548}
schlank	{'Alpha': 0,00181289, 'l1_ratio': 0,0, 'Iterationen': 3}

Quelle: eigene Studie.

4.5. Modellbewertung

Wir bewerten unsere Modelle hinsichtlich Genauigkeit, Vielfalt und Effizienz.

Die Genauigkeit der Empfehlungsmethode bei impliziten Feedback-Datensätzen ist typischerweise ausgewertet, indem untersucht wird, welche Elemente aus der Liste der Top-Empfehlungen der Benutzer im Testset interagiert [8]. Viele Bewertungsmetriken bestimmen die Metriken für jedes Benutzer separat und geben den Mittelwert als endgültige Bewertungsnote der Methode an [23]. Die Empfehlungssystem kann als Klassifikator betrachtet werden, bei dem die Interaktionen des Benutzers in Die Testreihe enthält positive Beispiele und die Top-Empfehlungen sind Elemente, die als positiv durch das Modell. Daher können die zur Bewertung von Klassifikatoren verwendeten Methoden direkt auf Empfehlungssysteme angewendet. Beispielsweise ist der Wert von precision@ für einen bestimmten Benutzer die Anzahl der empfohlenen Elemente, mit denen der Benutzer im Testset interagiert hat, geteilt durch die Anzahl der diesem Benutzer empfohlenen Artikel. Weitere in dieser Arbeit diskutierte Metriken waren beschrieben von Tamm et al. [23].

Im Kleinanzeigenbereich, vor allem im Stellenbereich, wird in der Regel nur ein Benutzer benötigt, um eine Transaktion bezüglich einer bestimmten Anzeige abschließen. In diesem Fall die Vielfalt der Empfehlungen ist noch kritischer als in anderen Bereichen (wie Musik, Filme oder andere E-Commerce, ohne Kleinanzeigen). Um diesen Aspekt zu bewerten, verwenden wir drei Maßnahmen: Testabdeckung, Shannon Entropie [56] und der Gini-Index [56]. Die Testabdeckung bezieht sich auf den Anteil der Testitems mindestens einem Benutzer empfohlen. Mit Shannon-Entropie und dem Gini-Index ignorieren wir Elemente außerhalb des Testsatzes. Größere Testabdeckung und Shannon-Entropiewerte deuten auf höhere Vielfalt, während höhere Werte des Gini-Index eine geringere Vielfalt anzeigen.

Machine Translated by Google

Der letzte Schritt der Laborbewertung konzentriert sich auf die Effizienz. Zu diesem Zweck verwendeten wir eine AWS SageMaker ml.m5.4xlarge-Instanz mit 64 GB RAM und 16 vCPUs, mit einer Intel(R) Xeon(R) Platinum 8259CL CPU @ 2,50 GHz, läuft auf Amazon Linux AMI 2018.03. Alle Methoden wurden hinsichtlich dreier Funktionen evaluiert: Datenvorverarbeitung, Modellanpassung und Artikelempfehlung. Die Ausführungszeit wurde berechnet für die Gesamtzeit, die zur Berechnung des Modells benötigt wurde (in Stunden). Der Speicherpeak wurde berechnet in GB und die maximale Speichernutzung wurde beim Ausführen einer bestimmten Funktion berücksichtigt. Diese Auswertung wurde mit der Bibliothek tracemalloc durchgeführt (https://docs.python.org/3/bibliothek/tracemalloc.html, abgerufen am 29. Mai 2024).

5. Laborauswertung

In diesem Abschnitt wird die Laborauswertung der im entwickelten Datensatz diskutierten Methoden beschrieben. Wir haben die im vorherigen Abschnitt ermittelten Hyperparameter mit der besten Leistung verwendet. Jedes Modell wurde mit dem gesamten Trainings- und Validierungsdatensatz trainiert und Empfehlungen für alle 619.389 Benutzer aus dem Testdatensatz erstellt. Um die Werte von Für jede Offline-Metrik verglichen wir die beschriebenen Methoden mit einfachen, nicht personalisierten Baselines: der beliebteste Produktansatz und zufällige Empfehlungen.

5.1. Richtigkeit

Alle Metriken in diesem Abschnitt berücksichtigen die ersten zehn Empfehlungen, z. B. precision@10, recall@10. Höhere Metrikwerte weisen auf eine höhere Qualität der Empfehlungen hin. Die Werte sollte nicht direkt mit den Ergebnissen aus anderen Datensätzen verglichen werden, da Die Metriken hängen stark von der Verteilung des Datensatzes ab (z. B. hohe Spärlichkeit) und die Train/Test-Splitting-Strategie [55].

5.1.1. Bewertung der Genauigkeit

Tabelle 7 stellt die Ergebnisse der Evaluierung dar. Die meisten der beobachteten Unterschiede sind statistisch signifikant aufgrund der großen Anzahl ausgewerteter Benutzer. Wir gehen näher darauf ein die statistische Signifikanz der Ergebnisse in Abschnitt 5.1.2.

Tabelle 7. Genauigkeit: Ergebnisse der Evaluierung. Alle vorgestellten Metriken wurden von Tamm et al. [23] beschrieben. Für Für die mAP-Metrik haben wir die Variante mit = min(,) verwendet.

Metric RP3Beta	SLIM	ALS	Prod2Vec	LightFM	Beliebteste	Zufällig
Präzision 0,0484	0,0472	0,0434	0,0368	0,0736	0,0359	0,0012	0,00006
Rückruf 0,0783 NDC	0,0657	0,0580	0,0721	0,0657	0,0564	0,0012	0,00005
0,0759 MAP 0,0393	0,0567	0,0365	0,0329	0,0282	0,0545	0,0016	0,00007
MRR 0,1365 LAUC	0,1314	0,1230	0,1065	0,5368	0,0264	0,0006	0,00002
0,5391 HR 0,3131	0,5328	0. 5289	0,3066	0,2878	0,1034	0,0038	0,00019
	0,2537				0,5281	0,5006	0,49999
					0,2547	0,0112	0,00059

Quelle: eigene Studie.

Unsere Ansätze übertreffen die zufälligen und beliebtesten Artikelempfehlungen deutlich . Die schlechte Leistung des beliebtesten Ansatzes resultiert aus der Natur von Kleinanzeigen, vor allem die Domäne des Jobs, wo die Benutzer meist an einem bestimmten Standort und Kategorie. Aufbau ausgefeilterer, personalisierter Empfehlungssysteme ist in einem solchen Szenario sinnvoll.

Es wird gezeigt, dass RP3Beta andere Ansätze in Bezug auf Genauigkeit übertrifft Metriken. Wir glauben, dass dies mit der hohen Spärlichkeit unseres Datensatzes zusammenhängt. RP3Beta berechnet die Empfehlungen deterministisch (durch Nutzung von Pfaden der Länge drei auf einer Benutzer-Element-bipartiter Graph). Alle anderen Ansätze nutzen maschinelle Lerntechniken, um Benutzer zu finden Darstellungen (ALS, LightFM), Artikeldarstellungen (ALS, LightFM, Prod2Vec) oder direkte Artikel-Artikel-Ähnlichkeiten (SLIM), die in Fällen auftreten können, in denen Benutzer oder Artikel eine geringe Anzahl der Interaktionen.

Machine Translated by Google

5.1.2. Statistischer Vergleich für Präzision@10

Wir präsentieren einen detaillierten statistischen Vergleich für unsere Hauptmetrik, precision@10, um Unterschiede zwischen den Methoden zu identifizieren. Zunächst testen wir die Nullhypothese, dass alle Methoden gleich abschneiden und die beobachteten Unterschiede lediglich zufällig sind (Omnibus- Test). Der Friedman-Test [57,58] mit der **Erweiterung von Iman und Davenport [59]** ist wahrscheinlich **der beliebteste Omnibus-Test und ist normalerweise eine gute Wahl, wenn mehr als fünf verschiedene Algorithmen** verglichen werden **[60,61]. Sei der Rang der -ten Methode auf der -ten Stichprobe und**

$$= \frac{1}{} \ddot{y}\ddot{y}_{=1} \quad .$$

In unserem Fall ist die Anzahl der Stichproben gleich der Anzahl der Benutzer im Testsatz = 619.389.

Der Test vergleicht die mittleren Ränge der Methoden und basiert auf den folgenden Statistiken:

$$= \frac{(\ddot{y} 1)^{2}}{(\ddot{y} 1) \ddot{y}^{2}},$$

Wo

$$^{2} = \frac{12}{(+ 1)} \ddot{y}\ddot{y}_{=1} ^{2} \ddot{y} 3 (+ 1)$$

ist die Friedman-Statistik, die die Verteilung mit ÿ 1 und (ÿ 1) (ÿ 1) Freiheitsgraden hat . **Der -Wert aus diesem Test ist gleich 0. Der erhaltene -Wert zeigt an, dass wir die Nullhypothese, dass alle Algorithmen gleich abschneiden, sicher ablehnen können. Daher** können wir mit den Post-hoc-Tests fortfahren, um signifikante paarweise Unterschiede zwischen allen **Methoden festzustellen .** Demšar **[62] schlägt die Verwendung des Nemenyi-Tests [63] vor, der alle Algorithmen paarweise vergleicht . Für ein Signifikanzniveau bestimmt der Test den kritischen Unterschied (CD). Wenn der Unterschied zwischen den durchschnittlichen Ranglisten zweier Algorithmen größer ist als**

$$\mathbf{CD} = \ddot{y}\ddot{y} \frac{(+ 1)}{6},$$

Die Nullhypothese, dass die Algorithmen die gleiche Leistung aufweisen, wird abgelehnt (auf der basiert studentisierten Spannweitenstatistik geteilt durch ÿ 2). Demšar [62] schlug ein Diagramm vor, um die Unterschiede zu überprüfen (das CD-Diagramm visuell). Algorithmen, die nicht durch eine Linie verbunden sind, können im Diagramm als unterschiedlich betrachtet werden.

In unserem Fall, mit einer Signifikanz von = 0,05, sind zwei beliebige Algorithmen mit einem Unterschied in der Ein Durchschnittsrang über 0,0114 wird als ungleich angesehen (Abbildung 5).

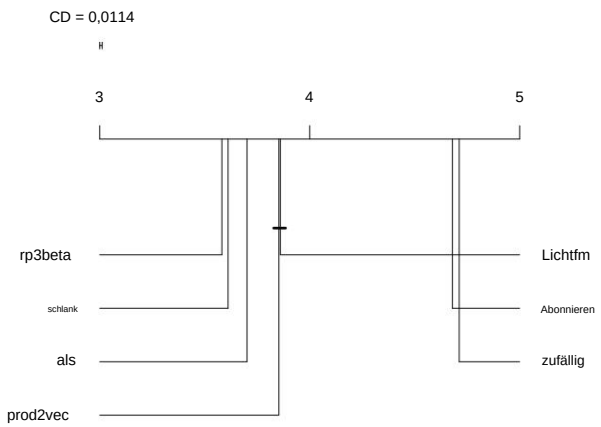


Abbildung 5. Diagramm der kritischen Differenz (für Präzision). Quelle: eigene Studie.

Machine Translated by Google

5.1.3. Genauigkeit für Benutzer mit unterschiedlicher Anzahl von Interaktionen

Die einzigen Informationen, die wir über unsere Benutzer haben, sind ihre Interaktionen. In Abbildung 6 untersuchen wir, wie die Anzahl der Elemente, mit denen ein Benutzer interagiert hat, unsere Hauptmetrik, precision@10, beeinflusst. Wir teilen unsere Benutzer basierend auf der Anzahl der Interaktionen in zehn ähnlich große Gruppen ein .

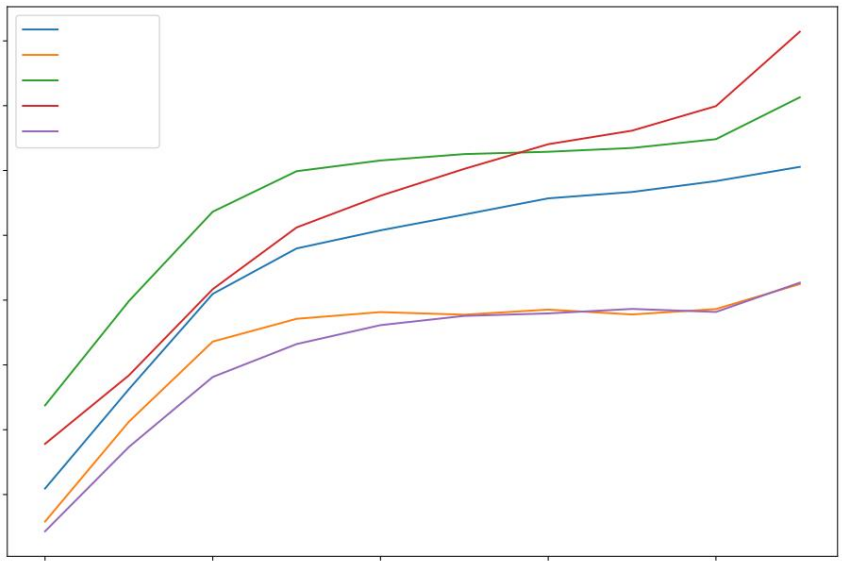


Abbildung 6. Präzision@10 für jedes Modell, abhängig von der Anzahl der Elemente, mit denen der Benutzer interagiert hat. Am besten lässt es sich in Farbe betrachten. Quelle: eigene Studie.

Wir können sehen, dass die Reihenfolge der nach precision@10 sortierten Modelle nicht von der Gruppe abhängt, außer dass SLIM bei Benutzern mit mindestens 22 Interaktionen besser abschneidet als RP3Beta. Der Unterschied ist für jede Benutzergruppe statistisch signifikant (siehe Abschnitt 5.1.4). Das RP3Beta-Modell kann als Sonderfall des itembasierten kollaborativen Filteransatzes angesehen werden, der ein deterministisches Ähnlichkeitsmaß zwischen Elementen verwendet [64]. Andererseits verwendet SLIM Methoden des maschinellen Lernens, um Ähnlichkeiten zu lernen.
Wir gehen davon aus, dass die so erlernten Ähnlichkeitswerte weniger verzerrt sind, aber eine höhere Varianz aufweisen, was zu einer besseren Leistung für Benutzer mit vielen Interaktionen führt, da die Mittelwertbildung über eine größere Anzahl von Werten erfolgt.

5.1.4. Statistische Auswertung der Genauigkeit für Benutzer mit unterschiedlicher Anzahl von Interaktionen

In diesem Unterabschnitt diskutieren wir den Unterschied zwischen den Modellen RP3Beta und SLIM abhängig von der Anzahl der Benutzerinteraktionen, wie in Abbildung 6 zu sehen ist.

Um zwei Methoden über mehrere Datensätze hinweg statistisch zu vergleichen, empfiehlt Demšar [62] den Wilcoxon-Vorzeichen-Rang-Test [65]. Der Wilcoxon-Vorzeichen-Rang-Test ist eine nicht-parametrische Alternative zum Paartest, der die Unterschiede in der Leistung zweier Algorithmen für jeden Datensatz unter Missachtung der Vorzeichen bewertet. Er vergleicht die Ränge der positiven und negativen Unterschiede.

Tabelle 8 zeigt, dass der Unterschied zwischen den Modellen RP3Beta und SLIM statistisch für jede Benutzergruppe von Bedeutung.

Machine Translated by Google

Tabelle 8. Ergebnisse des Wilcoxon-Vorzeichen-Rang-Tests zwischen RP3Beta- und SLIM-Modellen in Abhängigkeit von Anzahl der Elemente, mit denen der Benutzer interagiert hat.

Behälter	p-Wert
[1,0, 3,0)	0
[3,0, 5,0)	0
[5,0, 8,0)	0
[8,0, 11,0)	0
[11,0, 16,0)	0
[16,0, 22,0)	1,25 × 10ÿ6
[22,0, 31,0)	9,27 × 10ÿ7
[31,0, 45,0)	0
[45,0, 74,0)	0
[74,0, 852,0]	0

Quelle: eigene Studie.

5.2. Vielfalt

In Tabelle 9 können wir beobachten, dass Prod2Vec und LightFM die größte Vielfalt bieten nach Ausschluss zufälliger Empfehlungen (die die größte Vielfalt aufweisen sollten). Die Diversitäten der genauesten Modelle, RP3Beta, SLIM und ALS, sind ähnlich, mit Ausnahme der Fall der geringeren Testabdeckung von ALS. Die Beziehung zwischen Testabdeckung und Präzision@10 ist in Abbildung 7 dargestellt .

Tabelle 9. Vielfalt: Bewertungsergebnisse.

Metrisch	RP3Beta	SLIM	ALS	Prod2Vec	LightFM	Beliebteste	Zufällig
Testabdeckung	0,5725	0,5171	0,3038	0,7400	0,7031	0,0002	0,9778
Shannon	9,5271	9,6728	9,6270	10,4031	10,1385	2,3296	11.7267
Gini	0,9083	0,9029	0,9120	0,7956	0,8397	0,9999	0,1159

Quelle: eigene Studie.

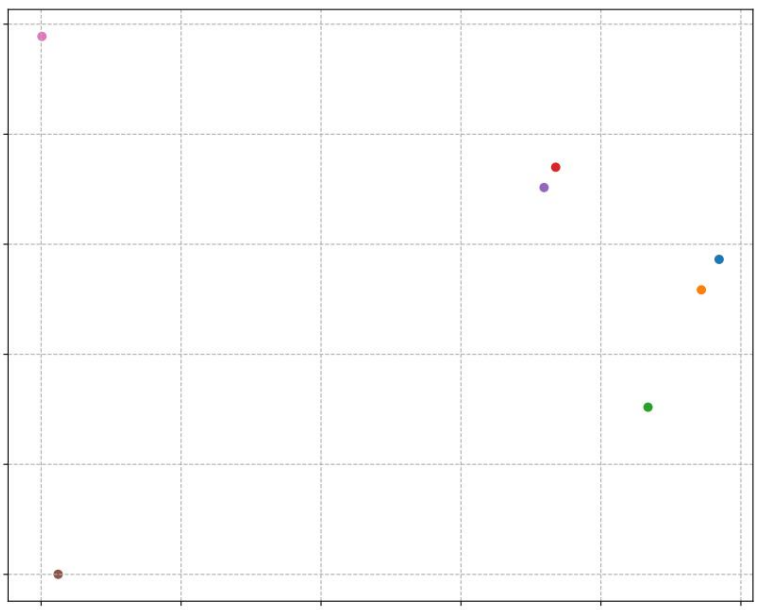


Abbildung 7. Testabdeckung im Vergleich zur Präzision@10 für die betrachteten Modelle. Quelle: eigene Studie.

Machine Translated by Google

5.3. Überschneidungen der Methoden

In unserer Forschung konzentrieren wir uns auf Methoden, die es wert sind, mit echten Benutzern getestet zu werden. Wir untersuchten die Überschneidungen zwischen den Empfehlungen der implementierten Methoden bezüglich der Benutzer-Item-Paare (Overlap-Koeffizient [66]). Wir erwarten signifikantere Unterschiede in den Empfehlungen beim Vergleich von Methoden mit geringer Überlappung.

Die Ergebnisse dieses Experiments sind in Tabelle 10 aufgeführt. RP3Beta und SLIM bieten ähnliche Empfehlungen. Obwohl sie unterschiedlich berechnet werden, verwenden sie eine spärliche Artikel Ähnlichkeitsmatrix. Die Vergleichsgruppe mit ALS weist mittlere Ähnlichkeiten mit anderen Methoden. Am unterschiedlichsten aus der Überlappungsperspektive sind auch die unterschiedlichsten Modelle, nämlich Prod2Vec und LightFM. Die geringe Überlappung von ALS und LightFM unterstreicht die Bedeutung der Wahl der Verlustfunktion bei Matrixfaktorisierungsansätzen.

Tabelle 10. Überschneidung der Modelle.

Modell	RP3Beta	SCHLANK	ALS	Prod2Vec	LichtFM
RP3Beta	100 %	73 %	53 %	37 %	38 %
SLIM	73 %	100 %	50 %	35 %	35 %
ALS	53 %	50 %	100 %	38 %	37 %
Prod2Vec	37 %	35 %	38 %	100 %	28 %
LightFM	38 %	35 %	37 %	28 %	100%

Quelle: eigene Untersuchung.

5.4. Bewertung der Effizienz

Die Ergebnisse der Effizienzbewertung sind in Abbildung 8 dargestellt. RP3Beta, ALS und LightFM übertrifft SLIM und Prod2Vec in Bezug auf die Ausführungszeit, wobei Prod2Vec den geringsten Speicherbedarf hat. Um das Modell häufig neu zu trainieren, haben wir die maximal akzeptable Ausführungszeit von 3 h, was von ALS, LightFM und RP3Beta erfüllt wurde. Aus Kostensicht Perspektive wurde die akzeptable Speicherauslastung auf 110 GB festgelegt, was alle Modelle erfüllt. Wir können beobachten, dass unsere Modelle sogar auf einem einzelnen Laptop trainiert werden konnten.

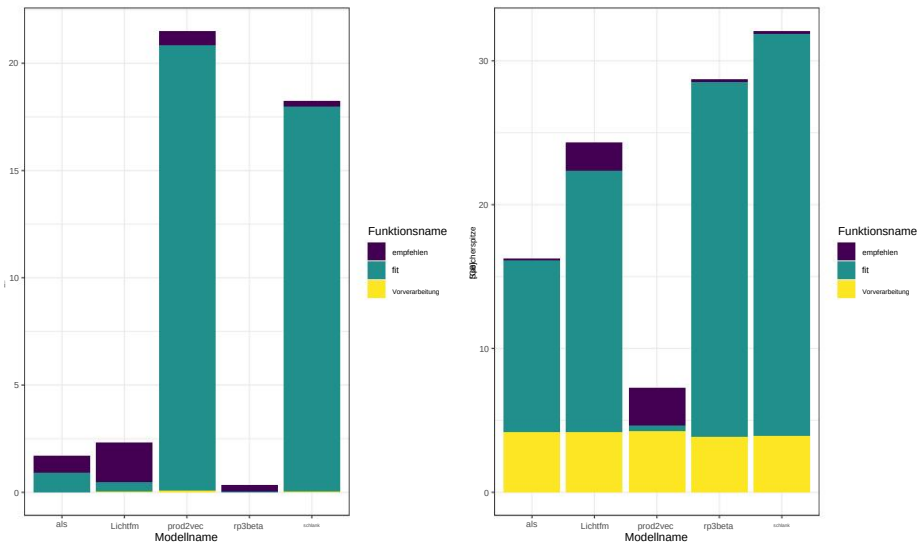


Abbildung 8. Ausführungszeit und Speicherauslastung. Am besten in Farbe betrachten. Quelle: eigene Studie.

5.5. Zusammenfassung

Durch die Laborforschung konnten wir Modelle auswählen, die online mit OLX getestet werden sollten Benutzer. Erstens haben wir das RP3Beta-Modell für Online-Tests gewählt, weil es der beste Ansatz ist hinsichtlich Genauigkeit und Ausführungszeit. Der relativ hohe Speicherverbrauch ist immer noch niedrig genug, um das Modell in der Produktion zu testen. Die Untersuchung der Auswirkungen dieser Die relativ geringe Diversität dieser Methode könnte ein spannender Forschungszweig sein.

Machine Translated by Google

SLIM ist das zweitbeste Modell in Bezug auf Präzision@10. Dennoch haben wir diese Methode aufgrund ihrer geringen Effizienz und der hohen Überschneidung mit dem bereits ausgewählten RP3Beta-Modell nicht für die Online-Evaluierung ausgewählt. Eine Verbesserung der Effizienz der SLIM-Implementierung könnte diese Entscheidung ändern.

Unter anderen Modellen erreicht ALS die höchste Präzision von 10. Darüber hinaus wurde ALS bereits bei OLX implementiert und für Stellenempfehlungen verwendet. Daher haben wir uns entschieden, es in den Online-Vergleich aufzunehmen.

Obwohl Prod2Vec und LightFM die vielfältigsten Empfehlungen liefern, haben wir uns aufgrund der deutlich schlechteren Präzision@10 und Effizienz von Prod2Vec dazu entschlossen, sie nicht online zu testen.

6. Online-Evaluation

Um die Wirksamkeit der ausgewählten Methoden zu bewerten, führten wir zwei Online-A/B- Tests mit OLX Jobs-Benutzern durch. Wir konzentrierten uns auf Benutzer, die kürzlich Stellenanzeigen auf unserer Plattform besucht hatten. Wir schickten ihnen E-Mails mit zehn empfohlenen Stellenanzeigen, und wenn ein Benutzer eine OLX-Anwendung installiert hatte, wurde zusätzlich eine Push-Benachrichtigung mit einer empfohlenen Stellenanzeige gesendet. Dann beobachteten wir, ob sich der Benutzer innerhalb der nächsten 48 Stunden auf eine Stelle bewarb (konvertierter Benutzer).

6.1. Auswirkungen des ALS-Modells

Das erste Experiment zielte darauf ab, die Frage zu beantworten: Erhöht die Einführung der Empfehlungsmethode die Zahl der Benutzer, die sich auf Stellen bewerben? In diesem Test haben wir die Benutzer in zwei Gruppen aufgeteilt: eine Kontrollgruppe ohne Empfehlungen und eine Gruppe mit Empfehlungen, die mithilfe des ALS-Modells generiert wurden. Nur ÿ 10 % der Benutzer wurden der Kontrollgruppe zugewiesen, da das ALS-Modell bereits vor dem Test implementiert und für Stellenempfehlungen bei OLX verwendet wurde. Wir zogen es vor, eine minimale Stichprobe zu sammeln, indem wir einen längeren Test durchführten, um eine längerfristige Auswirkung auf unsere Benutzer zu beobachten. Der Test wurde im März 2021 25 Tage lang durchgeführt. Die Ergebnisse sind in Tabelle 11 aufgeführt. **Die statistische Signifikanz des Vorteils von ALS gegenüber der Kontrollgruppe wurde mithilfe des Chi-Quadrat-Tests untersucht, und der -Wert ergab 0. Dieses Experiment bestätigte unsere Hypothese, dass Empfehlungen die Aktivität beeinflussen, da der Prozentsatz der konvertierten Benutzer nach Erhalt einer Mitteilung signifikant um (16,83 % ÿ 15,98 %)/15,98 % ÿ 5,3 % anstieg.**

Der beobachtete Anstieg bedeutet, dass sich täglich Hunderte zusätzlicher Benutzer über OLX Jobs auf Stellen bewerben.

Tabelle 11. Vergleich der Aktivität mit oder ohne Empfehlungen.

Variante	Benutzer	% Konvertierte Benutzer
Kontrolle	129.308	15,98 %
ALS	1.170.262	16,83 %

Quelle: eigene Studie.

6.2. Vergleich der ALS- und RP3Beta-Modelle

Im zweiten Experiment testeten wir, ob das beste Modell aus Labortests ALS übertraf. Wir teilten unsere Benutzer in drei Gruppen auf, die Empfehlungen von verschiedenen Empfehlungssystemen erhielten: ALS, RP3Beta und eine gemischte Variante. Wir generierten die Hälfte der Empfehlungen mit ALS und die andere Hälfte mit RP3Beta. Das zweite Experiment wurde 28 Tage lang im März und April 2021 durchgeführt. Die E

Der Prozentsatz der konvertierten Benutzer wird hier relativ zu allen Benutzern berechnet. Die Änderung des Empfehlungssystems hatte jedoch keine Auswirkungen auf Benutzer, die die Nachricht nicht geöffnet haben. Wir haben uns entschieden, diese herauszufiltern, da ihr Verhalten unseren Daten unnötiges Rauschen hinzufügte. Tabelle 13 zeigt die Auswirkungen auf Benutzer, die die Nachricht geöffnet haben. Der Vorteil von RP3Beta gegenüber ALS ist statistisch signifikant (p-Wert ÿ 10ÿ6). Der Unterschied zwischen den Varianten RP3Beta und ALS+RP3Beta ist statistisch nicht signifikant (der p-Wert beträgt 0,19).

Machine Translated by Google

Tabelle 12. Vergleich verschiedener Empfehlungsmethoden.

Variante	Benutzer	% Konvertierte Benutzer
ALS	343.892	15,25 %
RP3Beta	345.273	15,40 %
ALS+RP3Beta	343.896	15,30 %

Quelle: eigene Studie.

Tabelle 13. Vergleich verschiedener Empfehlungsmethoden für Benutzer, die die Nachricht geöffnet haben.

Variante	Benutzer	% Konvertierte Benutzer
ALS	44.775	19,66 %
RP3Beta	46.097	20,94 %
ALS+RP3Beta	45.469	20,59 %

Quelle: eigene Studie.

Von den Benutzern, die die Nachricht geöffnet haben, haben wir 1,28 Prozentpunkte beobachtet. mehr konvertierte Benutzer von RP3Beta als von ALS. Da die Nachrichtenöffnungsrate 13,2 % betrug, Dies entspricht einem Anstieg von 1,28 × 0,132 ÿ 0,17 Prozentpunkten aller Zielnutzer. Basierend auf der Ergebnisse des ersten Experiments, ALS bringt eine Steigerung von 0,85 Prozentpunkten gegenüber der Kontrollgruppe. Das bedeutet dass der Ersatz von ALS durch RP3Beta unsere Wirkung auf alle Zielbenutzer um ca. 0,17/0,85 = 20 %.

6.3. Diskussion

Wir haben nachgewiesen, dass das Senden von Jobempfehlungen aus dem ALS-Modell die Anzahl der Nutzer, die auf Stellenanzeigen antworten, um mehr als 5%. Darüber hinaus hat RP3Beta übertrifft ALS.

Zusätzlich führten wir einen weiteren Online-A/B-Test durch, bei dem wir das RP3Beta-Modell mit eine Variante, bei der wir die Hälfte der RP3Beta-Empfehlungen durch Empfehlungen ersetzt haben generiert aus dem Deep-Learning-Hybrid-Empfehlungssystem, das bei OLX verwendet wird (https://tech.olx.com/item2vec-neural-item-embeddings-to-enhance-recommendations-1fd948a6f293, abgerufen am 29. Mai 2024). Wir haben uns entschieden, diesen Test in diesem Artikel nicht zu beschreiben, da diese Arbeit konzentriert sich auf kollaborative Filteransätze. Wir wollten auch, dass unsere Ergebnisse auf dem veröffentlichten Datensatz reproduzierbar bleiben. Leider konnten wir nicht alle Funktionen veröffentlic von unserem internen Modell genutzt. Wir konnten keine Verbesserungen gegenüber RP3Beta feststellen, obwohl das zweite Modell zusätzliches Wissen über Stellenanzeigen nutzt. Daher bestätigte, dass ein richtig bestimmtes und abgestimmtes einfaches Modell mehr fortgeschrittene Tiefenlösungen. Dies steht im Einklang mit den Ergebnissen von Dacrema et al. [64] und Anelli et al. [30] verglichen mehrere nicht-neuronale Ansätze, darunter RP3Beta, mit fortgeschrittenere neuronale Modelle.

Die Effizienz des RP3Beta-Modells ermöglicht es OLX Jobs, es von Grund auf zu trainieren, mehrere mal pro Tag mit sehr geringen Kosten (auf einer CPU). Die hochmoderne kollaborative Filterung Das Modell bei OLX Jobs wird derzeit verwendet, um E-Mail- und Push-Empfehlungen für unsere Nutzer. Wir haben dieses Modell kürzlich angepasst, um Empfehlungen vor Ort in Echtzeit bereitzustellen, unter Berücksichtigung der jüngsten Benutzerinteraktionen mit Elementen [67]. Wir schlugen auch das P3-Lernen vor, um Rangmodell (P3LTR) [68], eine Verallgemeinerung des RP3Beta, bei der wir die Bedeutung von Benutzer-Element-Beziehungen basierend auf den Typen vergangener Benutzer-Element-Interaktionen und Zeitstempeln.

In dieser Arbeit haben wir uns mit der Forschungslücke des Mangels an Vergleichen von klassische kollaborative Filtering-Ansätze im Jobbereich durch umfassende Laborforschung an einem realen Datensatz und eine Online-Auswertung mit Millionen von Benutzern. Darüber hinaus haben wir den Quellcode für die Offline-Bewertung und den Datensatz öffentlich gemacht verfügbar. Wir ermutigen andere Forscher, neue Empfehlungsmethoden zu entwickeln, um helfen Benutzern, schneller ihren Traumjob zu finden.

Die in dieser Arbeit betrachteten Methoden erfordern nur (Benutzer-, Artikel- und Zeitstempel-)Tripletts als Eingangsdaten für Training und Evaluation; daher können sie in mehreren Bereichen eingesetzt werden,

Machine Translated by Google

sogar für Datensätze mit mehreren zehn Millionen Interaktionen. Aufgrund unserer ersten Ergebnisse in anderen OLX-Kategorien erwarten wir gute Ergebnisse im Bereich Kleinanz

7. Zusammenfassung

In diesem Artikel wurde das Thema Stellenempfehlungen in Online-Kleinanzeigen am Beispiel von OLX Jobs behandelt. Wir haben eine umfassende Bewertung von Ansätzen durchgeführt, die verschiedene Familien von Empfehlungsmethoden repräsentieren.

Während der Laborforschung haben wir gezeigt, dass RP3Beta, SLIM und ALS deutlich bessere Ergebnisse erzielen als LightFM und Prod2Vec. Wir haben auch festgestellt, dass Prod2Vec die außergewöhnlichste Vielfalt bietet, die bei der Arbeit an der Neuartigkeit von Empfehlungen erforderlich sein kann . Darüber hinaus haben wir untersucht, wie ähnlich sich die **von verschiedenen Modellen generierten Empfehlungen** sind.

Wir haben außerdem eine Auswertung durchgeführt, indem wir Millionen von Nachrichten an Online-Benutzer gesendet haben.

Es wurden A/B-Tests mit ALS und RP3Beta durchgeführt. Die Ergebnisse dieser Tests zeigten, dass das Senden von Jobempfehlungen, die von diesen Modellen generiert wurden, die Anzahl der Benutzer erhöht, die Werbetreibende von OLX Jobs kontaktieren. Mit RP3Beta in der Produktionsumgebung konnten wir unsere Effektivität um 20 % steigern (unter Berücksichtigung der Benutzerreaktionen).

Bei der Bewertung der Methoden aus der Rubrikenperspektive erwies sich ein Modell, das **die beliebtesten (noch nicht gesehenen) Artikel empfiehlt, für Stellenempfehlungen als ungeeignet.** Um **den Vorteil der untersuchten Methoden aufzuzeigen, wurde zum Vergleich zusätzlich ein Empfehlungsmodell auf Basis zufälliger Empfehlungen herangezogen .**

Ein weiteres wesentliches Ergebnis dieser Untersuchung ist der Datensatz, der weitere Untersuchungen zu Empfehlungen für klassifizierte Bedürfnisse ermöglicht.

Eine der Schwächen unserer Arbeit ist das Fehlen eines Vergleichs mit tiefen neuronalen Netzwerken (DNNs) oder graphisch neuronalen Netzwerken (GNNs). Bei solchen Methoden ist es eine Herausforderung, unsere Zeitbeschränkungen ohne kostspieligere Infrastruktur (z. B. mit GPU) einzuhalten. Für künftige Arbeiten wäre es interessant, unsere Beschränkungen zu lockern oder die effizientesten DNNs oder GNNs zu vergleichen . Eine weitere Schwäche besteht darin, dass der Vergleich nur auf einem einzigen Datensatz basiert. Die Größe des Datensatzes hat einen erheblichen Einfluss auf die Methodenauswahl. Es wäre interessant, dieselben Methoden an Datensätzen unterschiedlicher Größe zu vergleichen, entweder an Teilmengen des ursprünglichen Datensatzes oder an anderen Datensätzen aus dem Berufsbereich (z. B. aus verschiedenen Märkten, in denen OLX tätig ist). Darüber hinaus wäre es interessant, den Zusammenhang zwischen Diversität und Online-Ergebnissen zu untersuchen , um eine Offline-Kennzahl zu entwicke

Beiträge der Autoren: Konzeptualisierung, RK; Methodik, RK und TG; Software, RK und VD; Validierung, RK, VD, TG und AF; formale Analyse, RK und TG; Untersuchung, RK und TG; Ressourcen, RK; Datenkuratierung, RK und VD; Schreiben – Vorbereitung des Originalentwurfs, RK, AF und TG; Schreiben – Überprüfung und Bearbeitung, RK, AF und TG; Visualisierung, RK und TG; Überwachung, TG und AF; Projektverwaltung, RK. Alle Autoren haben die veröffentlichte Version **des Manuskripts gelesen und stimmen ihr zu .**

Finanzierung: Diese Forschung erhielt keine externe Finanzierung.

Datenverfügbarkeitserklärung: Datensatz: https://www.kaggle.com/datasets/olxdatascience/olx- jobs-interactions, abgerufen am 29. Mai 2024; Quellcode: https://github.com/rob-kwiec/olx-jobs-recommendations , abgerufen am 29. Mai 2024

Interessenkonflikte: Die Autoren Robert Kwieci ´nski, Agata Filipowska und Viacheslav Dubrov waren bei der Firma OLX Group beschäftigt. Der andere Autor erklärt, dass die Forschung ohne kommerzielle oder finanzielle Beziehungen durchgeführt wurde, die als potenzieller Interessenkonflikt ausgelegt werden könnten.

Verweise

1. Davis, M. Was sind Online-Kleinanzeigen? Big Easy Magazine, **4. Januar 2021. 2.**
Jones, S. Online-Kleinanzeigen; **Pew Internet & American Life Project: Washington, DC, USA, 2009.**
3. Gourville, J.; Soman, D. Überauswahl und Sortimentstyp: Wann und warum Vielfalt nach hinten losgeht. Mark. Sci. **2005,** 24, **382–395.**
[Querverweis]

Machine Translated by Google

4. Chernev, A. Wenn mehr weniger ist und weniger mehr ist: Die Rolle der idealen Punktverfügbarkeit und des Sortiments bei der Wahl des Verbrauchers. J. Consum. Res. **2003,** 30, **170–183. [CrossRef]**

5. Boatwright, P.; Nunes, J. Reducing Assortment: Ein attributbasierter Ansatz. J. Mark. **2001,** 65, **50–63. [CrossRef]**

6. Gómez-Uribe, C.; Hunt, N. Das Netflix-Empfehlungssystem. ACM Trans. Manag. Inf. Syst. **2015, 6, 1–19. [CrossRef]**

7. Smith, B.; Linden, G. Zwei Jahrzehnte Empfehlungssysteme bei Amazon.com. IEEE Internet Comput. **2017,** 21, **12–18. [CrossRef]**

8. Aggarwal, CC Recommender Systems: The Textbook, **1. Auflage; Springer: Berlin/Heidelberg, Deutschland, 2016. 9.**

Li, S.; Karahanna, E. Online-Empfehlungssysteme im B2C-E-Commerce-Kontext: Ein Überblick und zukünftige Richtungen. J. Assoc. Inf. Syst. **2015,** 16, **72–107. [CrossRef]**

10. Twardowski, B. Modellierung kontextbezogener Informationen in sitzungsabhängigen Empfehlungssystemen mit neuronalen Netzwerken. In Proceedings der 10. ACM-Konferenz zu Empfehlungssystemen, Boston, MA, USA, 15.–19. September 2016; S. 273–276. [Querverweis]

11. Alotaibi, S. Eine Untersuchung von Job-Empfehlungssystemen. Int. J. Phys. Sci. **2012, 7, 5127–5142. [CrossRef]**

12. Abel, F.; Benczúr, A.; Kohlsdorf, D.; Larson, M.; Pálovics, R. RecSys Challenge'16: Proceedings of the Recommender Systems Challenge; **Verband für Rechenmaschinen: New York, NY, USA, 2016.**

13. Abel, F.; Deldjoo, Y.; Elahi, M.; Kohlsdorf, D. RecSys Challenge'17: Proceedings der Recommender Systems Challenge 2017; **Association for Computing Machinery: New York, NY, USA, 2017.**

14. Li, S.; Shi, B.; Yang, J.; Yan, J.; Wang, S.; Chen, F.; He, Q. Tiefes Jobverständnis bei LinkedIn. In Proceedings der 43. Internationalen ACM SIGIR-Konferenz zu Forschung und Entwicklung im Bereich Informationsabruf, virtuell, 25.–30. Juli 2020; S. 2145–2148. [CrossRef]

15. Zhu, G.; Chen, Y.; Wang, S. Graph-Community-gestützte personalisierte Kurs-Job-Empfehlungen mit domänenübergreifenden Daten Integration. Sustainability **2022,** 14, **7439. [CrossRef]**

16. Lacic, E.; Reiter-Haas, M.; Kowald, D.; Reddy Dareddy, M.; Cho, J.; Lex, E. Verwendung von Autoencodern für sitzungsbasierte Jobempfehlungen dations. Benutzermodell. Benutzer-Adapt. Interagieren. **2020,** 30, **617–658. [CrossRef]**

17. Pazzani, MJ; Billsus, D. Inhaltsbasierte Empfehlungssysteme. In The Adaptive Web: Methoden und Strategien des Web Personalisierung; **Springer: Berlin/Heidelberg, Deutschland, 2007; S. 325–341. [CrossRef]**

18. Chen, R.; Hua, Q.; Chang, YS; Wang, B.; Zhang, L.; Kong, X. Eine Untersuchung kollaborativer, filterbasierter Empfehlungssysteme: Von traditionellen Methoden zu hybriden Methoden auf Basis sozialer Netzwerke. IEEE Access **2018, 6, 64301–64320. [CrossRef]**

19. Wang, X.; He, X.; Wang, M.; Feng, F.; Chua, TS Neural Graph Collaborative Filtering. In Proceedings der 42. Internationalen ACM SIGIR-Konferenz zu Forschung und Entwicklung im Bereich Informationsabruf, Paris, Frankreich, 21.–25. Juli 2019; S. 165–174. [Querverweis]

20. Schafer, JB; Frankowski, D.; Herlocker, J.; Sen, S. Collaborative filtering recommender systems. In The Adaptive Web; **Springer: Berlin/Heidelberg, Deutschland, 2007; S. 291–324. [CrossRef]**

21. Takács, G.; Pilászy, I.; Németh, B.; Tikk, D. Skalierbare kollaborative Filteransätze für große Empfehlungssysteme. J. Mach. Lernen. Res. **2009,** 10, **623–656. [CrossRef]**

22. Jawaheer, G.; Szomszor, M.; Kostkova, P. Vergleich von implizitem und explizitem Feedback eines Online-Musikempfehlungsdienstes . In Proceedings des 1. Internationalen Workshops zu Informationsheterogenität und -fusion in Empfehlungssystemen, Barcelona, Spanien, 26. September 2010; S. 47–51. [CrossRef]

23. Tamm, YM; Damdinov, R.; Vasilev, A. Qualitätsmetriken in Empfehlungssystemen: Berechnen wir Metriken konsistent? In Proceedings der fünfzehnten ACM-Konferenz zu Empfehlungssystemen, Amsterdam, Niederlande, 27. September–1. Oktober 2021; S. 708–713. [CrossRef]

24. Kaminskas, M.; Bridge, D. Diversität, Serendipität, Neuheit und Reichweite: Eine Umfrage und empirische Analyse von über die Genauigkeit hinausgehenden Zielen in Empfehlungssystemen. ACM Trans. Interact. Intell. Syst. **2016, 7, 1–42. [CrossRef]**

25. Kunaver, M.; Pozrl, T. Vielfalt in Empfehlungssystemen, eine Untersuchung. Knowl.-Based Syst. **2017,** 123, **154–162. [CrossRef]**

26. Ge, M.; Delgado-Battenfeld, C.; Jannach, D. Mehr als Genauigkeit: Bewertung von Empfehlungssystemen anhand von Abdeckung und Serendipität. In Proceedings der vierten ACM-Konferenz zu Empfehlungssystemen, Barcelona, Spanien, 26.–30. September 2010; S. 257–260. [Querverweis]

27. Mogenet, A.; Pham, TAN; Kazama, M.; Kong, J. Vorhersage der Online-Leistung von Job-Empfehlungssystemen mit Offline- Bewertung. In Proceedings der 13. ACM-Konferenz zu Empfehlungssystemen, Kopenhagen, Dänemark, 16.–20. September 2019; S. 477–480. [CrossRef]

28. Anelli, VW; Bellogín, A.; Di Noia, T.; Jannach, D.; Pomo, C. Top-N-Empfehlungsalgorithmen: Eine Suche nach dem neuesten Stand der Technik. In Proceedings der 30. ACM-Konferenz zu Benutzermodellierung, Anpassung und Personalisierung, Barcelona, Spanien, 4.–7. Juli 2022; S. 121–131. [CrossRef]

29. Dacrema, MF; Boglio, S.; Cremonesi, P.; Jannach, D. Eine beunruhigende Analyse der Reproduzierbarkeit und des Fortschritts bei Empfehlungs Systemforschung. ACM Trans. Inf. Syst. **2021,** 39, **1–49. [CrossRef]**

30. Anelli, VW; Bellogín, A.; Di Noia, T.; Pomo, C. Neugestaltung des Vergleichs zwischen neuronalem kollaborativem Filtern und Matrixfaktorisierung. In Proceedings der fünfzehnten ACM-Konferenz zu Empfehlungssystemen, Amsterdam, Niederlande, 27. September–1. Oktober 2021; S. 521–529. [CrossRef]

Machine Translated by Google

31. Zhu, J.; Dai, Q.; Su, L.; Ma, R.; Liu, J.; Cai, G.; Xiao, X.; Zhang, R. BARS: Auf dem Weg zum offenen Benchmarking für Empfehlungssysteme. In Proceedings der 45. Internationalen ACM SIGIR-Konferenz zu Forschung und Entwicklung im Bereich Informationsabruf, Madrid, Spanien, 11.–15. Juli 2022; S. 2912–2923. [CrossRef]
32. Dong, Y.; Li, J.; Schnabel, T. Wenn neuer nicht besser ist: Profitiert Deep Learning wirklich von Empfehlungen aus implizitem Feedback? In Proceedings der 46. Internationalen ACM SIGIR-Konferenz zu Forschung und Entwicklung im Bereich Informationsabruf , Taipeh, Taiwan, 23.–27. Juli 2023; S. 942–952. [CrossRef]
33. Anelli, VW; Malitesta, D.; Pomo, C.; Bellogin, A.; Di Sciascio, E.; Di Noia, T. Den Mythos des Graph Collaborative Filtering in Frage stellen: Eine begründete und reproduzierbarkeitsorientierte Analyse. In Proceedings der 17. ACM-Konferenz zu Empfehlungssystemen , Singapur, 18.–22. September 2023; S. 350–361. [CrossRef]
34. Koren, Y.; Bell, R.; Volinsky, C. Matrixfaktorisierungstechniken für Empfehlungssysteme. Computer **2009**, 42, **30–37.** [CrossRef]
35. Kula, M. Metadateneinbettungen für Benutzer- und Artikel-Kaltstart-Empfehlungen. arXiv **2015, arXiv:1507.08439.**
36. Rendle, S.; Freudenthaler, C.; Gantner, Z.; Schmidt-Thieme, L. BPR: Bayesian Personalized Ranking from Implicit Feedback. arXiv **2012, arXiv:1205.2618 37. Weston,**
J.; Bengio, S.; Usunier, N. Wsabie: Scaling up to large vocabulary image annotation. In Proceedings der 22. Internationalen Gemeinsamen Konferenz für Künstliche Intelligenz, Barcelona, Spanien, 16.–22. Juli 2011; S. 2764–2770.
38. Weston, J.; Yee, H.; Weiss, RJ Lernen, Empfehlungen mit dem K-Order-Statistikverlust zu bewerten. In Proceedings der 7. ACM-Konferenz über Empfehlungssysteme, Hongkong, China, 12.–16. Oktober 2013; S. 245–248. [CrossRef]
39. Hu, Y.; Koren, Y.; Volinsky, C. Kollaboratives Filtern für implizite Feedback-Datensätze. In Proceedings der 2008 8. IEEE Internationale Konferenz zum Thema Data Mining, Pisa, Italien, 15.–19. Dezember 2008; S. 263–272. [CrossRef]
40. Takács, G.; Pilászy, I.; Tikk, D. Anwendungen der konjugierten Gradientenmethode für kollaboratives Filtern mit implizitem Feedback. In Proceedings der fünften ACM-Konferenz zu Empfehlungssystemen, Chicago, IL, USA, 23.–27. Oktober 2011; S. 297–300. [Querverweis]
41. Ning, X.; Karypis, G. SLIM: Sparse Linear Methods for Top-N Recommender Systems. In Proceedings der 2011 IEEE 11th International Conference on Data Mining, Vancouver, BC, Kanada, 11.–14. Dezember 2011; S. 497–506. [CrossRef]
42. Christoffel, F.; Paudel, B.; Newell, C.; Bernstein, A. Blockbusters and Wallflowers: Accurate, Diverse, and Scalable Recommen- dations with Random Walks. In Proceedings der 9. ACM-Konferenz zu Empfehlungssystemen, Wien, Österreich, 16.–20. September 2015; S. 163–170. [CrossRef]
43. Cooper, C.; Lee, SH; Radzik, T.; Siantos, Y. Zufallswanderungen in Empfehlungssystemen: Genaue Berechnung und Simulationen. In Proceedings der 23. Internationalen Konferenz zum World Wide Web, New York, NY, USA, 7.–11. April 2014; S. 811–816. [Querverweis]
44. Grbovic, M.; Radosavljevic, V.; Djuric, N.; Bhamidipati, N.; Savla, J.; Bhagwan, V.; Sharp, D. E-Commerce in Ihrem Posteingang: Produktempfehlungen im großen Maßstab. In Proceedings der 21. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australien, 10.–13. August 2015; S. 1809–1818. [CrossRef]
45. Barkan, O.; Koenigstein, N. ITEM2VEC: Neurale Item-Einbettung für kollaboratives Filtern. In Proceedings des 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP), Vietri sul Mare, Italien, 13.–16. September 2016; S. 1–6. [Querverweis]
46. Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.; Dean, J. Verteilte Darstellungen von Wörtern und Phrasen und ihre Kompositionalität. In Proceedings der 26. Internationalen Konferenz über neuronale Informationsverarbeitungssysteme – Band 2, Red Hook, NY, USA, 5.–10. Dezember 2013; S. 3111–3119.
47. De Ruijt, C.; Bhulai, S. Job-Empfehlungssysteme: Eine Überprüfung. arXiv **2021, arXiv:2111.13576.**
48. Kula, M. LightFM—Github Repository. 2023. Online verfügbar: https://github.com/lyst/lightfm (abgerufen am 29. Mai 2024).
49. Frederickson, B. ALS—Github Repository. 2022. Online verfügbar: https://github.com/benfred/implicit/blob/main/implicit/als.py (abgerufen am 29. Mai 2024).
50. Levy, M.; Grisel, O. SLIM—Github Repository. 2014. Online verfügbar: https://github.com/Mendeley/mrec/blob/master/ mrec/item_similarity/slim.py (abgerufen am 29. Mai 2024).
51. Quadrana, M. SLIM—Github Repository. 2021. Online verfügbar: https://github.com/MaurizioFD/RecSys2019_DeepLearning_ Evaluation/blob/master/SLIM_ElasticNet/SLIMElasticNetRecommender.py (abgerufen am 29. Mai 2024).
52. Bernardis, C. RP3Beta—Github Repository. 2021. Online verfügbar: https://github.com/MaurizioFD/RecSys2019_DeepLearning_ Auswertung/blob/master/GraphBased/RP3betaRecommender.py (abgerufen am 29. Mai 2024).
53. Paudel, B.; Christoffel, F.; Newell, C.; Bernstein, A. Aktualisierbare, genaue, vielfältige und skalierbare Empfehlungen für interaktive Anwendungen. ACM Trans. Interact. Intell. Syst. **2016, 7, 1–34.** [CrossRef]
54. Reh ůuřek, R.; Sojka, P. Software Framework for Topic Modelling with Large Corpora. In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, Valletta, Malta, 22. Mai 2010; S. 45–50, ISBN 2-9517408-6-7. Online verfügbar: http://is.muni.cz/publication/884893/en (abgerufen am 29. Mai 2012).
55. Meng, Z.; McCreadie, R.; Macdonald, C.; Ounis, I. Erforschung von Datenaufteilungsstrategien zur Bewertung von Empfehlungsmodellen . In Proceedings der 14. ACM-Konferenz zu Empfehlungssystemen, virtuell, 22.–26. September 2020; S. 681–686. [Querverweis]
56. Shani, G.; Gunawardana, A., Evaluating Recommendation Systems. In Recommender Systems Handbook; **Springer: Boston, MA, USA, 2011; S. 257–297.** [CrossRef]

Machine Translated by Google

57. Friedman, M. Die Verwendung von Rängen zur Vermeidung der impliziten Annahme der Normalität in der Varianzanalyse. J. Am. Stat. Assoc. **1937,** 32, **675–701.** [CrossRef]

58. Friedman, M. Ein Vergleich alternativer Signifikanztests für das Problem der Rangfolge. Ann. Math. Stat. **1940,** 11, **86–92.** [Querverweis]

59. Iman, RL; Davenport, JM Näherungen an den kritischen Bereich der Friedman-Statistik. Commun. Stat.-Theory Methods **1980,** 9, 571–595. [CrossRef]

60. García, S.; Herrera, F. Eine Erweiterung zum „Statistischen Vergleich von Klassifikatoren über mehrere Datensätze“ für alle paarweisen Vergleiche. J. Mach. Learn. Res. **2008, 9, 2677–2694.**

61. García, S.; Fernández, A.; Luengo, J.; Herrera, F. Erweiterte nichtparametrische Tests für Mehrfachvergleiche bei der Versuchsplanung in Computational Intelligence und Data Mining: Experimentelle Leistungsanalyse. Inf. Sci. **2010,** 180, **2044-** [Querverweis]

62. Demšar, J. Statistische Vergleiche von Klassifikatoren über mehrere Datensätze. J. Mach. Learn. Res. **2006, 7, 1–30.**

63. Nemenyi, P. Verteilungsfreie Mehrfachvergleiche. Dissertation, Princeton University, Princeton, NJ, USA, 1963.

64. Dacrema, MF; Cremonesi, P.; Jannach, D. Machen wir wirklich große Fortschritte? Eine beunruhigende Analyse der jüngsten neuronalen Empfehlungsansätze. In Proceedings der 13. ACM-Konferenz zu Empfehlungssystemen, Kopenhagen, Dänemark, 16.–20. September 2019; S. 101–109. [CrossRef]

65. Wilcoxon, F. Individuelle Vergleiche durch Rankingmethoden. Biom. Bull. **1945, 1, 80–83.** [CrossRef]

66. Vijaymeena, M.; Kavitha, K. Eine Untersuchung zu Ähnlichkeitsmaßen im Text Mining. Mach. Learn. Appl. Int. J. **2016, 3, 19–28.** [CrossRef]

67. Kwieci ´nski, R.; Melniczak, G.; Górecki, T. Vergleich von Empfehlungen für Echtzeit- und Batch-Jobs. IEEE Access **2023,** 11, **20553–20559.** [CrossRef]

68. Kwieci ´nski, R.; Górecki, T.; Filipowska, A. Lernen der Kantenwichtigkeit in bipartiten, graphenbasierten Empfehlungen. In Proceedings der 17. Konferenz für Informatik und Intelligenzsysteme (FedCSIS) 2022, Sofia, Bulgarien, 4.–7. September 2022; S. 227–233. [CrossRef]

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Blockchain und homomorphe Verschlüsselung für Datensicherheit un

Statistischer Datenschutz

Rahul Raj 1,*, Ye¸sem Kurt Peker [1] und Zeynep Delal Mutlu [2]

[1] TSYS School of Computer Science, Columbus State University, Columbus, GA 31907, USA;
[2] peker_yesem@columbusstate.edu Unabhängiger Forscher, 1200 Brüssel, Belgien; zeynepdelalmutlu@gmail.com * Korrespondenz: raj_rahul@students.columbusstate.edu

Zusammenfassung: Diese Studie schlägt ein Blockchain-basiertes System vor, das vollständig homomorphe Verschlüsselung verwendet, um Datensicherheit und statistische Privatsphäre zu gewährleisten, wenn Daten zu Analyse- oder Forschungszwecken an Dritte weitergegeben werden. Das vorgeschlagene System bietet nicht nur Sicherheit für Daten während der Übertragung, im Ruhezustand und bei der Verwendung, sondern gewährleistet auch Privatsphäre und Rechenintegrität für einfache statistische Berechnungen. Dies wird erreicht, indem die Eigenschaften der Blockchain-Technologie genutzt werden, die Verfügbarkeit und Datenintegrität bietet, kombiniert mit homomorpher Verschlüsselung, die die Vertraulichkeit der verwendeten Daten gewährleistet. Die Berechnungen werden auf Smart Contracts durchgeführt, die sich auf der Blockchain befinden, was Rechenintegrität gewährleistet. Das vorgeschlagene System ist auf der Zama-Blockchain implementiert und führt statistische Operationen einschließlich Mittelwert, Median und Varianz auf verschlüsselten Daten durch. Die Ergebnisse zeigen, dass es möglich ist, vollständig homomorphe Berechnungen auf der Blockchain durchzuführen. Auch wenn die aktuellen Rechenbeschränkungen der Blockchain es nicht zulassen, das System für große Datensätze auszuführen, ist die Technologie verfügbar und mit den Fortschritten bei effizienteren homomorphen Operationen auf Blockchains wird das vorgeschlagene System eine ultimative Lösung für die Bereitstellung der begehrten Sicherheitseigenschaften in Anwendungen bieten, einschließlich Datenschutz und statistischer Privatsphäre, Vertraulichkeit

Schlüsselwörter: Datenschutz; Datenintegrität; rechnerische Integrität; homomorphe Verschlüsselung; Blockchain; statistischer Datenschutz; Datensicherheit; statistische Vertraulichkeit

check for updates

Zitat: Raj, R.; Kurt Peker, Y.; Mutlu, ZD Blockchain und Homomorphe Verschlüsselung für Datensicherheit und statistischen Datenschutz. Electronics **2024,** 13, **3050. https:// doi.org/ 10.3390/electronics13153050**

Wissenschaftlicher Redakteur: Simeone Marino

Erhalten: 8. Juni 2024
Überarbeitet: 20. Juli 2024
Akzeptiert: 26. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

Der Schutz der Privatsphäre der Benutzer und die Vertraulichkeit sensibler Daten ist in allen Systemen, in denen Benutzerdaten gespeichert, übertragen oder verarbeitet werden, von entscheidender Bedeutung. Häufig müssen Daten für geschäftliche Zwecke oder zum Nutzen der Öffentlichkeit analysiert werden. Volkszählungsdaten erfassen beispielsweise Informationen über die Bevölkerung eines Landes, eines Staates, einer Stadt oder anderer genau definierter geografischer Gebiete. Sie enthalten in vielen Fällen sensible Informationen wie Alter, Rasse, Geschlecht, religiöse Überzeugungen und langfristige Gesundheitszustände. Volkszählungsdaten sind die wichtigsten Daten, die verwendet werden, um die sozialen, wirtschaftlichen und demografischen Bedingungen auf lokaler und nationaler Ebene zu verstehen , und liefern Informationen, die Regierungen benötigen, um Richtlinien zu entwickeln, öffentliche Dienste zu planen und bereitzustellen und Mittel zuzuweisen. Der Schutz von Volkszählungsdaten ist von der Erfassung bis zur Verarbeitung von wesentlicher Bedeutung. Eine Schutzmaßnahme besteht darin, nur autorisierten Personen Zugriff auf die Daten zu gewähren. Verschlüsselung ist ein gängiger Mechanismus zur Umsetzung dieser Schutzmaßnahme. Daten werden im Speicher verschlüsselt und zwischen den autorisierten Benutzern übertragen. Es besteht jedoch das Risiko unbeabsichtigter Fehler autorisierter Benutzer oder eines böswilligen Benutzers, der die Sicherheit und den Datenschutz der Daten sowie die Integrität der Analyse gefährdet. In manchen Fällen müssen Daten Dritten (z. B. Forschern) zugänglich gemacht werden, damit eine ordnungsgemäße und objektive Analyse durchgeführt werden kann. In diesen Fällen werden die Daten zum Schutz der Privatsphäre der Benutzer vor der Weitergabe anonymisiert

Machine Translated by Google

Bedrohungen: Wenn große Datensätze verfügbar gemacht werden, können Benutzerinformationen aufgrund ausgefeilterer Angriffe verloren gehen. Dies gilt insbesondere, wenn große Datensätze verfügbar gemacht werden. Die Studie „Exposed! A Survey of Attacks on Private Data“ von Dwork et al. aus dem Jahr 2016 bietet eine Übersicht über gängige Angriffe auf aggregierte Daten. Sie beschreibt Rekonstruktions- und R Bei Rekonstruktionsangriffen geht es darum, sensible Merkmale von Personen in einem Datensatz zu approximieren. Bei Tracing-Angriffen wird ermittelt, ob die Daten einer Zielperson im Datensatz enthalten sind oder nicht [1]. Auf Dworks Studie folgen zahlreiche Studien, die sich mit Angriffen auf den Datenschutz und Abwehrmechanismen dagegen beschäftigen [2–6].

Zwei Begriffe, die mit der Sicherheit von Daten im statistischen Sinne in Zusammenhang stehen, sind statistischer Datenschutz und statistische Vertraulichkeit. Statistischer Datenschutz ist „die Kunst, einen Datenschutzmechanismus zu entwickeln, der sensible Daten in Daten umwandelt, die gleichzeitig nützlich und nicht sensibel sind“ [7]. Das Ziel statistischer Datenschutzmethoden besteht darin, „einen breiten Datenaustausch über verschiedene Datenkontexte und Domänen hinweg zu ermöglichen, in denen der Schutz der Identität oder sensibler Attribute von Einzelpersonen gewünscht oder erforderlich ist (z. B. Volkszählung, Gesundheit, Genomdaten , soziale Netzwerke)“ [8].

Unter statistischer Geheimhaltung versteht man den „Schutz von Informationen einzelner statistischer Einheiten“ oder, gleichbedeutend, „den Schutz von Daten, die für statistische Zwecke erhoben v Die statistische Geheimhaltung ist eng mit dem statistischen Datenschutz verbunden. Obwohl in der Literatur keine klare Unterscheidung zwischen diesen Begriffen besteht, verstehen wir unter statistischer Geheimhaltung auf Grundlage der in diesem Artikel zitierten Definitionen den Schutz statistischer Einheiten, d. h. der über Einzelpersonen oder Unternehmen erhobenen Daten, und die Gewährleistung ihrer genehmigten Verwendung . Statistischer Datenschutz ist das umfassendere Sicherheitskonzept, das die Umwandlung sensibler Daten in nicht sensible und nützliche Daten umfasst. Die statistische Geheimhaltung ist eines der Grundprinzipien der offiziellen Statistik, das von der Generalversammlung der Vereinten Nationen gebilligt wurde **[10]. Es ist auch ein Grundprinzip der europäischen Statistik, das Prinzipien, Regeln und Verfahren definiert, um vertrauliche Daten zu schützen und gleichzeitig ihre Verwendung für statistische Zwecke** zu ermöglichen **[11].**

Die erste Veröffentlichung von Innovations in Federal Statistics [12] befasst sich mit den Herausforderungen und der Notwendigkeit eines neuen Paradigmas im föderalen Statistiksystem. Kapitel 5 der Veröffentlichung mit dem Titel „Schutz von Privatsphäre und Vertraulichkeit bei gleichzeitiger Bereitstellung des Zugangs zu Daten für Forschungszwecke“ konzentriert sich auf den Schutz der Vertraulichkeit der für statistische Zwecke gesammelten Informationen. Nach der Erörterung der Sicherheitsimplikationen der Verknüpfung mehrerer Datensätze für die Analyse, einschließlich der rechtlichen Grundlagen für Privatsphäre und Vertraulichkeit, ziehen die Autoren zwei wichtige Schlussfolgerungen: „Wenn die föderalen Statistikbehörden mit der Verknüpfung **mehrerer Datensätze fortfahren ,** müssen sie sich gleichzeitig mit der Quantifizierung und Kontrolle des Risikos **des Datenschutzverlusts befassen“ und „Datenschutzverbessernde Techniken und eine datenschutzwahrende statistische Datenanalyse kön Aufgrund dieser Schlussfolgerungen empfehlen sie: „Die statistischen Bundesbehörden sollten moderne Datenbank-, Kryptografie- und datenschutzschützende sowie datenschutzverbessernde Technologien einführen.“**

In dieser Studie schlagen wir ein System vor, das Datensicherheit, statistische Privatsphäre und Vertraulichkeit bietet und dabei ausschließlich Techniken und Werkzeuge aus der Mathematik und Informatik verwendet . Wir untersuchen die Verwendung eines Blockchain-basierten Systems, das die Durchführung statistischer Berechnungen mit verschlüsselten Daten ermöglicht, indem mithilfe von Smart Contracts eine vollständig homomorphe Verschlüsselung auf der Blockchain verwendet wird . Das vorgeschlagene System bietet End-to-End-Sicherheit der Daten von dem Zeitpunkt an, an dem sie den Dateneigentümer verlassen, bis zu dem Zeitpunkt, an dem der Forscher die Ergebnisse der statistischen Berechnungen erhält. Die bereitgestellte Sicherheit umfasst die Integrität und Vertraulichkeit der Daten während der Übertragung, im Ruhezustand und bei der Verwendung. Nicht nur werden die einzelnen Datenpunkte nie im Klartext offengelegt, sondern die Ergebnisse sind nur autorisierten Benutzern im Klartext verfügbar. Da die Berechnungen auf der Blockchain durchgeführt werden, ist es außerdem nahezu unmöglich, die Algorithmen für die statistischen Berechnungen zu manipulieren. Darüber hinaus ist es nicht möglich, andere Statistiken anzufordern als die, die im Smart Contract genehmigt sind. Somit bietet das vorgeschlagene System die statistische Privatsphäre und Vertraulichkeit, die in verschiedenen Anwendungen, einschließlich Bundesstatistiksystemen, gewünscht wird. In Abschnitt 2 geben wir kurze Beschreibu

Machine Translated by Google

In **Abschnitt 3 beschreiben wir unser vorgeschlagenes System und die Implementierung eines Proof of Concept für das System. In Abschnitt 4 präsentieren wir die Ergebnisse der Durchführung des Proof of Concept und schließlich liefern wir in Abschnitt 5 eine Diskussion der Ergebnisse.**

2. Verwandte Arbeiten

Zahlreiche Studien haben versucht, Datensicherheit und -integrität nicht nur im Ruhezustand, sondern auch während der Nutzung durch verschiedene Mittel, darunter homomorphe Verschlüsselung und Blockchain-Technologie, zu gewährleisten. Eine dieser Studien schlägt ein System für ein kollaboratives Datentrainingsparadigma für den Austausch medizinischer Bilddaten vor, bei dem das Blockchain-Ledger die Dezentralisierung der föderierten Lernmodelle ohne Abhängigkeit von einem zentralen Server ermöglicht und die homomorphe Verschlüsselung die Bedenken im Zusammenhang **mit dem Austausch von Rohdaten lindert [13]. Ihr System verschlüsselt die Gradienten homomorph und gibt sie über die Blockchain frei .**

Liang et al. [14] schlagen die Integration von Blockchain und homomorpher Verschlüsselung vor, um die Herausforderungen beim Urheberrechtsschutz von Schaltkreisen zu bewältigen. Ihre Studie schlägt ein auf homomorpher Verschlüsselung basierendes mathematisches Modell innerhalb der Blockchain vor, das die Transaktionen sichert und gleichzeitig durch die Verwendung von Smart Contracts die Integrität und Vertraulichkeit der Da

Yaji et al. [15] schlugen ein System vor, das Goldwasser-Micali- und Paillier-Verschlüsselungsschemata für die vergleichende Evaluierungsstudie mit Schwerpunkt auf Datenschutztechniken unter Verwendung der Blockchain-Technologie für KI-Anwendungen verwendet. Die Studie ergab, dass Angriffe auf die Blockchain wie Kollision, Preimage und Angriffe auf das Wallet durch **die Verschlüsselung von Blöcken mit den vorgeschlagenen Goldwasser-Micali- und Paillier-**

Mutlu et al. [16] schlugen ein System vor, das Blockchain-Technologie und homomorphe Verschlüsselung verwendet und es Dritten (Forschern) ermöglicht, lineare Regressionen auf verschlüsselten Daten durchzuführen. Sie verwenden den Pallier-Algorithmus, um die für die lineare Regression erforderliche Summe durch Smart Contracts zu berechnen. Der Dateneigentümer verschlüsselt die Daten mit dem öffentlichen Schlüssel des Forschers und sendet sie an den Smart Contract, wo die Berechnung durchgeführt wird. Das verschlüsselte Ergebnis kann dann vom Forscher abgerufen werden, der es auf seinem System mit seinem privaten Schlüssel entschlüsseln kann .

Vanin et al. [17] schlagen ein Modell zur Sicherung von persönlichen Gesundheitsdaten (PHR) vor, das ein interplanetares Protokolldateisystem basierend auf verteilten Hashtabellen (DHTs) zusammen mit der Blockchain verwendet. PHR-Metadaten werden auf der Blockchain gespeichert und über das Netzwerk geteilt, während PHR-Daten außerhalb der Blockchain über das IPFS-Netzwerk gespeichert werden. Sie verwenden zwei Elemente: Data Steward (DS), der für die Speicherung von PHR im Namen des Einzelnen verantwortlich ist, und Shared Data Vault (SDV), einen temporären IPFS-Speicherbereich, in dem Gesundheitseinrichtungen mit Zustimmung des Einzelnen auf PHR zugreifen können. Verschlüsselte Daten sind der Öffentlichkeit über statistische Portale zugänglich, wo sie mit homomorpher Verschlüsselung Operationen an den Daten durchführen können, um aussagekräftige Ergebnisse zu erhalten. Zu diesem Zweck verwenden sie die Microsoft Simple Encrypted Arithmetic Library (SEAL), die den BFV-Algorithmus in JavaScript implementiert.

Umar et al. [18] schlugen ein Modell für E-Voting unter Verwendung des Paillier-Algorithmus vor. Sobald der Wähler seine Stimme abgibt, wird diese homomorph verschlüsselt. Anschließend wird ein neuer Block der Transaktion erstellt, der die verschlüsselten Stimmzettel, die pseudonymen Adressen des Wählers und des Administrators, den Zeitstempel der Blockerstellung, den Hash des vorherigen Blocks der Transaktion sowie den Hash des aktuellen Blocks enthält. Anschließend wird der neue Block der Transaktion mithilfe des Konsensmechanismus abgebaut. Nach dem Abbau wird der neue Block in das Hauptbuch der Blockchain eingetragen. Dieser Prozess wird bis zum Ende der Wahl fortgesetzt. Danach zählt der Administrator die verschlüsselten Stimmen mithilfe des Paillier-Algorithmus aus, wodurch die Endsumme ermittelt wird , die dann entschlüsselt werden kann, um das Ergebnis zu ermitteln.

Shrestha et al. [19] haben eine Studie durchgeführt, die die Sicherheitsbedenken im Zusammenhang mit dem Internet der Dinge (IoT) analysiert. Eines der größten Bedenken, das hervorgehoben wurde, ist der Datenschutz . Diese Studie untersucht auch die Möglichkeit, das IoT mit der Blockchain **und mit homomorpher Verschlüsselung zu integrieren .** Zu den Vorteilen gehören die Unveränderlichkeit und Unvergesslichkeit der Daten, **die Beseitigung einzelner Ausfallpunkte und die Vertraulichkeit der Daten.**

Machine Translated by Google

Caldarola et al. [20] schlugen das Neural Fairness Protocol vor, einen Konsensmechanismus , der den Elliptic Curve Digital Signature Algorithm (ECDSA) und kryptografische Techniken integriert. Dieses Protokoll verbessert die Anonymität und Verantwortlichkeit bei Blockchain-Transaktionen, insbesondere durch die Entwicklung eines robus Über die anfängliche Anwendung hinaus bietet diese Innovation Potenzial zur Verbesserung der Sicherheit verschiedener Kryptowährungen wie Litecoin und Mastercoin.

Das in dieser Studie vorgeschlagene System kann mit der Arbeit von Liang [14], Yaji [15] und Mutlu [16] verglichen werden , **da es Blockchain-Technologie zusammen mit homomorpher Verschlüsselung verwendet, um die Vertraulichkeit der Daten während ihrer Verwendung zu gewährleisten. Darüber hinaus** verwendet das von Mutlu **[14] vorgeschlagene System eine teilweise homomorphe Verschlüsselung über den Smart Contract, die der Architektur unseres vorgeschlagenen Systems ähnelt. Die oben genannten Studien verwenden jedoch den Pallier-Algorithmus, der teilweise homomorph ist, während diese Studie ein vollständig homomorphes System verwendet, das nicht nur Addition, sondern auch Multiplikation der verschlüsselten Daten ermöglicht, wodurch eine größere Bandbreite an Berechnungen sicher durchgeführt werden kann.**

3. Materialien und Methoden

In diesem Abschnitt beschreiben wir ein System, das den sicheren Datenaustausch und die Datenanalyse zwischen Dateneigentümern und Dritten ermöglicht. Die Dateneigentümer sind Unternehmen oder Einzelpersonen, die Zugriff auf die Klardaten haben. Man könnte sie auch als Sensoren betrachten, die Daten verarbeiten und übertragen können. Die Eigentümer sind autorisiert oder geben ihre Zustimmung zur Datenanalyse unter der Bedingung, dass die Vertraulichkeit der Daten gewahrt bleibt. Die Dritten im System sind Forscher. Dies können Einzelpersonen oder Unternehmen sein, die autorisiert sind, bestimmte Arten von Analysen der Daten durchzuführen. Sie sind möglicherweise nicht autorisiert, die Daten im Klartext anzuzeigen.

3.1. Beschreibung des vorgeschlagenen Systems

Das System verschlüsselt die Daten mithilfe eines homomorphen Verschlüsselungsschemas, bevor sie den Dateneigentümer verlassen. Die verschlüsselten Daten werden in einem Smart Contract auf der Blockchain ges Die Analysen, die an den Daten durchgeführt werden können, werden ebenfalls im Smart Contract gespeichert. Wenn ein Forscher eine Analyse der Daten durchführen muss, wird die Analyse – vorausgesetzt, dass die Art der Analyse im Smart Contract verfügbar ist – an den verschlüsselten Daten mithilfe homomorpher Berechnungen durchgeführt. Die Ergebnisse der homomorphen Berechnungen werden ebenfalls verschlüsselt. Um das Ergebnis auf seiner Seite zu erhalten, fordert der Forscher an, dass das Ergebnis mit dem Schlüssel des Forschers erneut verschlüsselt wird, damit es sicher an den Forscher übertragen und nur vom Forscher entschlüsselt werden kann. Dabei ist zu beachten, dass die Klartextdaten während des gesamten Prozesses, auch während der erneuten Verschlüsselung, niemals offengelegt werden. Wenn der Forscher seinen öffentlichen Schlüssel zur erneuten Verschlüsselung an den Smart Contract sendet, initiiert der Smart Contract ein verteiltes Entschlüsselungsprotokoll, durch das Teile des Ergebnisses von verschiedenen Validierern entschlüsselt, mit dem öffentlichen Schlüssel des Forschers erneut verschlüsselt und dann kombiniert werden, um das endgültige erneut verschlüsselte Ergebnis zu erhalten **[21]. Abbildung 1 zeigt eine Übersicht über das System mit einem Dateneigentümer und einem Forscher als Hauptakteuren.**

Machine Translated by Google

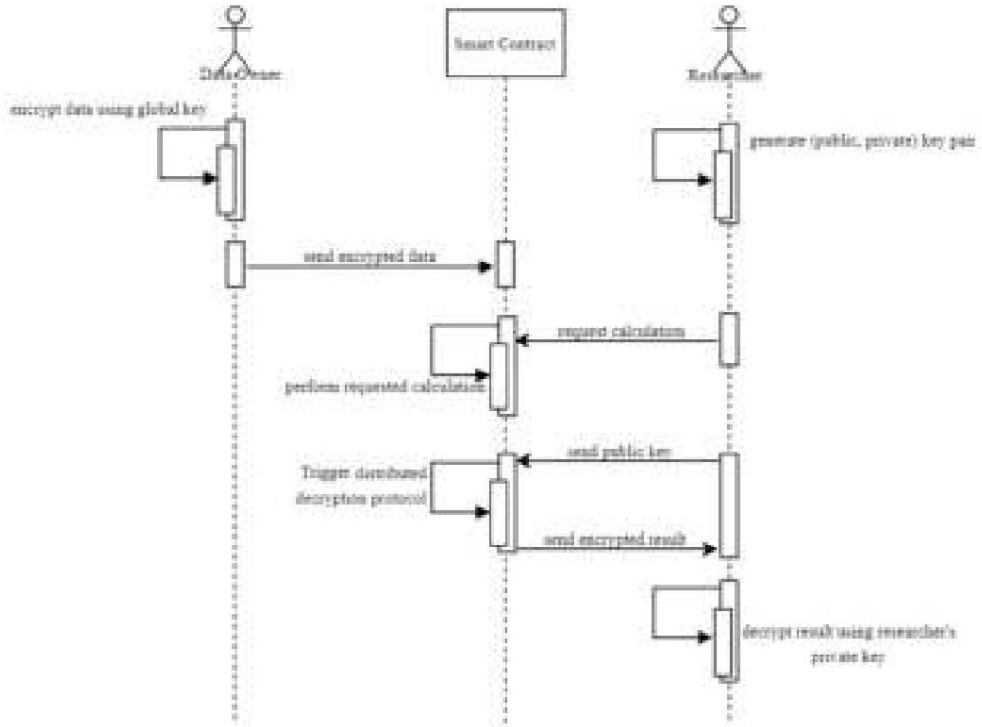


Abbildung 1. Eine Übersicht über das vorgeschlagene System.

3.2. Implementierung

Das vorgeschlagene System wird auf der Zama Blockchain implementiert, einer auf der Ethereum Virtual Machine (EVM) basierenden Blockchain, die Berechnungen mit verschlüsselten Werten unte

Wie Ethereum, das eine Währung namens Ether hat, hat Zama seine eigene Währung namens ZAMA. Die Grundidee hinter Zama besteht darin, vertrauliche Smart Contracts bereitzustellen. Dies bedeutet, dass die an den Smart Contract gesendeten oder von ihm empfangenen Daten verschlüsselt sind und nicht gelesen werden können, wenn die Datenübertragung abgefangen wird. Zama verwendet asymmetrische Verschlüsselung und daher zwei Schlüssel: einen öffentlichen und einen privaten. Der öffentliche Schlüssel, auch als globaler Schlüssel bekannt, wird öffentlich in der Kette gespeichert und von jedem Benutzer verwendet, um seine Daten zu verschlüsseln und Berechnungen mit diesen verschlüsselten Daten durchzuführen. Der globale Schlüssel wird während einer Einrichtungsphase von den ersten Validierern generiert und sicher erneut freigegeben, wenn der Validierersatz geändert wird [21]. Dies ermöglicht das Mischen verschlüsselter Daten von mehreren Benutzern und über mehrere Smart Contracts hinweg. Der private Schlüssel wird zum Entschlüsseln der Daten verwendet und gehört keinem einzelnen Benutzer. Stattdessen verwendet Zama ein Schwellenwertprotokoll. In einem Schwellenwertprotokoll werden Teile des privaten Schlüssels unter den Validiererknoten im Netzwerk verteilt, und eine bestimmte Anzahl von Validierern muss zusammenarbeiten, um die Daten zu entschlüsseln. Im Falle von Zama ist die Teilnahme von mindestens einem Drittel der Prüfer notwendig, um die Entschlüsselung durchzuführen **[18]. Diese Methode erhöht die Sicherheit durch Dezentralisierung und verhindert die missbräuchliche Verwendung des Schlüssels. Zama verwendet ein probabilistisches Verschlüsselungsschema, was bedeutet, dass der verschlüsselte Wert für einen einzelnen Klartext nicht immer derselbe ist. Betrachten Sie beispielsweise ein Array mit 4 Elementen [2, 4, 1, 4]. Sobald die Elemente dieses Arrays verschlüsselt sind, würde es hypothetisch zu [d5048, e2314, x3215, o9849] werden. Beachten Sie, dass die Zahl 4 im Klartext-Array zweimal an den Index 1 und 3 vorkommt, die verschlüsselte Version jedoch keine Wiederholungen aufweist und die Werte an den Index 1 und 3 unterschiedlich sind. Dieses Schema erhöht die Sicherheit weiter, indem es Angreifer daran hindert, anhand von Mustern im Geheimtext** Informationen über den **tt-** Klartext abzuleiten.

ff

tt

Das vorgeschlagene System verwendet die von Zama bereitgestellten Bibliotheken für homomorphe Berechnungen . Sobald der Vertrag kompiliert ist, erzeugt er die Application Binary Interface (A Der ABI gibt die im Smart Contract verfügbaren Funktionen zusammen mit ihren Parametern und Rückgabetypen an, sodass Benutzer mit dem Smart Contract interagieren können. Sowohl der Dateneigentümer als auch der Forscher haben ihre eigenen JavaScript-Anwendungen, die den ABI und die Me

Machine Translated by Google

wird von der Ethers-Bibliothek bereitgestellt, um mit dem Smart Contract zu kommunizieren. Der Dateneigentümer sendet die Daten an den Smart Contract und der Forscher fordert eine Analyse vom Smart Contract an. Abbildung 2 zeigt die Interaktion der Akteure mit dem Smart Contract.

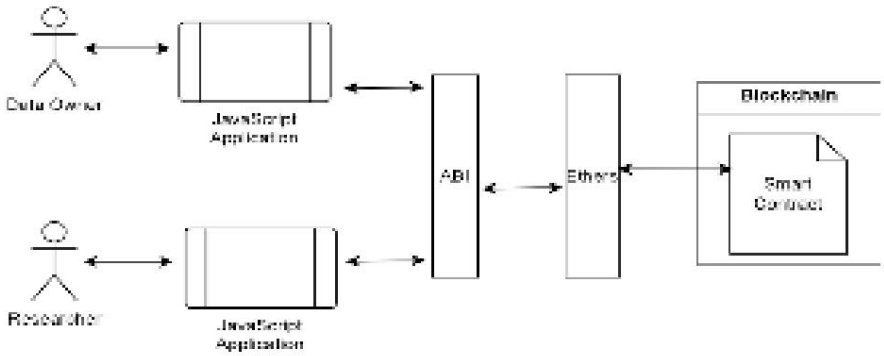


Abbildung 2. Interaktion mit dem Smart Contract.

Nachfolgend finden Sie eine Liste der Tools und Technologien, die zur Implementierung des Systems verwendet wurden:

- **Solidity: Eine Programmiersprache, die zum Schreiben von Smart Contracts auf verschiedenen Plattformen, einschließlich Ethereum.**
- **Node.js: Dies ist eine JavaScript-Laufzeitumgebung, die die Ausführung von JavaScript- Code außerhalb des Webbrowsers ermöglicht. Für dieses System wurde Version 18.18.0 verwendet.**
- **Remix: Dies ist eine browserbasierte integrierte Entwicklungsumgebung (IDE), die für die Entwicklung und Einsatz von Smart Contracts.**
- **MetaMask: Dies ist eine Kryptowährungs-Wallet, die zur Verbindung mit und Interaktion mit die Blockchain.**
- **Visual Studio Code: Es ist eine IDE, die die Entwicklung in verschiedenen Programmiersprachen unterstützt. Sprachen, einschließlich JavaScript.**
- **Ethers.js: Hierbei handelt es sich um eine JavaScript-Bibliothek, die die Interaktion mit Smart Contracts ermöglicht. Für dieses System wurde die Version 6.10.0 verwendet.**
- **FhEVM: Dies ist eine von Zama bereitgestellte Bibliothek, die die Erstellung vertraulicher Smart Contracts auf EVM mithilfe von Solidity ermöglicht. Für dieses System wurde Version 0.4.0 verwendet.**
- **FhEVMjs: Dies ist eine von Zama bereitgestellte JavaScript-Bibliothek, die die Interaktion mit Smart Contracts ermöglicht, einschließlich der Verschlüsselung der Daten und der Generierun Für dieses System wurde die Version 0.4.0 verwendet.**

Im Rahmen dieser Studie wird ein Proof of Concept für das vorgeschlagene System implementiert. Aufgrund von Einschränkungen in den Bibliotheken für homomorphe Berechnungen auf Zama und der Arbeit an einer öffentlichen Blockchain, wie in Abschnitt 3.2.1 näher erläutert wird, umfasste die Implementierung Berechnungen einfacher beschreibender Statistiken wie Mittelwert, Median und Varianz. Der Proof of Concept wurde mit einer kleinen Anzahl von Datenpunkten im Zahlenbereich zwischen 1 und 50 getestet.

3.2.1. Herausforderungen

Bei der Implementierung des Systems wurden die folgenden Einschränkungen und Herausforderungen in Solidity und Zama festgestellt: • Datentypen:

- **fhEVM unterstützt nur vorzeichenlose Ganzzahlen, die entweder 8-Bit, 16-Bit, 32-Bit oder 64-Bit sind. Darüber hinaus geben die von fhEVM bereitgestellten Methoden für Berechnungen mit verschlüsselten Daten wie Addition, Subtraktion, Multiplikation und Division eine verschlüsselte vorzeichenlose Ganzzahl zurück. Im Falle einer Division gibt es eine zusätzliche Einschränkung, da eine verschlüsselte Ganzzahl nur durch eine Ganzzahl im Klartext geteilt werden kann. Dies bedeutet, dass es derzeit nicht möglich ist, zwei verschlüsselte Ganzzahlen zu dividieren. Selbst wenn eine solche Division möglich wäre , wäre das Ergebnis aufgrund des Rückgabetyps immer eine verschlüsselte Ganzzahl ; daher ist es nicht möglich, mit Dezimalzahlen zu arbeiten. Darüber hinaus sind die Ganzzahlen**

fh fh

Machine Translated by Google

sind vorzeichenlos; daher ist es auch nicht möglich, negative Ganzzahlen zu verarbeiten. Diese Einschränkungen begrenzen die Anzahl der Analysen, die auf dem Smart Contract durchgeführt werden können. • Gaslimit: Berechnungen auf der Blockchain laufen auf Gas, einer Kosteneinheit, die mit der Durchführung einer Transaktion oder Berechnung im Netzwerk verbunden ist. Gas wird verwendet, um Validierer für die Ressourcen zu bezahlen, die sie zur Durchführung von Transaktionen verwenden. Das Gaslimit bezieht sich auf die maximale Gasmenge, die für eine Transaktion oder Berechnung ausgegeben werden kann. Das Festlegen eines Limits für den Gasverbrauch ist wichtig, da es sicherstellt, dass Transaktionen eine vordefinierte Obergrenze für den Ressourcenverbrauch haben, um vor Szenarien zu schützen, in denen die Codeausführung in eine Endlosschleife geraten könnte. Derzeit gibt es auf Zama devnet ein Gaslimit von 10.000.000 [22]. Homomorphe Berechnungen sind jedoch recht ressourcenintensiv, und da diese Berechnungen auf dem Smart Contract durchgeführt werden, verbrauchen sie viel Gas. Der Gasverbrauch hängt von den Operationen ab, die auf den verschlüsselten Daten durchgeführt werden, und steigt mit der Komplexität des Algorithmus und den Bits des Datentyps (32-Bi Das aktuelle Gaslimit schränkt die Funktionalität hinsichtlich der Anzahl der verarbeitbaren Datenpunkte ein.

- **Solidity: Im Gegensatz zu traditionellen Programmiersprachen wie Python und Java weist Solidity verschiedene Einschränkungen auf. Eine davon ist, dass die Datenstruktur, in der Schlüsselpaarwerte (Mappings) gespeichert werden, in ihrer Funktionalität ziemlich eingeschränkt ist und es nicht möglich ist, über Schlüssel oder Werte zu iterieren. Dies erschwert die Implementierung verschiedener Algorithmen, wie beispielsweise die Berechnung des Mod**

3.2.2. Berechnung des Mittelwertes

Für eine Wertemenge {x1, x2, ..., xn} ist **das arithmetische Mittel, bezeichnet mit x, die Summe der Werte geteilt durch die Anzahl der Werte, wie in Gleichung (1) gezeigt.**

$$\overline{x} = \frac{\sum_{i=1}^{N} x_i}{N} \quad (1)$$

Im entworfenen System wird im Smart Contract nur der Zähler, also die Summe der verschlüsselten Werte, berechnet. Die Division wird auf der Seite des Forschers durchgeführt. Dies liegt an der in Abschnitt 3.2.1 erwähnten Einschränkung , **dass eine Division immer eine Ganzzahl zurückgibt und daher an Genauigkeit im Ergebnis verliert. Zusammenfassend sendet der Smart Contract sowohl die verschlüsselte Summe als auch die verschlüsselte Anzahl der Elemente an den Forscher Diese beiden Werte werden dann vom Forscher entschlüsselt und eine abschließende Division durchgeführt, um das Ergebnis zu erhalten.**

3.2.3. Berechnung des Medians

Die Berechnung des Medians sortierter Daten ist eine einfache Aufgabe. Angenommen, die Daten bestehen aus n Werten, dann ist der Median das mittlere Element, wenn n ungerade ist, und der Durchschnitt der beiden mittleren Elemente, wenn n gerade ist. Die Formel für den Median eines sortierten Datensatzes mit n Datenpunkten ist in Gleichung (2) angegeben.

$$\text{Med} = \begin{cases} \left(\frac{n+1}{2}\right)^{\text{th}} \text{ Element wenn n ungerade ist} \\ \frac{\left(\frac{n}{2}\right) \text{ das Element} + \left(\frac{n}{2}+1\right) \text{ das Element}}{2} \text{ wenn n gerade ist} \end{cases} \quad (2)$$

Im System wird die Differenzierung bei der Berechnung des Medians einfach mithilfe einer if-Anweisung gehandhabt. Wenn n ungerade ist, wird der Median auf dem Smart Contract berechnet und verschlüsselt an den Forscher zurückgegeben. Wenn n gerade ist, wird die Summe der beiden mittleren Elemente auf dem Smart Contract berechnet und an den Forscher gesendet. Das Ergebnis wird dann vom Forscher entschlüsselt und durch 2 geteilt, um den Median zu erhalten. Somit trifft der Forscher die Entscheidung, die Division durchzuführen, basierend auf der Anzahl der Elemente.

Die obige Medianberechnung geht davon aus, dass die Daten sortiert sind. Die größte Herausforderung bei der Berechnung des Medians unsortierter Daten ist jedoch die Implementierung eines Sortieralgorithmus auf dem Smart Contract, da das Sortieren verschlüsselter Daten keine triviale Aufgal

Machine Translated by Google

Das vorgeschlagene System verwendet den Bubblesort-Algorithmus zum Sortieren der Daten. Dieser Algorithmus wurde aufgrund seiner Einfachheit gewählt. Er funktioniert, indem er zwei benachbarte Werte in einem Array vergleicht und die Zahlen basierend auf dem Ergebnis dieses Vergleichs vertauscht, was die Implementierung vereinfacht. Darüber hinaus erfordert die Implementierung von Bubblesort keine zusätzliche Datenstruktur. Daher eignet er sich zum Schreiben in Solidity, wenn man die Einschränkungen der Programmiersprache sowie der fhEVM-Bibliothek berücksichtigt.

3.2.4. Berechnung der Varianz

Die Formel zur Berechnung der Varianz für eine Reihe von Werten {x1, x2, . . . , xn} ist in **Gleichung (3) angegeben ,** wobei x **den Mittelwert der Daten darstellt.**

$$V = \frac{\ddot{y}_{ich=1}^{N} (xi \ddot{y} \bar{x})^2}{n \ddot{y} 1} \quad (3)$$

Wie in Abschnitt 3.2.2 **erläutert , ist es nicht möglich, den genauen Mittelwert einer beliebigen Sequenz im Smart Contract zu berechnen, da die Divisionsfunktion in der fhEVM-Bibliothek immer eine verschlüsselte Ganzzahl zurückgibt. Um die Berechnung des Mittelwerts in der Varianzformel zu eliminieren, haben wir eine äquivalente Formel abgeleitet, die weder einen Mittelwert noch einen anderen Divisionsoperator als eine Hauptdivisionsoperation im letzten Schritt enthält. Die schrittweise Herleitung der Formel beginnend mit Gleichung (3) ist unten dargestellt.**

$$V = \frac{\ddot{y}_{ich=1}^{N} (xi \ddot{y} \bar{x})^2}{n \ddot{y} 1}$$

$$= \frac{\ddot{y}_{ich=1}^{N} (xi \ddot{y} \frac{\ddot{y}_{j=1}^{N} xj}{N})^2}{n \ddot{y} 1}$$

$$= \frac{\ddot{y}_{ich=1}^{N} (xi^2 \ddot{y} \frac{2xi \ddot{y}_{j=1}^{N} xj}{N} + (\frac{\ddot{y}_{j=1}^{N} xj}{N})^2)}{n \ddot{y} 1}$$

$$= \frac{\frac{2n}{2n} \ddot{y}_{ich=1}^{N} (xi^2 \ddot{y} \frac{2xi \ddot{y}_{j=1}^{N} xj}{N} + \frac{(\ddot{y}_{j=1}^{N} xj)^2}{2n})}{n \ddot{y} 1}$$

$$= \frac{\ddot{y}_{ich=1}^{N} (n xi^2 \ddot{y} 2n xi \frac{\ddot{y}_{j=1}^{N} xj}{N} + \frac{(\ddot{y}_{j=1}^{N} xj)^2}{2n})}{n(n \ddot{y} 1)}$$

$$= \frac{n \ddot{y}_{i=1}^{N} xi^2 \ddot{y} 2n \ddot{y}_{i=1}^{N} (xi \ddot{y} \frac{1}{n} \ddot{y}_{j=1}^{N} xj) + \ddot{y}_{ich=1}^{N} (\ddot{y}_{j=1}^{N} xj)^2}{n(n \ddot{y} 1)}$$

$$= \frac{n \ddot{y}_{i=1}^{N} xi^2 \ddot{y} 2n \ddot{y}_{i=1}^{N} (xi \ddot{y} \frac{1}{n} \ddot{y}_{j=1}^{N} xj) + n(\ddot{y}_{i=1}^{N} xj)^2}{n(n \ddot{y} 1)} \quad (4)$$

Dieser Ansatz entfernte die Berechnung des Mittelwerts aus dem Algorithmus und damit auch die Division. Wie in Gleichung (4) gezeigt, war es jedoch nicht möglich, die Division vollständig zu eliminieren. Unter Berücksichtigung dieser Einschränkungen wird im Smart Contract nur die Berechnung des Zählers implementiert. Das verschlüsselte Ergebnis des Zählers wird zusammen mit dem verschlüsselten Wert von n an den Forscher gesendet, wo die endgültige Division durchgeführt wird. Bei näherer Betrachtung von Gleichung (4) kann festgestellt werden, dass der Zähler selbst drei Hauptkomponenten enthält. Darüber hinaus beinhaltet der Zähler auch eine Subtraktion. Wie in Abschnitt 3.2.1 erläutert, liefert Zama nur vorzeichenlose Ganzzahlen, was bedeutet, dass jedes Ergebnis unter Null die Genauigkeit des Ergebnisses beeinträchtigen würde. Um das Problem zu mildern, werden diese drei Komponenten separat berechnet und ihre Ergebnisse in separaten Variablen gespeichert . Dann wird die erste Komponente zur dritten addiert und die zweite Komponente wird dann vom Ergebnis der Addition abgezogen. Dadurch wird sichergestellt, dass das Ergebnis nicht unter Null fällt.

Machine Translated by Google

3.3. Sicherheitsgarantien für das vorgeschlagene System

Das vorgeschlagene System erfüllt die Sicherheitsanforderungen für alle Grundsätze der Informationssicherheit. Es bietet Sicherheitsgarantien für die Übertragung, Speicherung und Verwendung von Daten . Als Nächstes beschreiben wir, wie Vertraulichkeit, Integrität und Verfügbarkeit im vorgeschlagenen System erreicht werden.

• Vertraulichkeit: Daten werden vor der Weitergabe verschlüsselt und bleiben von da an verschlüsselt. Daten werden während der Übertragung, im Ruhezustand und bei der Verwendung verschlüsselt, was statistische Privatsphäre sowie statistische Vertraulichkeit gewährleistet. Die verschlüsselten Daten werden an den Smart Contract übertragen, wo sie gespeichert werden, sodass Berechnungen mit ihnen durchgeführt werden können. Das Ergebnis dieser Berechnungen wird ebenfalls verschlüsselt. Bevor das Ergebnis mit dem Forscher geteilt wird, wird es erneut verschlüsselt, nachdem es intern auf der Blockchain mithilfe des verteilten Entschlüsselungsprotokolls entschlüsselt wurde. Dieser Vorgang ist auf den Blöcken der Kette nicht sichtbar und daher für die Benutzer nicht sichtbar. Um das System zu betrügen und den Entschlüsselungsschlüssel zu erhalten, müsste ein Drittel der Knoten auf der Blockchain zusammenarbeiten, was bei Blockchains ein unwahrscheinliches Szenario ist. • Integrität der Daten: Jeder Block in einer Blockchain enthält den Hash des vorherigen Blocks, wodurch eine Kette entsteht. Das Manipulieren eines Blocks ändert den Hash dieses Blocks und aller Blöcke, die darauf folgen. Da das Erstellen der Kette mit den aktualisierten Blöcken eine fast unmögliche Aufgabe ist, sind die Daten auf der Blockchain manipulationssicher. Dies gewährleistet die Integrität der Daten. • Integrität der Berechnungen: Die Algorithmen für die Berechnungen, die für die Daten zulässig sind, sind im Smart Contract implementiert. Da Smart Contracts Teil der Blockchain sind, ist eine unbefugte Änderung nicht möglich. Dies gewährleistet Integrität der Berechnungen und damit Vertrauen in die für die Berechnungen implementierten Algorithmen. • Autorisierung: Die Autorisierung im vorgeschlagenen System kann aus zwei Perspektiven betrachtet werden. Eine ist die Kontrolle darüber, welche Art von Analyse an den Daten durchgeführt werden kann. Da die Funktionen für die Analyse im vorgeschlagenen System im Smart Contract verfügbar sind und nur autorisierte Benutzer/Einheiten die Funktionalität des Smart Contracts ändern können , ist ein Forscher auf die Verwendung der im Smart Contract verfügbaren Funktionen beschränkt. Vorausgesetzt, im Smart Contract sind nur genehmigte Funktionen implementiert, kann ein Forscher keine nicht genehmigte Analyse der Daten durchführen. Die andere Autorisierungsüberlegung im System betrifft, wer Daten an den Smart Contract senden und wer eine Analyse der Daten anfordern kann. Blockchain-Anwendungen können so programmiert werden, dass sie den Zugriff auf die im Smart Contract bereitgestellten Funktionen kontrollieren. Jeder Benutzer wird durch seinen öffentlichen Schlüssel eindeutig identifiziert und über diesen Schlüssel kann sein Zugriff auf die **Funktionen des Smart Contracts bestimmt werden .**

• Verfügbarkeit: Verfügbarkeit ist eine Funktion, die die Blockchain aufgrund ihrer dezentralen Natur selbst bietet . Jeder Knoten im Netzwerk hat seine eigene Kopie des Ledgers, sodass das gesamte System auch dann noch zugänglich ist, wenn ein Knoten Dadurch wird ein einzelner Ausfallpunkt beseitigt und Fehlertoleranz in das System eingeführt.

4. Ergebnisse

Wir haben die Proof-of-Concept-Implementierung des Systems durchgeführt, bei der der Smart Contract die Funktion hat, den Mittelwert, den Median und die Varianz des Datensatzes zu berechnen. Wir haben die Zeit aufgezeichnet, die zum Ausführen verschiedener Operationen für die Datenpunkte 4, 8, 10, 16, 24 und 32 erforderlich ist. Diese Operationen umfassen das Festlegen der Daten (d. h. das Senden der Daten an den Smart Contract), die Mittelwertberechnung, das Sortieren der Daten (was zum Ermitteln des Medians erforderlich ist), die Medianberechnung und die Varianzberechnung. Alle Experimente wurden mit einem 16-Bit-Datentyp durchgeführt. Jeder Datensatz wurde zufällig mit Punkten Jeder erstellte Datensatz wurde sofort an den Smart Contract gesendet. Jede Datensatzgröße wurde achtmal getestet, um die Zuverlässigkeit der Ergebnisse sicherzustellen. Die Tests wurden zu unterschiedlichen Zeiten und an unterschiedlichen geografischen Standorten mit unterschiedlicher Konnektivität durchgeführt. Tabelle 1 zeigt den durchschnittlichen Zeitpunkt der acht Versuche für jede Datensatzgröße.

Machine Translated by Google

Tabelle 1. Dauer der Vorgänge von der Anforderung bis zum Ergebnis in Millisekunden.

Anzahl der Datenpunkte (n)	Versendung Daten	Bedeuten Berechnung	Sortieren von Daten	Median Berechnung	Varianz Berechnung
4	8088.66679	7512.111693	9063.634214	6312.926094	18.109,45269
8	6642,155596	8792,629253	22.330,79229	7679.880587	28.161,12535
10	7243.960528	10,486.1769	34.412,11508	7358.729393	
16	7111.459872	11,948.44701			
24	9990.13181	16.596,62563			
32	13.358,13821	21.479,25259			

Wie aus Tabelle 1 hervorgeht , waren die Größen einiger Datensätze für bestimmte Betrieb aufgrund der von Zama auferlegten Gasgrenze. Während beispielsweise der Sortierprozess wurde an Datensätzen mit 4, 8 und 10 Datenpunkten durchgeführt, der Varianzprozess war nur auf Datensätze mit 4 und 8 Datenpunkten anwendbar. Das Senden der Datensätze an die Blockchain und die Mittelwertbildung könnten über alle Datensätze durchgeführt werden.

Abbildung 3 vergleicht die Zeitabläufe für jeden Vorgang basierend auf der Anzahl der Datenpunkte. Außerdem werden die Zeitabläufe für verschiedene Operationen verglichen. Insgesamt, da die Anzahl der Datenpunkte erhöht sich auch die für eine Operation benötigte Zeit. Einige Ausnahmen hiervon sind Senden von Daten für eine kleinere Anzahl von Datenpunkten und die Medianberechnung. Wir schreiben Dies liegt daran, dass Operationen, die eine kurze Zeit in Anspruch nehmen, anfälliger sind durch das Transaktionsvolumen und die Qualität des Netzwerks beeinflusst werden Konnektivität. Wie in der Grafik dargestellt, benötigen Varianz und Sortierung mehr Zeit als das Einstellen Daten, Mittelwertberechnung und Medianberechnung. Die Medianberechnung dauert am wenigsten Zeit da es sich im Vergleich zu den anderen Operationen um wenige einfache Operationen handelt, unabhängig von die Datengröße.

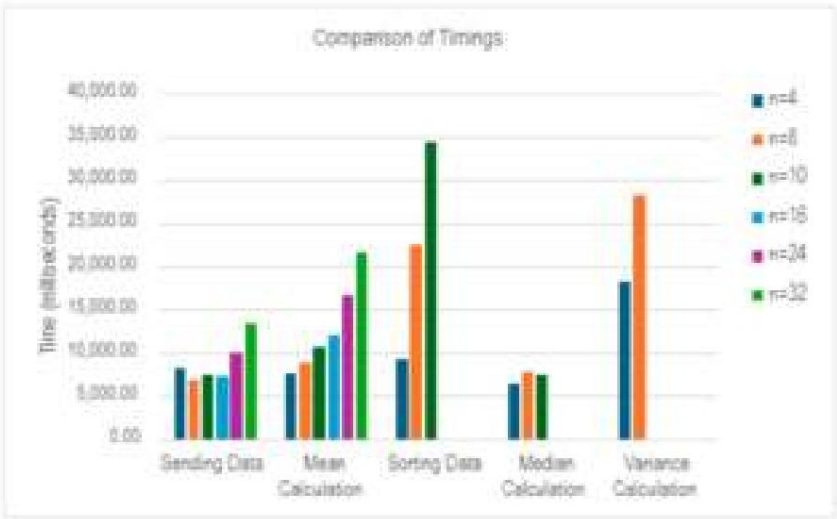


Abbildung 3. Vergleich der zeitlichen Abläufe bei Operationen in Bezug auf die Anzahl der Datenpunkte.

Wir haben auch den Gasverbrauch jedes Vorgangs für jeden Test aufgezeichnet und überprüft, ob die Datensortierung und Varianzberechnung erreichten das von Zama auferlegte Gaslimit bei 10 Datenpunkten und 8 Datenpunkte. Tabelle 2 zeigt die Durchschnittswerte der acht Versuche für jeden Datenpunkt. Größe festlegen.

Machine Translated by Google

Tabelle 2. Gasverbrauch im Betrieb.

Anzahl der Datenpunkte (n)	Versendung Daten	Bedeuten Berechnung	Sortierung Daten	Median Berechnung	Varianz Berechnung
4	290.491,5	577.618	1.367.854	169.517	5.022.035
8	565.458,875	1.124.706	6.007.343	169.517	8.213.723
10	709.171,5	1.398.250	9.565.005	169.517	
16	1.166.246,625	2.218.883			
24	1.835.907,625	3.313.060			
32	2.574.607,375	4.407.237			

Es ist kein Wunder, dass der Gasverbrauch für die angegebene Anzahl von Datenpunkten konstant war für den gleichen Vorgang. Abbildung 4 zeigt einen Vergleich des Gasverbrauchs im Verhältnis zur Anzahl von Datenpunkten für jeden Vorgang sowie ein Vergleich des Gasverbrauchs in Bezug auf die verschiedenen Operationen.

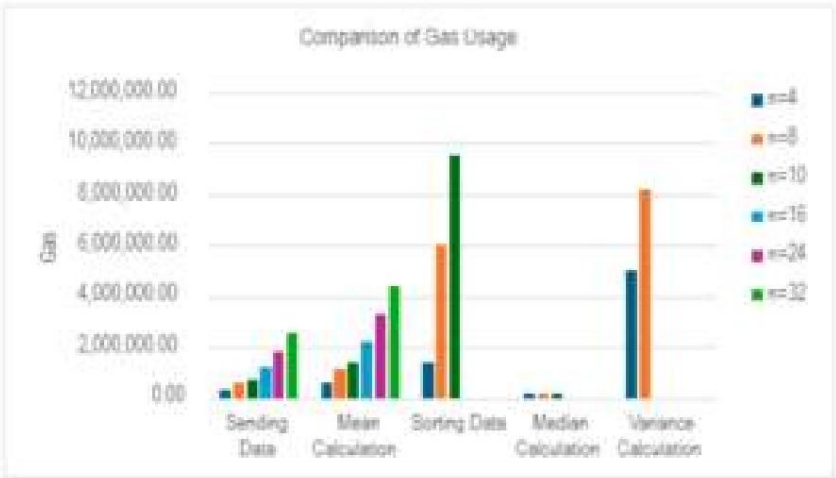


Abbildung 4. Vergleich des Gasverbrauchs der Betriebe hinsichtlich der Anzahl der Datenpunkte.

Wie Abbildung 4 zeigt, stieg der Gasverbrauch mit der Anzahl der Datenpunkte, mit Ausnahme der Median. Die Berechnung des Medians umfasste eine einfache Operation, nämlich die Ermittlung des Index des mittleres Element, das nur eine Addition und eine Division erfordert, unabhängig von der Anzahl von Datenpunkten. Daher war der Gasverbrauch für alle drei Datensatzgrößen gleich. Die Berechnung der Median erfordert eine Sortierung der Daten. Da die Sortierung der Daten jedoch zu viel Gas verbraucht, kann nicht für mehr als 10 Datenpunkte berechnet werden. Ebenso ist eine Varianzberechnung erforderlich rund 10 Millionen Gas für acht Datenpunkte. Das Hinzufügen eines weiteren Datenpunktes würde erfordern mehr als 10 Millionen Gas, was über dem von Zama auferlegten Gaslimit liegen würde. Wie aus der Grafik und der Beschreibung der Durchführung der Operationen ersichtlich in Abschnitt 3: Die Operationen, die mehr Berechnungen erfordern, benötigen mehr Gas.

5. Diskussion

Diese Studie schlug ein System vor, das Datensicherheit, statistische Privatsphäre und Vertraulichkeit unter Verwendung von Techniken und Werkzeugen ausschließlich aus der Mathematik und Informatik. Das System verwendet eine vollständig homomorphe Verschlüsselung über eine Blockchain und gewährleistet die End-to- End-Sicherheit der Daten vom Zeitpunkt, an dem sie den Dateneigentümer verlassen, bis zum Zeitpunkt, an dem der Fo empfängt die Berechnungsergebnisse. Die Daten sind während der Übertragung, im Ruhezustand sowie in Verwendung. Darüber hinaus bietet das vorgeschlagene System auch Rechenintegrität. Mit allen Durch den gebotenen Schutz ist das System eine ultimative Lösung für die Einhaltung der Datensicherheit Anforderungen.

Machine Translated by Google

Eine Proof-of-Concept-Implementierung für das vorgeschlagene System wird auf der öffentlichen Blockchain Zama durchgeführt . Bei der Implementierung kann ein Forscher sichere Berechnungen von Mittelwert, Median und Varianz anfordern. Die zeitlichen Ergebnisse für die Durchführung der erforderlichen Operationen an verschiedenen Datenpunktzahlen werden zusammen mit den Einschränkungen bei der Arbeit mit homomorphen Verschlüsselungsbibliotheken und einer öffentlichen Blockchain präsentiert. Obwohl diese Studie erfolgreich die Möglichkeit demonstriert, Daten und Berechnungen auf einer Blockchain mit vollständig homomorpher Verschlüsselung zu sichern, gibt es bestimmte Herausforderungen und Bereiche, die verbessert werden müssen.

Vollständig homomorphe Operationen sind in Bezug auf Ressourcen recht teuer, und es gibt Beschränkungen hinsichtlich der Anzahl der Berechnungen, die auf der Blockchain durchgeführt werden können, bevor die Gas- und Kostengrenzen überschritten werden. Dies ist besonders kritisch, wenn die Anwendung im Hauptnetzwerk ausgeführt wird, wo für Transaktionen Gebühren anfallen. Diese Einschränkung kann in einem privaten oder genehmigungspflichtigen Blockchain-Netzwerk teilweise gemildert werden. Eine solche Blockchain bietet ähnliche Garantien wie die öffentliche Blockchain und bietet mehr Kontrolle darüber, wer die Anwendung verwenden kann.

Zukünftige Studien in diesem Bereich können komplexere Algorithmen aus verschiedenen Disziplinen, von Statistik bis hin zu maschinellem Lernen, einbeziehen und ein breiteres Datenspektrum abdecken, da mehr Ressourcen verfügbar werden und die Technologie voranschreitet. Darüber hinaus müssen traditionelle Algorithmen, die für Klartextdaten entwickelt wurden, wie etwa Suchen und Sortieren, sorgfältig überarbeitet werden, um homomorphe Berechnungen zu unterstützen, abhängig von den Fähigkeiten der verfügbaren Bibliotheken. Eine weitere mögliche Richtung für zukünftige Studien könnte die Entwicklung eines Frameworks sein, das vollständig homomorphe Berechnungen über diesen Smart Contract unterstützt und ein breiteres Spektrum an Funktionen bietet und gleichzeitig die Einschränkungen des aktuellen Systems überwindet, wie etwa die Division zweier verschlüsselter Zahlen und die Einbeziehung von Dezimal- oder Gleitkommazahlen.

Beiträge der Autoren: Konzeptualisierung, YKP; Datenkuratierung, RR, YKP und ZDM; Formale Analyse, YKP und ZDM; Untersuchung, RR, YKP und ZDM; Methodik, RR und YKP; Projektverwaltung, YKP; Ressourcen, RR, YKP und ZDM; Software, RR und ZDM; Überwachung , YKP; Validierung, RR, YKP und ZDM; Visualisierung, YKP und ZDM; Schreiben – Originalentwurf , RR und YKP; Schreiben – Überprüfung und Bearbeitung, RR, YKP und ZDM. Alle Autoren haben **die veröffentlichte Version des Manuskripts gelesen** und stimmen ihr zu.

Finanzierung: Diese Forschung erhielt keine externe Finanzierung.

Datenverfügbarkeitserklärung: Die in dieser Studie vorgestellten Originalbeiträge sind im Artikel enthalten. Weitere Anfragen können an den/die entsprechenden Autor(en) gerichtet werden.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Verweise

1. **Dwork, C.; Smith, A.; Steinke, T.; Ullman, J. Aufgedeckt! Eine Untersuchung über Angriffe auf private Daten.** Annu. Rev. Stat. Appl. **2017, 4, 61–84.** [Querverweis]
2. **Kitamura, K.; Irvan, M.; Yamaguchi, RS Anonymitätstestangriffe und Schwachstellenindikatoren für die Offenlegung von „Patientenmerkmalen“ in medizinischen Artikeln. In Proceedings des 2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), Genua, Italien, 6.–10. Juni 2022; S. 186–193.** [CrossRef]
3. **Nissim, K. Datenschutz: Von der Datenbankrekonstruktion zu Rechtstheoremen. In Proceedings des 40. ACM SIGMOD-SIGACT-SIGAI- Symposiums zu Prinzipien von Datenbanksystemen, in PODS'21, virtuelle Veranstaltung, China, 20.–25. Juni 2021; Association for Computing Machinery: New York, NY, USA, 2021; S. 33–41.** [CrossRef]
4. **Wang, Y.-R.; Tsai, Y.-C. Der Schutz der Datenfreigabe zum Schutz der Privatsphäre in Financial Vision.** Appl. Sci. **2022,** 12, **7408.** [CrossRef]
5. **Agarwal, S. Sray und Mishra Daten- und Modelldatenschutz. In** „Verantwortungsvolle KI: Implementierung ethischer und unvoreingenommener Algorithmen“; **Springer International Publishing: Cham, Schweiz, 2021; S. 153–170.** [CrossRef]
6. **Sokolovska, A.; Kocarev, L. Integration technischer und rechtlicher Konzepte des Datenschutzes.** IEEE Access **2018, 6, 26543–26557.** [CrossRef] 7. **Kifer, D.; Lin, B.-R. Eine axiomatische Sicht auf statistische Privatsphäre und Nutzen.** J. Priv. Confidentiality **2012, 4.** [CrossRef]
8. **Slavkovi´c, A.; Seeman, J. Statistischer Datenschutz: Ein Lied von Datenschutz und Nutzen.** Annu. Rev. Stat. Appl. **2023,** 10, **189–218.** [Querverweis]
9. **C.3. Statistische Vertraulichkeit – MSITS 2010 Compilers Guide – UN Statistics Wiki. Online verfügbar: https://unstats.un.org/ wiki/display/M2CG/C.3.+ +Statistische+Vertraulichkeit (abgerufen am 4. Juni 2024).**

Machine Translated by Google

10. **UNSD – Grundprinzipien nationaler amtlicher Statistiken. Online verfügbar: https://unstats.un.org/fpos/ (abgerufen am 5. Juni 2024).**
11. **Statistische Vertraulichkeit und Schutz personenbezogener Daten – Eurostat. Online verfügbar: https://ec.europa.eu/eurostat/web/ Mikrodaten/statistische Geheimhaltung und Schutz personenbezogener Daten (abgerufen am 18. Juni 2024).**
12. **Nationale Akademien der Wissenschaften, Ingenieurwissenschaften und Medizin; Abteilung für Verhaltens- und Sozialwissenschaften und Bildung; Ausschuss für nationale Statistiken; Gremium zur Verbesserung der Bundesstatistik für Politik- und Sozialforschung unter Verwendung mehrerer Datenquellen und modernster Schätzmethoden; Groves, RM; Harris-Kojetin, BA (Hrsg.)** Innovationen in der Bundesstatistik: Kombinieren von Datenquellen bei gleichzeitigem Schutz der Privatsphäre; **The National Academies Press: Washington, DC, USA, 2017. [CrossRef]**
13. **Kumar, R.; Kumar, J.; Khan, AA; Ali, H.; Bernard, CM; Khan, RU; Zeng, S. Blockchain und homomorphe Verschlüsselung basierend auf datenschutzerhaltender Modellaggregation für medizinische Bilder.** Comput. Med. Imaging Graph. **2022,** 102, **102139. [Cross**
14. **Liang, W.; Zhang, D.; Lei, X.; Tang, M.; Li, K.-C.; Zomaya, A. Circuit Copyright Blockchain: Blockchain-basierte homomorphe Verschlüsselung zum Schutz von IP-Schaltungen.** IEEE Trans. Emerg. Top. Comput. **2020, 9, 1410–1420. [CrossRef]**
15. **Yaji, S.; Bangera, K.; Neelima, B. Datenschutz in Blockchain basierend auf einem partiell homomorphen Verschlüsselungssystem für KI-Anwendungen. In Proceedings der 2018 IEEE 25th International Conference on High Performance Computing Workshops (HiPCW), Bengaluru, Indien, 17.–20. Dezember 2018; S. 81–85. [CrossRef]**
16. **Mutlu, ZD; Peker, YK; Aydÿn, A. Selçuk Blockchain-basierte, Privatsphäre bewahrende lineare Regression.** J. Millimeterwave Com-mun. Optim. Model. **2023, 3, 45–49. Online verfügbar: https://www.jomcom.org/index.php/1/article/view/81 (abgerufen am 5. Juni 2024).**
17. **Vanin, FN; Policarpo, LM; Righi, RD; Heck, SM; da Silva, VF; Goldim, J.; da Costa, CA Ein Blockchain-basiertes End-to-End-Datenschutzmodell für den Austausch persönlicher Gesundheitsakten: Ein vollständig homomorpher Verschlüsselungsansatz.** Sensors **[Querverweis] [PubMed]**
18. **Umar, B.; Olaniyi, O.; Olajide, D.; Dogo, E. Paillier Kryptosystembasierter ChainNode für sichere elektronische Abstimmungen.** Vorderseite. Blockchain **2022, 5, 927013. [CrossRef]**
19. **Shrestha, R.; Kim, S. Integration von IoT mit Blockchain und homomorpher Verschlüsselung: Herausfordernde Probleme und Chancen.** Adv. Comput. **2019,** 115, **293–331. [CrossRef]**
20. **Caldarola, F.; d'Atri, G.; Zanardo, E. Neural Fairness Blockchain Protocol Using an Elliptic Curves Lottery.** Mathematik **2022,** 10, **3040. [CrossRef]**
21. **fhevm/fhevm-whitepaper.pdf unter main·zama-ai/fhevm·GitHub. Online verfügbar: https://github.com/zama-ai/fhevm/blob/ main/fhevm-whitepaper.pdf (abgerufen am 4. Juni 2024).**
22. **Schätzung Gas|0.3|fhEVM. Online verfügbar: https://docs.zama.ai/fhevm/v/0.3-2/how-to/gas (abgerufen am 4. Juni 2024).**

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Modellanalyse der Kanalanordnung in CRNs basierend auf Prioritätsplanungsstrategie mit reservierter Warteschlange

Qianyu Xu [1], Suoping Li 2,*, Jaafar Gaber 3 und Yuzhou Han [2]

[1] Fakultät für Elektro- und Elektroniktechnik, Xi'an Aeronautical University, Xi'an 710077, China; 202202001@xaau.edu.cn

[2] Fakultät für Elektro- und Informationstechnik, Technische Universität Lanzhou, Lanzhou 730050, China; 19119376252@163.com

[3] Fakultät für Informatik und Computertechnik, Universite de Technologie Belfort-Montbeliard, 90010 Belfort, Frankreich; gaber@utbm.fr

* Korrespondenz: lsuop@lut.edu.cn

Zusammenfassung: In kognitiven Funknetzwerken können Sekundärnutzer (SUs) durch Kanalzusammenstellung ihre Netzwerkkapazität erweitern und die Frequenznutzung verbessern. Planungsstrategien, die ausschließlich auf heterogener Dienstklassifizierung basieren, können die Bereitstellungspriorität wichtiger elastischer Dienste in speziellen Szenarien wie der Notfallrettung nicht garantieren. Daher wird in dieser Arbeit eine Prioritätsplanungsstrategie mit reservierter Warteschlange (Ps-rq) vorgeschlagen. Ein statischer Faktor wird definiert, um SUs basierend auf dem Nachrichtentyp in elastische Dienste und Echtzeitdienste zu klassifizieren, während ein dynamischer Faktor definiert wird, um elastische Dienste mit hoher Priorität basierend auf Informationsgültigkeit, Nachrichtenkorrelation und Nachrichtengröße zu unterscheiden. Die Benutzer mit hoher Priorität in den unterbrochenen elastischen Diensten werden in die reservierte Warteschlange gestellt, um ihre Dienste sicher
Es wird eine kontinuierliche Markov-Kettenanalyse durchgeführt und alle möglichen Übergangszustände, Auslöseereignisse, Übergangsraten und Übergangsbedingungen des Systems ausgehend von einem allgemeinen Zustand abgeleitet. Darüber hinaus werden Bewertungsindizes der Systemleistung ermittelt. Fallstudien und Simulationsergebnisse belegen, dass die vorgeschlagene Strategie die Netzwerkkapazität erhöhen, die Wahrscheinlichkeit von Blockierungen und Zwangsabbrüchen für sekundäre Benutzer verringern und insbesondere die Leistung von elastischen Diensten mit hoher Priorität verbessern kann. Darüber hinaus analysieren wir die Eigenschaften von Ps-rq durch einen umfassenden Vergleich mit vier anderen Schemata. Das Experiment beweist, dass die Ps-rq-Strategie die Servicequalität der wichtigen elastischen Dienste auf der Grundlage einer fairen Planung effektiv verbessern

Schlüsselwörter: kognitive Funknetze; Kanalzusammenstellung; Prioritätsplanungsstrategie; reservierte Warteschlange; Markow-Kette; Spektrumnutzung; Übergangsbedingungen

check for updates

Zitat: Xu, Q.; Li, S.; Gaber, J.; Han, Y. Modellanalyse der Kanalzusammenstellung in CRNs basierend auf einer Prioritätsplanungsstrategie mit Reservierte Warteschlange. Elektronik **2024,** 13, **3051. https://doi.org/10.3390/Elektronik13153051**

Wissenschaftliche Herausgeber: Duc Tai Le und Syed Muhammad Raza

Empfangen: 11. Juni 2024
Überarbeitet: 29. Juli 2024
Akzeptiert: 30. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz. Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Einleitung

Herkömmliche drahtlose Kommunikationsnetze basieren hauptsächlich auf einer statischen Frequenzzuweisung. Der lizenzierte Benutzer belegt jedoch immer ein festes Frequenzband, was zu einer übermäßigen Verschwendung von Spektrum führt [1]. Kognitive Funknetze (CRNs) ermöglichen es Sekundärbenutzern (SUs), das ungenutzte Spektrum von Primärbenutzern (PUs) durch Spektrumserfassung zu erkunden und zu nutzen, um die Spektrumsnutzung zu verbessern [2,3].

Zur Verbesserung der Frequenznutzung sind kognitive Funksysteme erforderlich, die verfügbare Frequenzressourcen flexibel planen und zuweisen [4,5]. Dynamischer Frequenzzugriff (DSA) kann sicherstellen, dass SUs Zugriff auf Frequenzlücken in der lizenzierten Frequenzbandbreite erhalten [6]. Die Flexibilität von DSA spiegelt sich darin wider, dass die laufenden SUs den Kanal verlassen können, sobald sie die Signale der PUs erkennen. Anschließend kann die Frequenzanpassung durch Frequenzübergabe oder dynamische Kanalauswahl durchgeführt werden. Die Entwicklung von Ressourcenplanungsregeln auf der Grundlage von DSA, die der Anwendungsumgebung entsprechen, kann zur Ver

Machine Translated by Google

In CRNs kann die Einführung von Kanalassemblierungstechnologie und Spektrumanpassungstechnologie im DSA-Prozess die Serviceeffizienz von SUs erheblich verbessern.
Mithilfe **der Kanalzusammenstellungstechnologie können SUs einen ungenutzten Kanal an mehrere Kanäle binden oder aggregieren, um die Kapazität des sekundären Netzwerks zu erhöhen. In Kombination mit der Spektrumanpassungstechnologie können die laufenden SUs die Anzahl der zusammengebauten Kanäle entsprechend den Aktivitäten anderer PUs und SUs adaptiv anpassen, um eine flexible Planung zu ermöglichen. Die Modellierung des DSA-Prozesses basierend auf der Ressourcenzuweisung ist jedoch komplex, sodass die Leistung sekundärer Netzwerke analysiert und bewertet w**

Die Serviceleistung von SUs ist untrennbar mit der Methode zur Ressourcenplanung und -zuweisung verbunden. Der Servicefluss bei der drahtlosen Kommunikation kann basierend auf der Verzögerungsempfindlichkeit in Echtzeitdienste (wie Audio und Video) und elastische Dienste (wie Daten) unterteilt werden . **Echtzeitdienste haben normalerweise eine höhere Priorität als elastische Dienste, während elastische Dienste eine geringere Verzögerungsempfindlichkeit aufweisen und ihre Übertragungsrate entsprechend der Netzwerkumgebung adaptiv angepasst werden kann.**

Forscher haben verschiedene Methoden zur Ressourcenzuweisung für heterogenen Verkehr vorgeschlagen, um die Frequenznutzung zu verbessern [8–14]. Die wichtigste Technik besteht darin, die verfügbaren Frequenzressourcen entsprechend der dynamischen Netzwerkumgebung in einem Frequenzpool zusammenzufassen **und sie dann bei Bedarf zuzuweisen. Unter diesen untersuchten Referenzen [10–13] das Ressourcenplanungsproblem von mehrkanaligen kognitiven Funknetzwerken mit heterogenem Verkehr basierend auf der Kanalzusammenstellung und verwendeten mathematische Methoden, um das Analysemodell zur Bewertung der Leistung sekundärer Netzwerke zu erstellen. Frühere Forschungen [11] haben gezeigt, dass die Anordnung einer variablen Anzahl zusammengestellter Kanäle für elastische Dienste und einer festen Anzahl zusammengestellter Kanäle für Echtzeitdienste entsprechend der Betriebsumgebung zur Verbesserung der Frequenzeffizienz beitragen kann.**

In [15] schlugen die Autoren vor, dass die Einführung des CTMC-Prozesses (Continuous-Time Markov Chain) eine Modellierungsmethode für die Leistungsanalyse von CRNs mit Kanalzusammenstellung bieten kann, die den Vorteil hat, dass das Problem der Zeitsynchronisierung von PUs und SUs auf demselben Kanal vermieden wird. In [11] schlugen die Autoren eine Kanalzusammenstellungsstrategie vor, die Spektrumanpassung und heterogene Dienste berücksichtigt, und CTMC wurde verwendet, um die Leistung des sekundären Netzwerks zu bewerten. Allerdings kann die durch die PU unterbrochene SU den Dienst nach Verlassen des Kanals nicht fortsetzen, was offensichtlich die Möglichkeit erhöht, dass die SU getrennt wird. Daher wurde in Ref. [12] vorgeschlagen, Warteschlangen für heterogene Dienste einzurichten und einen Warteschlangenplanungsalgorithmus einzuführen, um die Netzwerkkapazität auf Basis von Priorisierung zu verbessern. Frühere Forschungen [12,16] haben gezeigt, dass das Einrichten von Warteschlangen für SUs die Netzwerkkapazität verbessern und die Blockierungswahrscheinlichkeit von SUs reduziere Unter Berücksichtigung einer Vielzahl von Faktoren, die die Kanalkapazität beeinflussen, einschließlich einer zeitabhängigen drahtlosen Verbindung, eines begrenzten Puffers, Verkehrsklassen und eines adaptiven Modulations- und Kodierungsschemas (AMC) , wurde in **[17] eine Kanalzusammenstellungsstrategie mit Warteschlangen vorgeschlagen und die Interaktion zwischen den Systemparametern durch die Erstellung eines CTMC Basierend auf der Kanalzusammenstellung wurde in [18]** ein adaptiver Algorithmus zur Spektrumsvermietung vorgeschlagen **und die DSA-Leistung von SU-Verkehr mit Zugriffspriorität durch die Entwicklung des CTMC-Modells diskutiert. Die oben genannte Forschung zeigt, dass die Einführung der Kanalzusammenstellungstechnologie für heterogene Dienste die Nutzung des ungenutzten Spektrums verbessern kann. Darüber hinaus kann die Einführung von CTMC dazu beitragen, die Planungsdynamik von Spektrumressourcen unter Benutzern im Problem der Ressourcenzuweisung zu modellieren. Die oben genannte Forschung stellt eine strengere und verfeinerte Methodik zur mathematischen Analyse der Leistung des sekundären Sy**

In CRNs mit Diensten mehrerer Typen kann die Ressourcenzuweisung auf Grundlage der Dienstklassifizierung die Übertragungszuverlässigkeit wichtiger Informationen verbessern. Die Verbesserungen führen jedoch normalerweise zum Starvation-Problem von Diensten mit niedriger Priorität, da der Dienst über einen langen Zeitraum nicht verfügbar ist . In **[19–21]** wurden **Ressourcenplanungsschemata vorgeschlagen, die das Starvation-Problem von Benutzern mit niedriger Priorität** lindern können, **und die Leistung des sekundären Systems wurde durch die Erstellung von Markov-Analysemodellen mit Warteschlangen bewertet. In [20]** wurde eine **Kanalbindungstechnologie mit Starvation-Minderung eingeführt und eine**

Machine Translated by Google

Das M/G/1-Warteschlangenmodell wurde verwendet, um den Spektrum-Übergabeprozess von SUs darzustellen und so die Ressourcennutzung zu verbessern. In [21] wurde ein Round-Robin-Prioritäts-Planungsalgorithmus (RRP) vorgeschlagen, der statische und dynamische Ressourcenzuweisung kombiniert. Dieser Algorithmus erhält die Übertragungsmöglichkeiten von Benutzern mit niedriger Priorität, indem er die Warteschlange mit hoher Priorität vorübergehend aussetzt. Ziel ist es, die Aushungerung von Diensten mit niedriger Priorität in CRNs zu minimieren. Tatsächlich verwendet der RRP-Algorithmus in [21] eine Warteschlangenabfrageplanung, aber wenn der Verkehr einer Warteschlange groß ist, wird eine Benutzerplanung, die nur auf Warteschlangenabfrage basiert, die Systemleistung erheblich reduzieren [22]. Die Planung heterogener Dienste, die auf der Aufteilung der Diensttypen basiert, kann tatsächlich ein bequemeres Prioritätsplanungsschema für die Kanalzuweisung und Prioritätsplanung in CRNs während der Systeminitialisierung bieten . Die bisherige Forschung hat jedoch die Übertragungsanforderungen von Benutzern mit hoher Priorität und hoher Flexibilität in praktischen Anwendungen nicht berücksichtigt und konnte daher keine grundlegende L

Dies ist folglich der Grund für unsere Untersuchung. In diesem Artikel wird die Ressourcenplanung und -zuweisung mit Serviceklassifizierung für CRNs untersucht, die in speziellen Szenarien eine Rangfolge der Dienste erfordern. Einige wichtige elastische Dienste müssen in speziellen CRN-Anwendungsszenarien eine höhere Priorität erhalten, darunter Notfallrettung, Drohnen [23] und UAVs. Angesichts der Tatsache, dass Planungsschemata mit hierarchischen Diensten aufgrund langer Zeiträume ohne Dienste häufig zum Problem der Aushungerung von Diensten mit niedriger Priorität führen, besteht das Ziel dieser Studie darin, die Bereitstellungseffizienz wichtiger Informationen durch die Bereitstellung eines möglichst fairen Verteilungsr

Im Gegensatz zu unseren früheren Studien [9] befasst sich dieser Beitrag mit einem allgemeineren Planungsschema, das auf heterogenen Diensten basiert. Ziel dieser Studie ist es, ein umfassenderes Ressourcenplanungs- und -zuweisungsschema für CRNs vorzuschlagen, das die QoS von elastischen Diensten mit hoher Priorität garantieren kann. Genauer gesagt wird basierend auf der Polling-Planung eine neue reservierte Warteschlange für die elastischen Dienste mit hoher Priorität bereitgestellt, die von PUs unterbrochen werden, und den SUs in der reservierten Warteschlange wir Basierend auf der Technologie der Kanalzusammenstellung und Spektrumanpassung wird der dynamische Spektrumzugriffsprozess heterogener Dienste analysiert und die Spektrumressourcenzuweisung zwischen SUs durch die Entwicklung eines CTMC untersucht. Die wichtigsten Beiträge

dieser Studie sind wie folgt: 1. Diese Studie schlägt eine Prioritätsplanungsstrategie mit reservierter Warteschlange (Ps-rq-Strategie) vor. Das Systemmodell und das Dienstklassifizierungsschema der vorgeschlagenen Strategie werden vorgestellt. Durch die Kombination der Technologie der Kanalzusammenstellung und Spektrumanpassung wird ein proportionales Prioritätszuweisungssc

2. Der Ressourcenflussprozess des sekundären Netzwerks wird auf einem CTMC abgebildet. Durch den Einsatz stochastischer Analysemethoden werden die komplexen Benutzeraktivitäten im CRN einzeln mit Zustandsübergängen verknüpft. Das Übergangsereignis, die Übergangsrate und die Übergangsbedingung, die Änderungen der Systemzustände auslösen können, werden bereitgestellt und bieten so eine theoretische Grundlage für die Verbesserung der Leistung sekundärer Systeme durch Kanalzusammenstellung.
3. Das Verfahren zur Dimensionsreduktion hochdimensionaler Markow-Ketten wird angegeben. Darauf aufbauend werden die analytischen Ergebnisse der sekundären Netzwerkleistung untersucht und numerisch simuliert. Wir präsentieren einen Vergleich des Ps-rq mit vier früheren Schemata und beweisen durch numerische Verifikation, dass die vorgeschlagene Strategie die Servicequalität für unterbrochene wichtige elastische Dienste bei gleichzeitiger Beibehaltung einer fairen Planung gewährleistet.

2. Prioritätsplanungsstrategie mit reservierter Warteschlange (Ps-rq-Strategie)

Dieser Abschnitt bietet ein heterogenes Serviceklassifizierungsschema für CRNs mit Multi-Type-Services. Auf dieser Grundlage wird eine Prioritätsplanungsstrategie mit reservierter Warteschlange (Ps-rq) vorgeschlagen. Ziel ist es, eine faire Planung für CRNs mit Multi-Type-Services bereitzustellen, die die Lieferzuverlässigkeit wichtiger Informationen garantieren und gleichzeitig das Hungerproblem von Warteschlangen mit niedriger Priorität lindern kann.

Machine Translated by Google

2.1. Systemmodell und Annahmen

Zentralisierte kognitive Funknetze (CRN) bestehen aus zwei Funktypen: Primärbenutzer (PUs) und Sekundärbenutzer (SUs). Die Benutzerplanung wird von der entsprechenden Basisstation gesteuert. Der zentrale Knoten ist für die Interaktion mit der Umgebung sowie für die Verwaltung der Frequenzzuweisung und Planungsentscheidungen für andere SUs über einen Steuerkanal verantwortlich. Der zentrale Knoten in zentralisierten CRN kann die Systemleistung verbessern, indem er effiziente Entscheidungen trifft und Interferenzen zwischen Benutzern reduziert. Im Gegensatz zu verteilten CRNs entfällt bei zentralisierten CRNs die Notwendigkeit weiterer Überlegungen zu Aspekten wie Interferenzen zwischen Benutzern und Knotenegoismus.

Wir gehen davon aus, dass sich M (M ÿ Z ÿ **) PUs die Spektrumbandbreite teilen und jede PU nur auf einen Kanal zugreifen kann. Im Vergleich zu SUs haben PUs absolute Kanalzugriffspriorität . SUs sind mit Mitteln zur spektralen Energieerkennung ausgestattet, vorausgesetzt, dass SUs über perfekte Erfassungsfähigkeiten verfügen. Und die erforderliche Erfassungszeit ist so kurz, dass PUs den Kanal innerhalb dieser Zeit nicht erreichen. Sobald die PU erfasst wird, gibt die laufende SU die zusammengestellten Kanäle sofort frei und beendet und führt gleichzeitig eine Spektrumanpassung durch . Unter der Annahme, dass weniger Zeit für die Spektrumanpassung aufgewendet wird, hat der Kanalwechsel keine Auswirkungen auf die Übertragung der SUs.**

2.2. Heterogene Serviceklassifizierung

Gestützt auf das Systemmodell schlägt dieser Abschnitt ein heterogenes Dienstklassifizierungsschema vor, um die Planungsanforderungen für CRNs mit multitypen Diensten wie Daten, Bildern, Audio und Video zu erfüllen. So kann beispielsweise eine verspätete Übermittlung oder ein Paketverlust lebenswichtiger Patienteninformationen bei einem medizinischen Notfall lebensbedrohlich sein, und dieser Teil der Informationen ist überwiegend ein elastischer Dienst. Eine frühere Studie [12] auf Basis der **heterogenen Dienstklassifizierung berücksichtigte jedoch wichtige elastische Dienste mit hoher Übermittlungspriorität unter besonderen Szenarien nicht. Die Übertragungspriorität und Zuverlässigkeit wichtiger elastischer Dienste können kaum gewährleistet werden. Auf Basis der heterogenen Dienstklassifizierung werden multitypen Dienste in CRNs auch nach dynamischen Fakt**

- **Statischer Faktor: Nachrichtentyp. SUs in CRNs werden je nach Verzögerungsempfindlichkeit** in elastische Dienste (SUe) **und Echtzeitdienste** (SUr) eingeteilt. Elastische Dienste zeichnen sich **durch kurze Framelänge, hohen Datenverkehr und hohe Verzögerungstoleranz aus. Echtzeitdienste erfordern die Aktualität der Datenübertragung. Je nach Verzögerungstoleranz wird die Dienstpriorität wie folgt definiert:**

Priorität: SUr > SUe ;

• Dynamischer Faktor: Elastische Dienste in CRNs werden basierend auf drei dynamischen Faktoren neu klassifiziert, darunter Informationsgültigkeit, Nachrichtenkorrelation und Nachrichtengrö

Die Gültigkeit der Informationen (Val) **bezieht sich auf die verbleibende Zeit bis zur Frist. Nachrichten mit kürzerer verbleibender Zeit haben eine höhere Priorität.**

$$\text{Wert} = \frac{\textbf{Die verbleibende Zeit bis zum Stichtag}}{\textbf{Die Lieferzeit von Informationen}}; \quad \textbf{(1)}$$

Die Nachrichtenkorrelation (Cor) **bezieht sich auf den Nachrichtentyp, der zuerst zugestellt werden muss.**

$$\text{Cor} = \begin{cases} \textbf{1,} & \textbf{Relevanz wichtiger Informationen} \\ \textbf{0,} & \text{Andere} \end{cases}; \quad \textbf{(2)}$$

Die Nachrichtengröße (Größe) **ist umgekehrt proportional zur Nachrichtenpriorität.**

$$\text{Größe} = \frac{\textbf{Die Nachrichtengröße der SU}}{\textbf{Durchschnittliche Nachrichtenlänge}}; \quad \textbf{(3)}$$

Machine Translated by Google

Nach der Kombination der drei oben genannten dynamischen Faktoren werden elastische Dienste in hohe, mittlere und niedrige Prioritäten eingeteilt. Ziel ist es, die elastischen Dienste mit hoher Priorität (SUh e) **zu identifizieren und sie dann gegenüber Echtzeitdiensten zu priorisieren. Die Dienstpriorität wird wie folgt vorgeschlagen:**

Priorität : SUh $_t$ > SUr > SUe

Darüber hinaus wird in **Abbildung** 1 **eine dreidimensionale (3D) Prioritätsansicht elastischer Dienste dargestellt. Zur Priorisierung elastischer Dienste wird die Takagi-Sugeno-Fuzzy-Steuerung verwendet [24]. Der Eingangswert des Fuzzy-Inferenzsystems reicht von 0 bis 1. Die Ausgabe des Systems ist die Priorität und wird in drei Fuzzy-Mengen aufgeteilt, deren Domänen [ÿ1, 1] sind. Entsprechend der Fitness wird die „Wenn-dann“-Regel übernommen, um 27 Fuzzy-Regeln festzulegen, und die Priorität elastischer Dienste wird in hohe, mittlere und niedrige Prioritäten aufgeteilt. Wenn beispielsweise** Val **= 0,** Cor **= 1 und** Size **= 1, dann** ist Priorität **= 1, was wichtige Informationen mit einer kurzen verbleibenden Zeit und einer kleinen Nachrichtengröße darstellt. Aus den Eingangs- und Ausgangsvariablen werden Gaußsche und trapezförmige Mitgliedschaftsfunktionen**

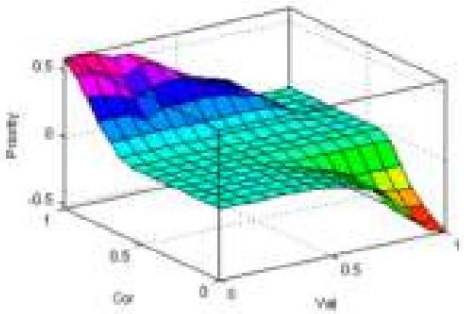


Abbildung 1. 3D-Prioritätsansicht elastischer Dienste.

2.3. Ps-rq-Strategie

Basierend auf dem oben genannten heterogenen Dienstklassifizierungsschema wird Verkehr mit Diensten unterschiedlicher Art, der in das System eingeht, durch die Prioritätsplanungsstrategie mit reservierter Warteschlange (Ps-rq-Strategie) bedient. Wie in Abbildung 2 dargestellt, besteht die Ps-rq-Strategie aus einer Informationsplanungseinheit und einer Datenübertragungseinheit. Benutzer, die in das sekundäre System (SUs) eintreten, werden durch einen Paketklassifizierer I **gemäß einem statischen** Faktor **in den elastischen Dienst** (SUe) **und den Echtzeitdienst** (SUr) unterteilt **(Paketklassifizierer werden zum Parsen eingehender Pakettypen verwendet.) Für** SUe **und** SUr werden unterschiedliche physische Warteschlangen eingerichtet, die mit Le und Lr bezeichnet werden. **Wenn der Dienst von SUs während der Übertragung durch eine PU unterbrochen wird , wird eine Spektrumanpassung durchgeführt. Da** SUr **eine hohe Aktualität erfordert und Jitter-Toleranz aufweist, wird bei einem Fehlschlagen der Spektrumanpassung der unterbrochene** SUr **gelöscht und die zusammengesetzten Kanäle freigegeben. E Das heißt,** Lr **hat keine Rückkopplungsschleife.**

Im Unterschied zu SUr hat SUe **geringere Anforderungen an die Aktualität und kann wichtige *L*** -Signale enthalten. Daher wird für Le **eine Rückkopplungsschleife eingerichtet . Die unterbrochenen** SUe **werden erneut in den Paketklassifizierer II eingegeben und anhand eines dynamischen Faktors in hohe, mittlere und niedrige Priorität klassifiziert. Der Zweck besteht darin, elastische Dienste mit hoher Priorität zu trennen, sie als** SUh $_e$, zu markieren **und in die reservierte Warteschlange** Le zu stellen. Währenddessen werden die unterbrochenen **elastischen Dienste mit mittlerer Priorität** (SUm e) **und niedriger Priorität** (SUl e) **von ihren Markierungen entfernt und an das Ende der Warteschlange** Le zurückgestellt **(Rückkopplungsschleife). Unter der Annahme, dass jede Warteschlange isomorph ist, warten die sekundären Benutzer in der Warteschlange darauf, entsprechend der Reihenfolge des „Wer zuerst kommt, mahlt zuerst“-Dienstes (FCFS) eingeplant zu werden. Die Priorität der Warteschlange wird auf** L > Lr > Le festgelegt . **SUs unterschiedlichen Typs können sich nur in der entsprechenden Warteschlange anstellen. Der Dienst der SUs wird blockiert, wenn in der Warteschlange** $_t$ **kein Platz mehr vorhanden ist. Wenn die Warteschlange voll ist und die unterbrochenen** SUe **und** SUh **in die entsprechen**

Machine Translated by Google

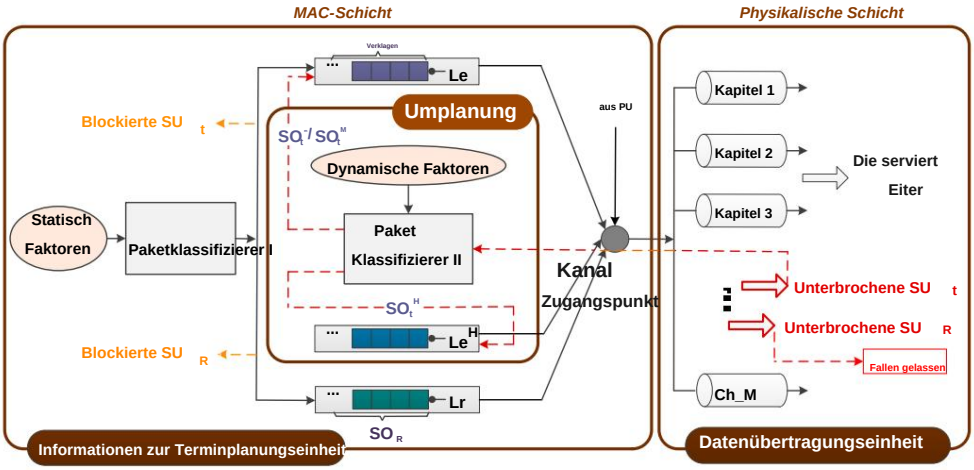


Abbildung 2. Ps-rq-Strategie.

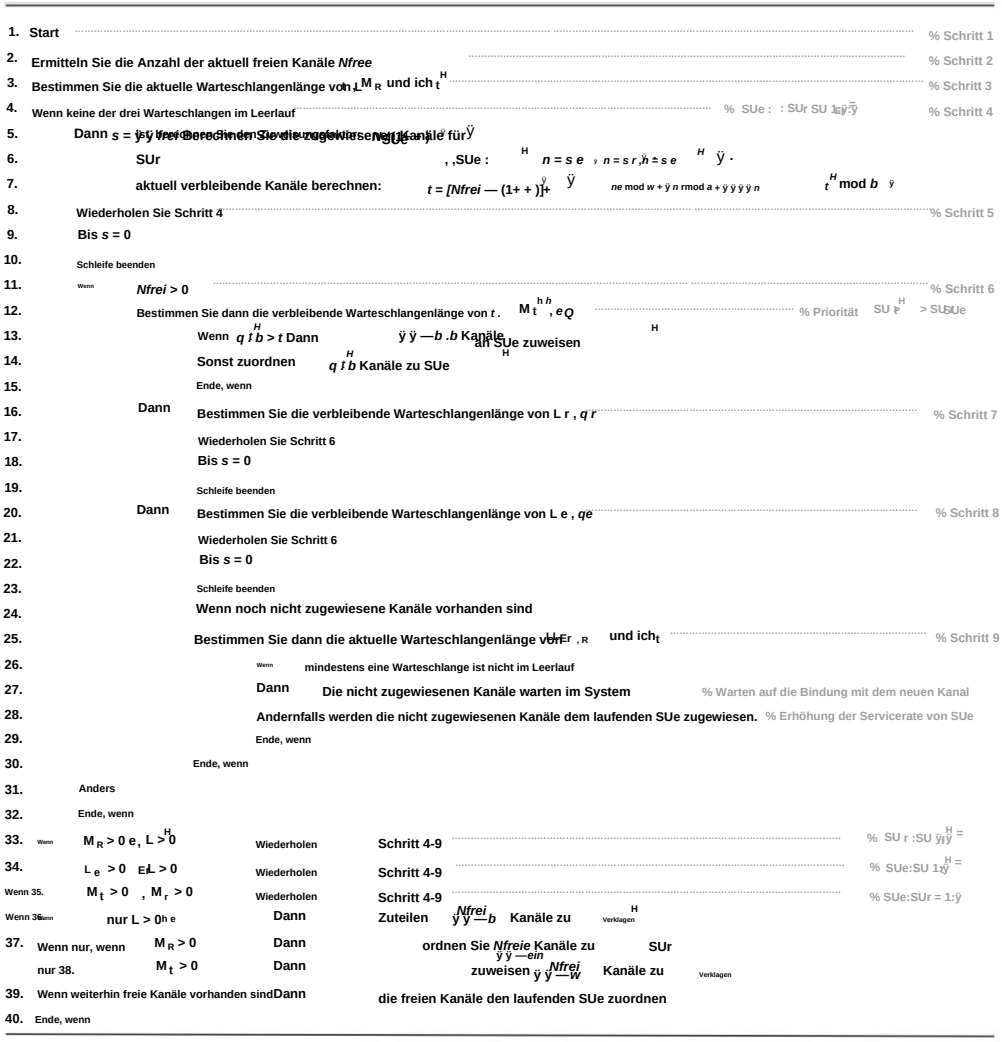
2.4. Kanalzuweisung basierend auf der Ps-rq-Strategie

Beim Planungsprozess wird eine Kanalassemblierungstechnologie angewendet, um die Servicerate der SUs durch Zusammenstellung eines oder mehrerer Kanäle zu SUs zu verbessern. Gleichzeitig wird eine Spektrumanpassungstechnologie verwendet, um eine adaptive Zuweisung der zusammengestellten Kanäle zu ermöglichen. Für CRNs mit mehrartigen Diensten läuft die Kanalzuweisung basierend auf der Ps-rq-Strategie wie folgt ab: Für SUr **und markierte** SUh **wird eine feste Anzahl von Kanälen zusammengestellt, um die Übertragungseffizienz sicherzustellen. Genauer gesagt kann der Dienst von** SUr **gestartet werden. Dementsprechend** kann SUh den Dienst starten, indem sie a Kanäle **aggregiert , wobei** a ÿ Z ÿ ist, indem **wobei** b ÿ Z ÿ **ist. Der Unterschied besteht darin, dass** SUe **die Anzahl der aggregierten Kanäle basierend auf der Spektrumanpassungstechnologie flexibel anpassen kann. Zu *. Konkret kann** der Dienst von SUe durch die Aggregation von w **Kanälen gestartet werden, und die Servicerate kann durch Erhöhen der aggregierten Kanäle bis zur Obergrenze v verbessert werden, wobei w,** v ÿ Z ÿ und w ÿ v. Um den Dienst von SUr **und der markierten** SUh **einfacher starten zu können, basiert die Parametereinstellung dieser Studie auf** a = b < **w. Angenommen, die Servicerate von SUs steigt linear an, nachdem die Technologie zur Kanalzusammenstellung übernommen wurde, dann steigt** die Servicerate **von** SUr **mit** a- Zusammenstellung von Kanälen um das a - **fache.**

Die Kanalzuweisung an die SUs erfolgt auf der Grundlage einer proportionalen Kanalzuweisung mit Priorität. Erkenne gemäß der Warteschlangenpriorität L > Lr > Le , **ob die drei Warteschlangen leer sind ,** wenn N freie Kanäle **vorhanden sind .** Wenn keine davon leer ist, wird der freie Kanal SUh e , SUr **und** SUe **in der Warteschlange im Verhältnis 1 :** *ÿ* : ÿ **zugewiesen , wobei ÿ,** *ÿ* **ÿ 1. Hier kann der Wert von ÿ,** *ÿ* **gemäß den Prioritätsanforderungen und der Verzögerungstoleranz von** SUh e , SUr **und** SUe bestimmt werden . **Dann werden die verbleibenden Kanäle gemäß der Priorität** SUh e zugewiesen, **gefolgt von** SUr **und** SUr .

Wenn nur zwei der Warteschlangen nicht leer sind, läuft der Zuweisungsprozess wie oben ab. Wenn nur eine nicht leer ist, werden dieser Warteschlange alle freien Kanäle z **Wenn es darüber hinaus noch freie Kanäle gibt, werden die verbleibenden Kanäle den laufenden** SUe zugewiesen . **Der obige Zuweisungsprozess wird in MATLAB realisiert und die proportionale Kanalzuweisung mit Priorität ist in Schema 1 dargestellt. Hier wird die Anzahl der freien Kanäle immer noch dadurch bestimmt, dass SUs Informationen über verfügbare freie Kanäle durch Technologie zur Erkennung der Spektrumenergie erhalten, um opportunistischen Zugriff zu erreichen.**

Machine Translated by Google



Schema 1. Proportionale Kanalzuweisung mit Priorität.

3. Dynamischer Kanalzugriffsprozess

In CRNs führt das Ankunfts- und Abgangsverhalten von PUs und SUs zu einer Änderung in den Systemzustand. Um den Ressourcenflussprozess der SUs in verschiedenen Systemzuständen ist der dynamische Kanalzugriffsprozess basierend auf der vorgeschlagenen Ps-rq-Strategie In diesem Abschnitt wird dies durch die Klassifizierung der Benutzeraktivitäten in CRNs in vier Ereignisse untersucht: PU Ankunft, PU Abfahrt, SU Ankunft und SU Abfahrt.

• Ereignis A: Ankunft der PU

SUs in CRNs werden in elastische Dienste und Echtzeitdienste eingeteilt nach Verzögerungsempfindlichkeit. Elastische Dienste haben die Eigenschaften einer kurzen Framelänge, Verkehrsfluss und hohe Verzögerungstoleranz. Echtzeitdienste erfordern die Aktualität der Daten Übertragung. Entsprechend der Verzögerungstoleranz wird die Dienstpriorität wie folgt definiert:

Wenn die PU eintrifft und ein freier Kanal im System vorhanden ist, wird dieser zugewiesen an die neu eingetroffene PU. In diesem Fall werden die Dienste anderer SUs nicht beeinträchtigt. Aber wenn Es gibt keine freien Kanäle im System. Die PU unterbricht jeden laufenden Kanal der SUs aufgrund der höheren Datenzugriffspriorität. Die unterbrochene SU verlässt den Kanal und führen Sie eine Spektrumanpassung durch. Das heißt, die laufende SUe mit der maximalen Anzahl von Kanälen wird den Kanal an die unterbrochene SU spenden, bis die gesammelten Kanäle von Kann die SUe den Dienst nicht unterstützen, dann wird die laufende SUe mit der zweiten Maximalzahl von Kanälen wird seinen Kanal spenden . . . bis alle SUe die erforderlichen Anzahl der Kanäle der unterbrochenen SU. Dann schlägt die Anpassung fehl, was bedeutet, dass alle

Machine Translated by Google

der laufende SUe **kann den Kanal nicht für den unterbrochenen SU spenden. Der** SUe, **der den**
Anpassung wird Paketklassifizierer II eingebunden und die Priorität durch einen Fuzzy-Regler markiert. Wenn
die SUe **gehört zu** SUh e , **es wird in die reservierte Warteschlange** L eingetragen H , **andernfalls kehrt es zurück zu**
das Ende der Warteschlange Le . **Wenn die Warteschlange voll ist, wird der Dienst der SU zwangsweise beendet.**

Wenn die PU den Kanal einer SUr **oder einer markierten** SUh **belegt** e , **Spektrumanpassung wird**
auf die gleiche Weise durchgeführt werden. Der Dienst wird fortgesetzt, wenn die Anpassung erfolgreich war;
andernfalls die unterbrochene SUr **oder** SUh t **wird gezwungen, zu beenden und freizugeben, die montiert**
Kanal.

Insbesondere die durch die zwangsweise beendete SUh freigegebenen Kanäle **a , oder** SUr **kann nicht**
der Einstellung a = b < **w nicht zum Starten des Dienstes der SUs in der Warteschlange verwendet wurden. Der**
Die von der zwangsweise beendeten SUe **freigegebenen Kanäle können für die wartende** SUh verwendet werden. t **und** SUr
in L e . **Die Kanalzuweisung der SUs in der Warteschlange folgt Schema 1. Wenn keine SU in der**
Regel des First-Come-First-Service (FCFS): Der Service kann nur gestartet werden, wenn y
freie Kanäle im System vorhanden sind. Wenn die in der Warteschlange befindlichen Kanäle den freien Kanal
nicht verwenden, werden **die minimalen aggregierten Kanäle bis zur** SUe **auf die Obergrenze v gesetzt. Wenn es**
können nur noch freie Kanäle, die restlichen Kanäle werden zur laufenden SUe zusammengeführt **mit**
wenn ÿ freie Kanäle im System vorhanden sind. Wenn beide in der Warteschlange nicht gebunden sind, v.
In
präsentieren ein Flussdiagramm, wie in Abbildung 3 dargestellt, um den dynamischen Kanalzugriff anzuzeigen
der Gültlage der Ps-rq-Strategie bei der Ankunft des PUs gestartet wird. Ebenso können andere Ereignisse, einschließlich des
Abfahrt, SU Ankunft und SU Abfahrt können auch das Flussdiagramm verwenden, um den Kanal zu beschreiben
der SUs eingerichtet.

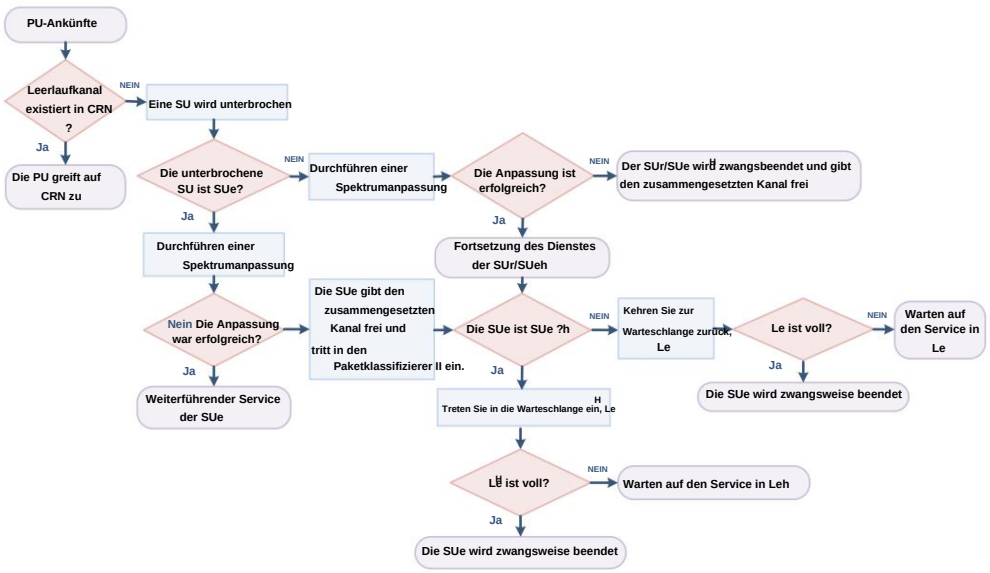


Abbildung 3. Der dynamische Kanalzugriffsprozess basierend auf der Ps-rq-Strategie bei PU-Ankunft.

Abbildung 3. Der dynamische Kanalzugriffsprozess basierend auf der Ps-rq-Strategie bei PU-Ankunft.

• Ereignis C: SU-Ankunft

Ereignis B: PU-Abgang

Es gibt drei Arten von SU-Ankünften im System, darunter die Ankunft von **. Der Abgang einer PU gibt**
für die in der Warteschlange wartenden SUs verwendet. Wenn die SUs in der Warteschlange den freien Kanal nicht verwenden,
Wenn die SU, die am System ankommt, ***b*** **freie Kanäle im System hat ,** kann der freie Kanal für die laufende SU
verwendet werden , deren zusammengesetzte Kanäle nicht erreicht haben.
dann der Service der Obergrenze mit der Kanalmontagetechnik durchgeführt werden.
kann gestartet werden. Ebenso kann der Dienst der
Insbesondere wird der freie Kanal bevorzugt SUh **zugewiesen** **in** L e , **die Regel befolgen**
neu eingetroffene Kanäle können gestartet werden, wenn die Anzahl der vorhandenen Kanäle in
von First Come First Service (FCFS). Der Service von SUh t **kann nur gestartet werden, wenn** b **ÿ 1 frei**
das System erreicht den Minimalwert *w* oder *a* . Wenn die freien Kanäle im System nicht vorhanden sind,
in der Warteschlange verwenden nicht den freien Kanal, den Kanal
unterstützen den Dienst, die laufenden werden Spektrum Anpassung durchführen, um die zu
spenden kann an SUr **in** Lr **montiert werden , und der Dienst von** SUr **kann nur gestartet werden, wenn** ein **ÿ 1 fr**
Kanal für die neu eingetroffenen SUs. Wenn alle laufenden SUs nicht in der Lage sind, die SUs in der
Kanäle im System vorhanden sind. Wenn beide SUh **Warteschlange bereitzustellen, verwenden** Sie **den freien Kanal nicht.**
Der Kanal kann zu SUe **in** Le **zusammengebaut werden . Ebenso kann der Dienst von** SUe **nur gestartet werden**
erforderliche Anzahl von Kanälen, wird die neu eingetroffene SU am Ende der entsprechenden
Warteschlange eingestuft. Wenn die Warteschlange voll ist, wird der Dienst der SU blockiert. Da jedoch
alle Kanäle voll sind,
in die Warteschlange kommen aus dem unterbrochenen ,
der Dienst von **wird zwangsweise beendet, das heißt, es gibt keine blockierte .**

• Ereignis D: SU-Abfahrt

Machine Translated by Google

wenn im System bereits w **ÿ** 1 freie Kanäle vorhanden sind. Wenn keine SU in der Warteschlange den freien **Kanal verwendet, wird der Kanal gemäß Regel 1 den aktiven SUs im System zugeordnet.**

• Ereignis C: SU Ankunft

Es gibt drei Arten von SU-Ankünften im System, darunter SUh e , SUe **und** SUr . **Wenn die SU, die im System ankommt,** SUh ist **und es** b **freie Kanäleim System gibt, kann der Dienst der neu angekommenen** SUh **gestartet werden. Ebenso kann der Dienst der neu angekommenen** SUe **oder** SUr **gestartet werden, wenn die Anzahl der im System vorhandenen Kanäle den Mindestwert** w **oder a erreicht. Wenn die freien Kanäle im System den Dienst nicht unterstützen können, führt die laufende** SUe **eine Spektrumanpassung durch, um den Kanal für die neu angekommenen SUs freizugeben. Wenn alle laufenden** SUe **nicht in der Lage sind, die erforderliche Anzahl von Kanälen bereitzustellen, wird die neu angekommene SU am Ende der entsprechenden Warteschlange eingestuft. Wenn die Warteschlange voll ist, wird der Dienst der SU blockiert. Da jedoch alle in die Warteschlange eingehenden** SUh e **voll sind, wird der Dienst der** SUh **gezwungen,**
L_e^H h **kommt von der unterbrochenen** SUe , **wenn** L t t **beenden, d.h. es gibt keine blockierte** SUh e .

• Ereignis D: SU-Abfahrt

Der Abgang von SU gibt alle aggregierten Kanäle frei. Das heißt, der Abgang von SUe **gibt** k (w **ÿ** k **ÿ** v) **Kanäle frei; der Abgang von** SUr **gibt** a **Kanäle frei; und der Abgang von** SUh
t gibt b Kanäle **frei . Schema 1 wird verwendet, um die von SUs freigegebenen Kanäle zuzuweisen, wobei die Summe der Anzahl der von SUs freigegebenen Kanäle und der Anzahl der freien** Nf- **Kanäle im System ist. Wenn nach der Zusammenstellung aller in der Warteschlange wartenden SUs noch freie Kanäle vorhanden sind , wird der freie Kanal gemäß Regel 1** d

4. CTMC-Analyse und QoS-Maßnahmen

Beim Kanalzugriff lösen Benutzeraktivitäten die Änderung des Systemzustandes aus. Kontinuierliche Markow-Ketten (CTMCs) werden verwendet, um eine entsprechende Beziehung zwischen dem möglichen Zustandsübergang des Systems und den auslösenden Ereignissen herzustellen, die die Änderung des Systemzustands verursachen, wodurch der Spektrumressourcenfluss zwischen SUs mehrerer Typen dargestellt werden kann. Daher kann die Leistungsanalyse des sekundären Systems mithilfe der CTMC durchgeführt werden. Die Aktivitäten der SUs können im mehrdimensionalen Wahrscheinlichkeitsraum abgebildet werden. Anschließend kann das mathematische Modell der vorgeschlagenen Ps-rq-Strategie erstellt werden, um die QoS-Metriken durch die stationäre Lösung zu erhalten, und die Leistung des sekundären Systems kann bewertet werden.

4.1. Planen dynamischer Analysen und CTMC-Modellierung **Beachten**

Sie, dass der allgemeine Zustand wie folgt aussieht:

$$x = lw, \ldots, lv, ga, gb, lpu, leq, lrq, lhq$$

wobei lm (w **ÿ** m **ÿ** v) **die Anzahl der** SUe **mit** m **aggregierten Kanälen,** ga **die Anzahl der** SUr **mit** a **aggregierten Kanälen,** gb **die Anzahl der** SUh **mit** b **aggregierten Kanälen,** lpu **die Anzahl der PUs im System und** leq, lrq **und** lhq **die aktuellen Warteschlangenlängen von** Le , Lr **bzw.** L e sind .
H

Weiterhin kann die Menge der möglichen Zustände des Systems wie folgt notiert werden:

$$S = lw, \ldots, lv, ga, gb, lpu, leq, lrq, lhq \ \ddot{y}\ \mathbf{0;}\ b(x)\ \ddot{y}\ \mathbf{M;}\ leq\ \ddot{y}\ Qeq;\ lrq\ \ddot{y}\ Qrq;\ lhq\ \ddot{y}\ Qhq; \quad \ddot{y}^{vjw}_{ich=1}\ ilw+i < \mathbf{w,}\ \text{wenn } leq > \mathbf{0} \tag{4}$$

Hier sind Qeq, Qrq **und** Qhq **die Gesamtkapazitäten von** Le , Lr **bzw.** L e . Wennleq > **0 bedeutet,** $\ddot{y}^{vjw}_{ich=1}$ ilw+i < **w, dass, wenn** Le **nicht leer ist, die Summe der maximalen Anzahl von Kanälen**
H

Machine Translated by Google

dass alle laufenden SUe **spenden können, wird immer kleiner sein als ihre minimale Aggregationszahl.** b(x) **ist die Gesamtzahl der genutzten Kanäle im Zustand** x , **die wie folgt beschrieben werden kann:**

$$b(x) = lpu + aga + bgb + \sum_{ym = w\ mlm} \quad (5)$$

Darüber hinaus ist M **ÿ** b(x) **die Anzahl der freien Kanäle im System (d. h., in Schema 1** wird Nf ree angenommen).

Um die Übergangsrate vom allgemeinen Zustand x **zu anderen Zuständen zu analysieren, nehmen wir an, dass die Benutzerankunft im System der Poisson-Verteilung und die Abgänge der Exponentialverteilung gehorchen. Beachten Sie, dass die Ankunftsrate von PU** ÿp **und die Abgangsrate** µp beträgt. **Die Ankunftsraten von** SUr **und** SUe **sind** ÿsr **und** ÿse **und die Abgangsraten sind** µsr **bzw.** µse . **Darüber hinaus kann die Planung mit Rückkopplungsschleife von** SUe **dazu führen , dass die externe Ankunft der in** Le **und** L eintretenden Dienste **nicht mehr der Poisson-Verteilung gehorcht. Um unsere Analyse zu vereinfachen, nehmen wir an, dass die Einstellung der Rückkopplungsschleife wenig Einfluss auf die Dienste in** Le **und** L hat , **und nehmen an, dass die exogene Ankunft von SU**
Alle Dienste unterliegen der Poisson-Verteilung.

Basierend auf den oben genannten Annahmen wird durch Analyse des Kanalzugriffsprozesses das CTMC- Modell basierend auf der Ps-rq-Strategie erstellt. Alle möglichen Zielzustände, Aktivitäten, Übergangsraten und Übergangsbedingungen des Systems ausgehend vom allgemeinen Zustand x **werden entwickelt.**

• Ereignis A: Ankunft der PU

Bei der Ankunft der PU werden die Zustandsübergänge des sekundären Systems ausgehend vom allgemeinen Zustand x **zu allen möglichen Auslöseereignissen wie in Abbildung 4 dargestellt abgeleitet, und die Übergangsbedingungen für alle möglichen Zustände sind in Abbildung 5 aufgeführt. Mit Ausnahme des Zustands t1 in der Abbildung gibt es im System keine freien Kanäle, die der angekommenen PU direkt zugewiesen werden können. Daher existiert** M **ÿ** b(x) = **0 im Zustand t2 bis zum Zustand**

Die Ankunft der PU kann den Dienst aller SUs unterbrechen. Für den Zustand t4 bis t14 und den Zustand t16 bis t17 der Dienstunterbrechung in Abbildung 5 bedeutet dies, dass die SUs in diesen Zuständen keine Spektrumanpassung durchführen können, d. h. ÿm (w < m **ÿ** v), lm = **0. Da** Lr keine **Rückkopplungsschleife hat ,** wird der Dienst des unterbrochenen SUr **zwangsweise beendet, wie beispielsweise im Zustand t30 bis t32.**

Im Zustand t3 und im Zustand t15 kann es zwei Fälle geben, in denen die SUs keine Spektrumanpassung durchführen : Es gibt keinen sekundären Benutzer im System, der eine Anpassung durchführen kann, das heißt ÿm (w < m **ÿ** v), lm = **0, oder die SUs im System können keine Spektrumanpassung durchführen und** ÿm (h = m = v), lm > 0. Ähnliche Bedingungen müssen für die Zustände **t18, t30**

Außerdem gilt, da a = b < **w,** a = b **ÿ** w **ÿ 1. Das heißt, die Kanäle, die durch die in unterbrochene** SUe kann den Dienst von SUe **in** Le **nicht starten , aber der Dienst von** SUr **und** SUh h Lr **und** L **freigegeben durch die zwangsweise** $_t$ **gestartet werden, wie im Zustand t6 bis zum Zustand t14. Dies ist dasselbe wie die Kanäle beendete** SUe , **Kanäle freigegeben durch die unterbrochene oder wie im Zustand t21 bis zum Zustand t29. Allerdings beendete** SUr **und** SUh **starten jeden Dienst von SUs in der Warteschlange. • Ereignis B: PU-Abgang** $_t$ **kann das nicht unterstützen die**

Beim Verlassen der PU werden in Abbildung 6 die Zustandsübergänge des Sekundärsystems ausgehend vom allgemeinen Zustand x **zu allen möglichen Auslöseereignissen abgeleitet und** in Abbildung **7** die Übergangsbedingungen **für alle möglichen Zustände aufgelistet .**

In diesem Fall, wenn die Warteschlange $_t$ L **nicht leer ist und** b = **1, dann ist der Dienst von** SUh $_t$ **in der Warteschlange können direkt gestartet werden. Wenn** b > **1, dann** ist M **ÿ** b(x) = b **ÿ 1 erforderlich, d. h. es existieren** b **ÿ 1 Kanäle im System, wie im Zustand t1. Dies ist dasselbe wie bei** SUr **wie im Zustand t2. Da** a = b < **w, d. h.** w ÿ = **1, kann der Dienst von** SUe **in** Le **nur unter der Bedingung** M **ÿ** b(x) = w **ÿ 1 gestartet werden, wie im Zustand t3.**

Machine Translated by Google

Abbildung 4. Der Übergang des Systems vom allgemeinen Zustand x zu allen erreichbaren Zuständen bei Ankunft der PU.

Machine Translated by Google

Zustand	Übergangsbedingungen	Zustand	Übergangsbedingungen
t1	*M* ÿ *xb* > 0)(	t18	*M* ÿ *xb* = ;0)(≥ ;0 *eq Qeq l* = ; (ÿ = = *wm lvm*), > 0 (ÿ < ÿ *wm lvm*), = 0;
t2	*M* ÿ *xb* = ;0)(l > ;0 *k* = max{ }v+1,...,v ;	t19	*M* ÿ *xb* = ;0)(*lw* > ;1 *eq Qeq l* = ; m= *2w* ÿÿ **v;1** = ;0 ÿ *(wm* ÿ<),lvm = ;0
t3	*M* ÿ *xb* = ;0)(lw > ;0 *gleich Qeq l* = ; (ÿ = =), *wm lvm* 0 (ÿ <), *wm lvm* 0;	t20	*M* ÿ *xb* = ;0)(*lw* > ;1 *gleich Qeq* ; *lq* = ;1 *lhq* = ;0 *rq* *w* (ÿÿ)(*lwv* +ÿ ;1)1 = ;0 1 << *vw ;* ÿ *(wm lvm* ÿ<),
t4	*M* ÿ *xb* = ;0)(> ;1 *eq Qeq l* < ; *m* = *2w* ÿÿ **v;1** = ;0 *wm lvm* ÿ<),	t21	*M* ÿ *xb* = ;0)(*gleich Qeq* ; ≥ ;0 *w* ÿÿ **b;1** ÿ *(wm* ÿ<),lvm = ;0
t5	*M* ÿ *xb* = ;0)(*lw* > ;1 *gleich Qeq l* = ; *lq* ÿ ;1 1 << *vw ;* *w* ÿ(ÿ)(+ÿ ;1)1 ÿ *(wm* ÿ<),lvm = ;0	t22	*M* ÿ *xb* = ;0)(*gleich Qeq* ; > ;0 *w* ÿÿ **a;1** = ;0 ÿ *(wm* ÿ<),lvm = ;0
t6	*M* ÿ *xb* = ;0)(*gleich Qeq l* = ; ≥ ;0 *w* ÿÿ **b;1** ÿ *(wm* ÿ<),lvm = ;0	t23	*M* ÿ *xb* = ;0)(> ;0 *gleich Qeq* ; > ;0 ≥ ;0 *w* **1** +ÿÿ *ba ;* ÿ *(wm lvm* = ;0
t7	*M* ÿ *xb* = ;0)(*gleich Qeq l* = ; > ;0 *w* ÿÿ **a;1** *lhq* = ;0 ÿ *(wm* ÿ<),lvm = ;0	t24	*M* ÿ *xb* = ;0)(*eq Qeq l* = ; *lhq* > ;0 *lw* > ;1 **1** << *vw ;* *wb* **1** +<+ÿ=ÿ< *abbwk ;* ÿ *(wm lvm* = ;0
t8	*M* ÿ *xb* = ;0)(*lw* > ;0 *gleich Qeq l* = ; > ;0 *lhq* > ;0 *w1* + ÿÿ *ba ;*	t25	*M* ÿ *xb* = ;0)(*eq Qeq l* = ; *lhq* > ;0 *wb* 1 +ÿ=ÿ< lwvb ;1 *lq* = *lw* ;1 ÿ *(wm* < ÿ),lvm = ;0 1< < *vw ;*
t9	*M* ÿ *xb* = ;0)(*Gleichung Qeq l* < ; *lhq* > ;0 *lw* > ;1 **1** << *vw ;* *wb* **1** +<+ÿ=ÿ< *abbwk ;* ÿ *(wm lvm* = ;0	t26	*M* ÿ *xb* = ;0)(*eq Qeq l* = ; > ;0 *lw* > ;1 *wb* **1** +ÿ=ÿ< *awk ;* *lhq* = ;0 ÿ *(wm* ÿ<),lvm = ;0 1 << *vw ;*
t10	*M* ÿ *xb* = ;0)(*gleich Qeq l* = ; > ;0 *wb* 1 +ÿ=ÿ<) ÿ ;1 = *lq* ÿ ;1 *lw* > ;1 ÿ *(wm* ÿ<),lvm = ;0 1 << *vw ;*	t27	ÿ *xbM* = ;0)(*eq Qeq l* = ; > ;0 = ;0(ÿ+=ÿ< ÿ ;1) *wb lwva lhq* *lw* > ;1 *lq* = ;1 ÿ (ÿ<),lvmwm = ;0 1 << *vw ;*
t11	*M* ÿ *xb* = ;0)(*Gleichung Qeq l* < : > ;0 *lw* > ;1 *wb* 1 +ÿ=ÿ< *awk ;* *lhq* = ;0 1 << *vw ;* ÿ *(wm* ÿ<),lvm = ;0	t28	ÿ *xbM* = ;0)(; eq=Ql *eq* > ;1 > ;0 ≥ ;0 *w1* ÿ ++=ÿ *wkba ;* ÿ (ÿ<),lvmwm = ;0 1 << *vw ;*
t12	*M* ÿ *xb* = ;0)(*gleich Qeq l* = ; > ;0 *wb* 1 +ÿ=ÿ< lwva ;1 *lhq* = ;0 *lq* = ;1 *lw* > ;1 ÿ *(wm* ÿ<),lvm = ;0 1 << *vw ;*	t29	ÿ *xbM* = ;0)(*eq Qeq l* = ; *lhq* > ;0 > ;0 *lq* = ;1 *w* 1 (ÿ++=ÿ)*(lwvba* ÿ1); *lwvwm* 1 ;0 << *vw ;*
t13	*M* ÿ *xb* = ;0)(*gleich Qeq l* = ; > ;1 > ;0 > ;0 *w* **1** ÿ++=ÿ *wkba ;* **1** << *vw ;* *wm lvm* ÿ (< ÿ), = ;0	t30	ÿ *xbM* = ;0)(*ga* > ;0 (ÿ = =), *lvmwm* 0 (ÿ <), *lvmwm* 0;
t14	*M* ÿ *xb* = ;0)(*gleich Qeq l* = ; > ;0 *lq* = ;1 > ;1 *w* 1 (ÿ++=ÿ)*(lwvba* ÿ1); 1 << *vw ;* *lwvwm* = ;0	t31	ÿ *xbM* = ;0)(*G* > ;0 *ak* ÿ= ;1 1 ÿ<< *vkw ;* > ;0 ÿ *(wm* ÿ<),lvmwm = ;0 *Abonnieren*
t15	*M* ÿ *xb* = ;0)(gb > ;0 *hq Qhq l* < ; (ÿ = =), *wm lvm* 0 (ÿ <), *wm lvm* 0;	t32	ÿ *xbM* = ;0)(*ga* > ;0 *lq* = ; *a* ÿ=(1 *lwv* > ;0 ÿ (ÿ<),lvmwm = ;0 1 << *vw ;*
t16	*M* ÿ *xb* = ;0)(gb > ;0 *lhq Qhq* ; *bk* ÿ= ;1 1 ÿ<< *vkw ;* *lw* > ;0 ÿ *(wm* ÿ<),lvm = ;0	t33	ÿ *xbM* = ;0)(gb > ;0 *hq Qhq l* = ; (ÿ = =), *lvmwm* 0 *Abonnieren* (ÿ <), *lvmwm* 0;
t17	*M* ÿ *xb* = ;0)(gb > ;0 *hq Qhq l* < ; *lq* = ; *b* ÿ=(1 *lwv* 1 << *vw ;* > ;0 ÿ *(wm* ÿ<),lvm = ;0	t34	ÿ *xbM* = ;0)(> ;0 *b g* *hq Qhq l* = ; *bk* ÿ= ;1 1 ÿ<< *vkw ;* > ;0 ÿ (ÿ<),lvmwm = ;0
		t35	ÿ *xbM* = ;0)(gb > ;0 *lhq Qhq* = *lq* ; *b* ÿ=(1 *lwv* *lw* > ;0 ÿ (ÿ<),lvmwm = ;0 1 << *vw ;*

Abbildung 5. Die Übergangsbedingungen des Systems vom allgemeinen Zustand x **zu allen erreichbaren Zuständen auf PU Ankunft.**

Machine Translated by Google

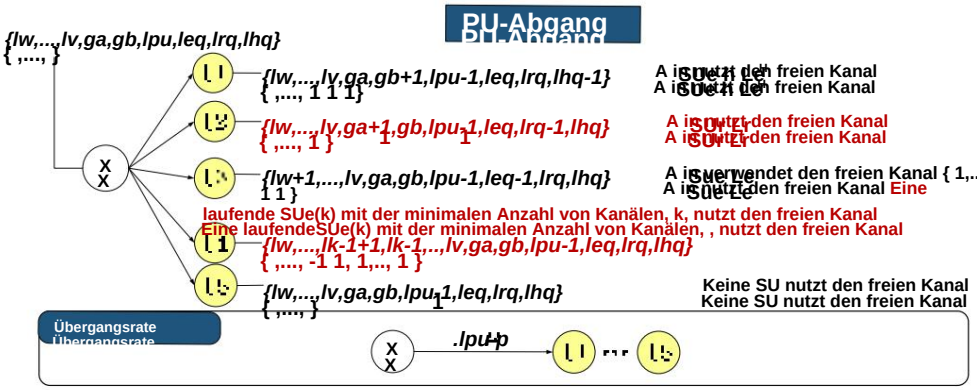


Abbildung 6. Der Übergang des Systems vom allgemeinen Zustand x **zu allen erreichbaren Zuständen beim Verlassen der PU.**

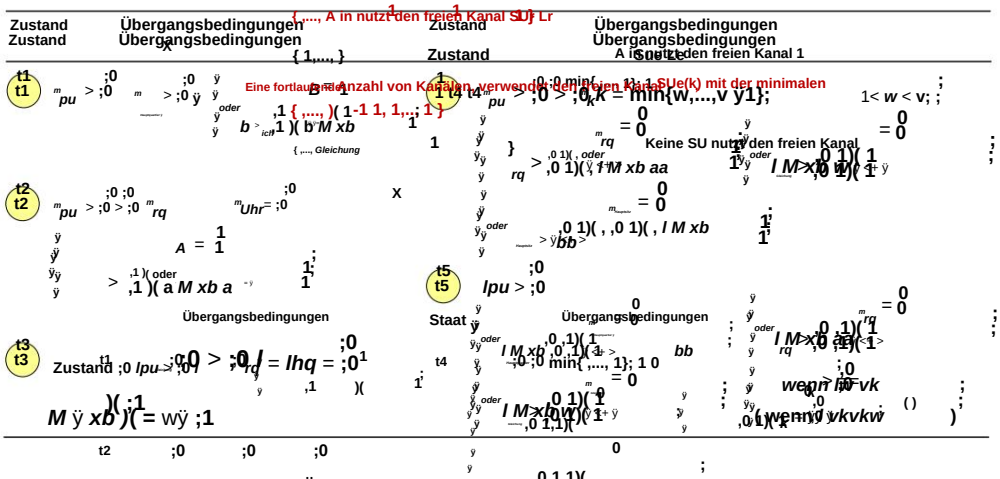


Abbildung 71Die Übergangsbedingungen des Systems vom allgemeinen Zustand x **zu allen erreichbaren Zuständen auf PU**

,1 Abfahrt.

• Ereignis 2: SU Ankunft

Bei der Ankunft des SU werden die Zustandsübergänge des Sekundärsystems beginnend mit dem

Der allgemeine Zustand x **zu allen möglichen auslösenden Ereignissen wird wie in Abbildung 8 dargestellt abgeleitet, und die**

Die Übergangsbedingungen für alle möglichen Zustände sind in Abbildung 9 aufgeführt.

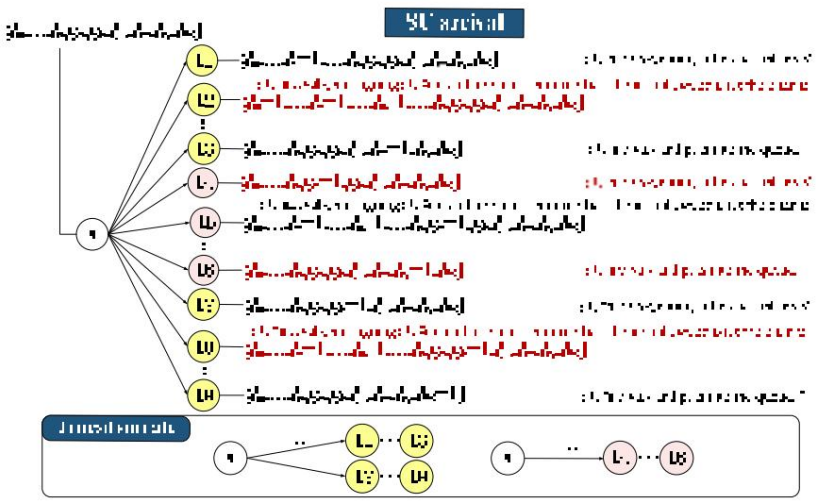


Abbildung 8. Der Übergang des Systems vom allgemeinen Zustand x **zu allen erreichbaren Zuständen bei Ankunft von SU.**

Machine Translated by Google

Zustand	Übergangsbedingungen	Zustand	Übergangsbedingungen
t1	*M* ÿ *xb*)(> w;	t6	*rq Qrq l* < *wenn* ÿ ÿ ÿ

Machine Translated by Google

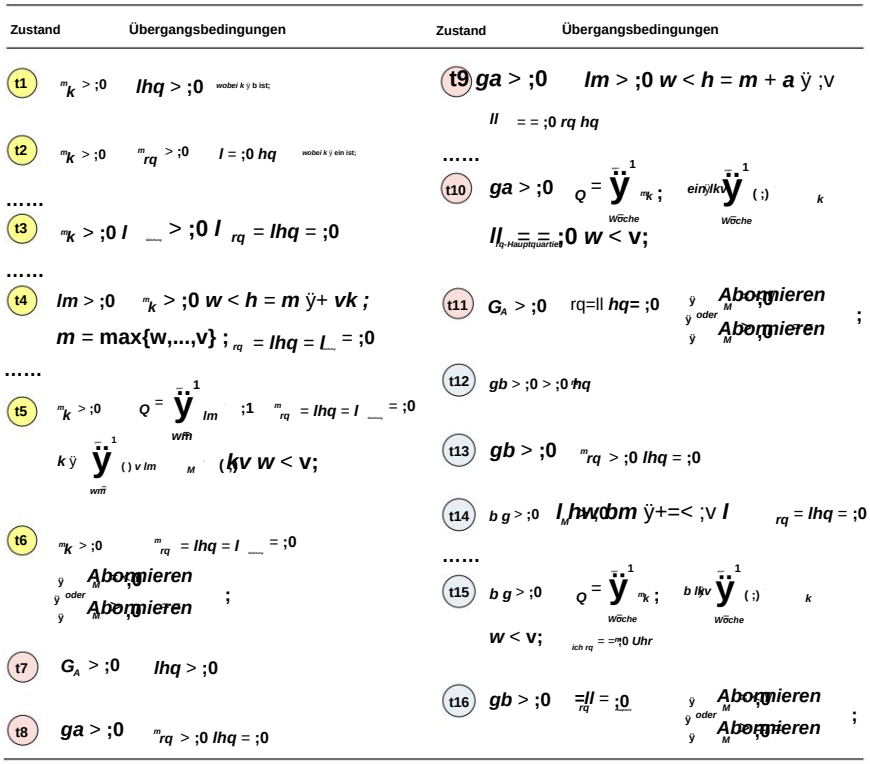


Abbildung 11. Die Übergangsbedingungen des Systems vom allgemeinen Zustand x **zu allen erreichbaren Zuständen auf SU-Abfahrt.**

In diesem Fall sollte die Gesamtzahl der von den laufenden SUs genutzten freien Kanäle nicht größer als die Anzahl der von den abgehenden SUs freigegebenen Kanäle, wie z. B. die Bedingung von k ÿ ÿ(v ÿ m)lm ÿ (v ÿ k) **im Zustand t5. Dies ist ähnlich wie im Zustand t10 und im Zustand t15.**

4.2. QoS-Kennzahlen

Durch CTMC-Modellierung kann jede mögliche Benutzeraktivität den Übergang auslösen von jeden erreichbaren Zustand ti **in einen anderen erreichbaren Zustand** tj , **wie in den Abbildungen 4–11 gezeigt.** qij (i, j ÿ S und i ÿ= j) **stellt die gesamte Übergangsrate dar, wobei** S **die Menge der möglichen Zustände des Systems ist. Sei,** ÿ(x) **die Steady-State-Wahrscheinlichkeit des Systems, die sich aus globalen Bilanzgleichungen und die Normalisierungsbedingung wie folgt.**

$$\ddot{y}(x)Q = \mathbf{0}, \ddot{y}x \ddot{y}(x) = \mathbf{1} \quad \mathbf{(6)}$$

wobei ÿ(x) **der stationäre Wahrscheinlichkeitsvektor und** Q **die Übergangsratenmatrix ist, die durch die Übergangsrate** qij erzeugt wird. **Die diagonalen Elemente der Übergangsratenmatrix Q, d findet sich als** qii = ÿ ÿ qij. **Somit basieren die QoS-Metriken des sekundären Systems auf dem**

Die vorgeschlagene Ps-rq-Strategie kann bezogen werden.

- **Die Netzwerkkapazität bezieht sich auf die durchschnittliche Anzahl der in CRNs abgeschlossenen SU-Dienste pro Zeiteinheit (Einheit: Paket/Zeiteinheit) [12]. Seien** Ce , Cr **die Kapazitäten von** SUe , SUr , **bzw. können wir erhalten:**

$$Cr = \ddot{y}_{x\ddot{y}S} \; agaµsr\ddot{y}(x), \quad \mathbf{(7)}$$

$$Ce = \ddot{y}_{x\ddot{y}S} \; \ddot{y}_{k=w} \; (klk + bgb)µse\ddot{y}(x) \; \mathbf{\ddot{y}} \; \mathbf{(8)}$$

wobei x **ÿ** S **die Menge aller möglichen Zustände in den in den Abbildungen 4–11 enthaltenen Systemen ist.**

Machine Translated by Google

• **Die Spektrumnutzung bezieht sich auf das Verhältnis der durchschnittlichen Anzahl der genutzten Kanäle zur Gesamtzahl der verfügbaren Kanäle [25], das wie folgt bestimmt werden kann:**

$$U = \ddot{y} \underset{x\ddot{y}S}{\ddot{y}(x)} \frac{b(x)}{M} \quad (9)$$

wobei b(x) **die Gesamtzahl der verfügbaren Kanäle des Zustands x,** M **die Gesamtzahl der Kanäle im System und** S **die Menge aller möglichen Übergangszustände in den Abbildungen 4–11 ist.**

• **Die Blockierungswahrscheinlichkeit bezieht sich auf die Wahrscheinlichkeit, dass die neu angekommene SU nicht bedient werden kann, weil alle Kanäle im System belegt sind. Da in dieser Studie die Warteschlange für SUs eingerichtet ist, wird der Service der neu angekommenen SU blockiert, wenn die entsprechende Warteschlange voll ist. Es ist zu beachten, dass die** SUh h die von der PU **unterbrochene eingehende** L SUe **ist , nicht jedoch die SUs, die neu im System ankommen.** t

Der Fall einer Sperrung liegt für die SUh **daher** nicht vor.

Die neu eingetroffene SUe **wird blockiert, wenn die beiden folgenden Bedingungen erfüllt sind: (1) Die Summe der Gesamtzahl der freien Kanäle im System und der Gesamtzahl der Kanäle ,** die von der laufenden SUe **gespendet werden können, reicht nicht aus, um den Dienst der** SUe zu starten **(d. h . die Untergrenze der Kanalzusammenstellung wird nicht erreicht) und (2) die Warteschlange** Le **ist voll. Somit ist die Blockierungswahrscheinlichkeit von** SUe , P_b^t, **erhalten Sie durch:**

$$P_b^t = \ddot{y} \underset{AE}{\ddot{y}(x)} \quad (10)$$

wobei der Satz Ae **die Auslösebedingung ist, für die** SUe **blockiert wird:**

$$Ae = x \ddot{y} S, leq = Qeq, M \ddot{y} b(x) + \ddot{y}_{k = w + 1}^{\infty} (k \ddot{y} w)lk < w$$

Ebenso **wird der neu eingetroffene** SUr blockiert, wenn die beiden folgenden Bedingungen **erfüllt sind: (1) Die Summe der Gesamtzahl der freien Kanäle im System und der Gesamtzahl** der Kanäle, die von der laufenden SUe **gespendet werden können, reicht nicht aus, um den Dienst des** SUr zu starten , **und (2) die Warteschlange** Lr ist voll. Dann **kann die Blockierungswahrscheinlichkeit von** SUr ,P_b^R, **wie folgt bestimmt werden:**

$$P_b^R = \ddot{y} \underset{Ar}{\ddot{y}(x)} \quad (11)$$

wobei der Satz Ae **die Auslösebedingung ist, für die** SUe **blockiert wird:**

$$Ar = x \ddot{y} S, lrq = Qrq, M \ddot{y} b(x) + \ddot{y}_{k = w + 1}^{\infty} (k \ddot{y} w)lk < a$$

• **Die Wahrscheinlichkeit einer erzwungenen Beendigung bezieht sich auf die Wahrscheinlichkeit, dass die bedienten SUs aufgrund der Ankunft einer PU zur Beendigung gezwungen werden. Als Satz der Rückkopplungsschleife von** SUe können **die** von der PU unterbrochenen SUe **über** den H Paketklassifizierer II in die Warteschlangen Le **und** L e **zurückkehren . Wenn die entsprechende Warteschlange voll ist, wird der Dienst zu**

Dementsprechend wird der Dienst von SUh **zwangsweise beendet, wenn die folgenden drei Bedingungen erfüllt sind: (1) Es gibt keine freien Kanäle im System; (2) die laufende** SUe **kann keine Spektrumanpassung durchführen, um Kanäle mit den unterbrochenen SUs zu teilen, und (3) die Warteschlange** h **ist voll. Genauer gesagt zeigen die Zustände t33 bis t35 in Abbildung 4 an, dass die PU alle** SUh L e , t **mit** b **aggregierten Kanälen und der** SUh t **wird zur Beendigung gezwungen. Die erzwungene Beendigung**

Machine Translated by Google

Wahrscheinlichkeit $_{e}$, $_{Pf}^{eh}$, **kann berechnet werden durch das Verhältnis der durchschnittlichen Zwangskündigung von** SUh **Wahrscheinlichkeit zur Ankunftsrate** von SUh e , ÿse, **wie folgt:**

$$P_f^{eh} = \ddot{y}\ \ddot{y}(x)_{B_t^H} \frac{\ddot{y}pbgb}{(M\ \ddot{y}\ lpu)\ddot{y}se} \quad (12)$$

h **wobei** B_t **ist der Satz von Bedingungen, unter denen** SUh **zur Beendigung gezwungen wird:**

$$B_t^H = x\ \ddot{y}\ \mathbf{S,}\ M = b(x),\ lpu < \mathbf{M,}\ gb > \mathbf{0,}\ lhq = Qhq,\ \text{oder}\ \begin{matrix} \ddot{y}m(w < m\ \ddot{y}\ v),\ lm = \mathbf{0} \\ \ddot{y}m(w = m = v),\ lm > \mathbf{0} \end{matrix}$$

Ebenso wird der Dienst von SUe **zwangsweise beendet, wenn die folgenden drei Bedingungen erfüllt sind: (1) Nach der Verdrängung reicht die Anzahl der verbleibenden Kanäle von** SUe **nicht aus, um den Dienst zu starten, d. h. die** SUe **mit** w **aggregierten Kanälen wird verdrängt; (2) die laufende** SUe **kann keine Spektrumanpassung durchführen, um Kanäle mit den unterbrochenen SUs zu teilen, und (3) die Warteschlange** Le **ist voll. Auf dieser Grundlage zeigen die Zustände t18 bis t29 in Abbildung 4 an, dass die** PU jede SUe **mit** w aggregierten Kanälen **verdrängt , und die** SUe **wird zwangsweise beendet. Die Wahrscheinlichkeit der erzwungenen Beendigung von** SUe , P, **kann wie folgt ermittelt werden:** $_{t}$ f ,

$$P_f^t = \ddot{y}\ \ddot{y}(x)_{Sei} \frac{\ddot{y}pwlw}{(M\ \ddot{y}\ lpu)\ddot{y}e} \quad (13)$$

wobei ÿe **die Ankunftsrate von** SUe darstellt , **die nicht blockiert ist: ÿe** = ÿse(1 **ÿ** P **die Menge** $_{B}^{t}$) **und** Be **ist der Bedingungen, unter denen** SUe **zum Beenden gezwungen wird:**

$$Be = x\ \ddot{y}\ \mathbf{S,}\ M = b(x),\ lpu < \mathbf{M,}\ lw > \mathbf{0,}\ leq = Qeq,\ \text{oder}\ \begin{matrix} \ddot{y}m(w < m\ \ddot{y}\ v),\ lm = \mathbf{0} \\ \ddot{y}m(w = m = v),\ lm > \mathbf{0} \end{matrix}$$

Da die Rückkopplungsschleife für die Warteschlange Lr nicht eingestellt ist, wird **der Dienst von** SUr **zwangsweise beendet, wenn die folgenden beiden Bedingungen erfüllt sind :** (1) Es gibt keine freien Kanäle im System, um **den Dienst fortzusetzen, und (2) die laufende** SUe **kann die Kanäle nicht dafür spenden. Auf dieser Grundlage zeigen die Zustände t30 bis t32 in Abbildung 4** an, dass die PU jedem SUr **mit** aggregierten Kanälen **zuvorkommt und der** SUr **zwangsweise beendet wird. Die Wahrscheinlichkeit einer erzwungenen** $_{Pf}^{R}$, **Beendigung von SUr** kann **wie folgt ermittelt werden:**

$$P_f^R = \ddot{y}\ \ddot{y}(x)_{Br} \frac{\ddot{y}paga}{(M\ \ddot{y}\ lpu)\ddot{y}r} \quad (14)$$

wobei ÿr **die Ankunftsrate von** SUr darstellt , **die nicht blockiert ist: ÿr** = ÿsr(1 **ÿ** P **die Menge der** $_{B}^{R}$) **und** Br **ist Bedingungen, unter denen** SUr **zur Beendigung gezwungen wird:**

$$Br = x\ \ddot{y}\ \mathbf{S,}\ M = b(x),\ lpu < \mathbf{M,}\ ga > \mathbf{0,}\ \text{oder}\ \begin{matrix} \ddot{y}m(w < m\ \ddot{y}\ v),\ lm = \mathbf{0} \\ \ddot{y}m(w = m = v),\ lm > \mathbf{0} \end{matrix}$$

5. Leistungsanalyse der vorgeschlagenen Ps-rq-Strategie 5.1.

Vergleich und Analyse der Modelleigenschaften **In diesem**

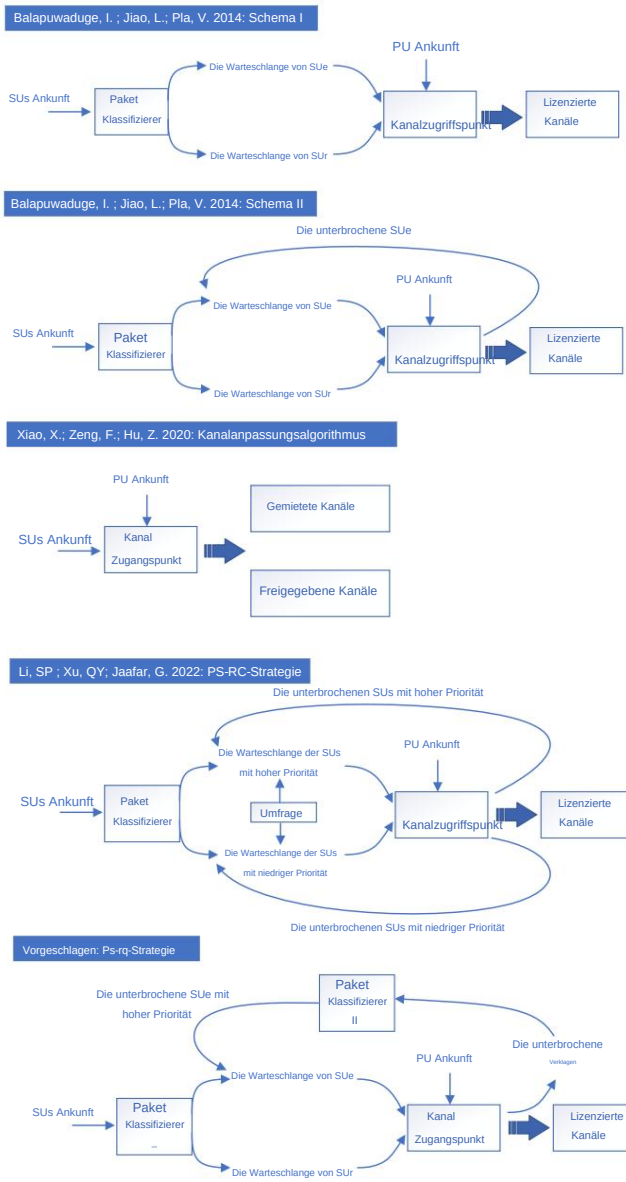
Abschnitt vergleichen wir die vorgeschlagene Ps-rq-Strategie mit zwei klassischen Kanalzuweisungsschemata (Referenz [12], 2014) sowie zwei neuartigen (Referenz [18], 2020 und Referenz [9], 2022), um die Eigenschaften der vorgeschlagenen Strategie zu veranschaulichen. Der Vergleich der Modelleigenschaften ist in Abbildung 12 dargestellt.

Machine Translated by Google

. Fehler, die nicht durch einen Vorschlag verursacht werden,

5.1. Vergleich und Analyse der Modelleigenschaften

In diesem Abschnitt vergleichen wir die vorgeschlagene Ps-rq-Strategie mit zwei klassischen Kanalzuweisungsschemata (Referenz [12], 2014) sowie zwei neuartigen (Referenz [18], 2020 und Referenz [9], 2022) mit dem Ziel, die Merkmale der vorgeschlagenen Strategie zu veranschaulichen. Der Vergleich der Modelleigenschaften ist in Abbildung 12 dargestellt.



Verkehrsart	Kapazitätserweiterung	Verdienste	Einschränkungen
Heterogener Dienst	/	· Einfaches Schema · Kritische Anwendungen verzögern	· Hohe Wahrscheinlichkeit einer Zwangsabbruchs · Nur anwenden, wenn PU nicht aktiv ist
Heterogener Dienst	Rückkopplungsschleife	· Reduzierung von Zwangsabbrüchen Wahrscheinlichkeit · Verzögerungstolerante Anwendungen	·Die Warteschlangen von SUe verursachen ein Service-Starvation-Problem
/	Frequenzleasing	· Einfaches Schema · Hohe Realisierbarkeit	·Die Netzwerkkapazität wird durch die Größe des gemieteten Spektrums beeinflusst · Heterogene Dienste sind nicht bedacht
Vorzugsservice	Umfragen, Feedbackschleife	· Sicherstellung der vordefinierten Dienste mit hoher Priorität	·Die Priorität muss während der Netzwerkinitialisierung definiert werden ·Schwierige Anpassung an die sich ändernde Netzwerkumgebung
Heterogener Dienst	Rückkopplungsschleife	· Hunger lindern Problem der SUe · Faire Terminplanung	· Die Servicetypen sind in Details unterteilt · Hohe Planungskomplexität

Abbildung 12. Der Vergleich der Modelleigenschaften [9,12,18].

Die genannten Forschungsmodelle, wie in Abbildung 12 dargestellt, zielen darauf ab, die Kapazität zu erhöhen

Die genannten Forschungsmodelle, wie in Abbildung 12 dargestellt, zielen darauf ab, die der sekundären Netzwerke. Durch den Einsatz der Zufallsanalysemethode und die Erstellung der

Kapazität der sekundären Netzwerke zu erhöhen. Durch den Einsatz der Methode der Zufallsanalyse und die Fe

entsprechenden kontinuierlichen Markov-Kette für die systematische Analyse wird die Leistung

Die entsprechende kontinuierliche Markov-Kette für die systematische Analyse wird verwendet, um

Leistungsmetriken abzuleiten, um die Wirksamkeit der Forschungsprogramme zu bewerten. Allerdings

Es werden Metriken abgeleitet, um die Wirksamkeit der Forschungsprogramme zu bewerten. Die

genannten Studien hatten jedoch Schwierigkeiten, die verschiedenen Servicetypen zu unterscheiden, wie gezei

Die genannten Studien hatten Schwierigkeiten bei der Unterscheidung von Servicetypen, wie in

[9,18] gezeigt, oder stießen auf Einschränkungen bei der Implementierung differenzierter Serviceanwendu

aufgrund von Einschränkungen der Netzwerkumgebung, wie in [12].

Diese Einschränkungen dienen als Inspiration für die in diesem Dokument beschriebene Strategie. Die in diesem Papier vorgeschlagene PS-RQ-Strategie zielt auf eine gerechtere Differenzierung ab Service, indem wir das Problem des Hungerns angehen. Um dies zu erreichen, verwenden wir Paketklassifizierer I und II, um zwei Ebenen der Dienstdifferenzierung durchzuführen. Der Zweck des Paketklassifizierers I ist die ankommenden SUs in SUe **und** SUr **aufzuteilen . Wenn man bedenkt, dass wenn die** SUe **durch die PU ist alle in die Warteschlange für den erneuten Service, es wird zwangsläufig zu dem Hungerproblem eines anderen führen Warteschlange. Das in diesem Artikel vorgeschlagene Modell berücksichtigt daher die Serviceanforderungen von** die teilweise unterbrochene SUe **mit hoher Priorität und integriert den Paketklassifizierer 2, um eine faire Dienstleistungserbringung bei gleichzeitiger Bereitstellung differenzierter Dienstleistungen, was einen bemerkenswerten Aspekt darstellt**

Machine Translated by Google

unseres Forschungsrahmens. Die Komplexität der vorgeschlagenen Planungsstrategie bleibt jedoch angesichts der Verfeinerung der Serviceklassifizierung eine Herausforderung. Als Nächstes werden wir eine detaillierte Leistungsanalyse der in diesem Artikel vorgestellten Forschungsstrategie Ps-rq durchführen.

5.2. Numerische Simulation und Analyse

Um die Leistung des sekundären Systems basierend auf der vorgeschlagenen Kanalzusammenstellungsstrategie zu veranschaulichen , werden in diesem Abschnitt numerische Simulationsergebnisse von elastischen Diensten und Echtzeitdiensten basierend auf MATLAB vorgestellt. Die Simulationsergebnisse basieren auf dem Verfahren zur Dimensionsreduzierung hochdimensionaler Markow-Ketten. Um das entwickelte Modell zu überprüfen, präsentieren wir die Simulationsergebnisse der Leistung des sekundären Systems bei unterschiedlichen Ankunftsraten der primären Benutzer. Die grundlegenden Parameter der numerischen Simulation werden wie folgt festgelegt.

Die Anzahl der PUs im System ist auf 6 (M = 6) eingestellt, **und das numerische Experiment wird in einem CRN mit sechs Kanälen durchgeführt, die sich die Spektrumbandbreite teilen. Die Ankunftsrate der PU ist auf** ÿp **= 1 und die Abgangsrate auf** µp **= 0,5 eingestellt. Die Ankunfts- und Abgangsrate des elastischen Dienstes sind auf** ÿse **= 2 bzw.** µse **= 1 eingestellt, und die des Echtzeitdienstes sind auf** ÿsr **= 1 und** µsr **= 1 eingestellt. Beachten Sie, dass die Einheit dieser Parameter „Sitzungen pro Zeiteinheit“ ist. Durch Zuweisen bestimmter Werte zu diesen Parametern, beispielsweise 0,5 Sitzungen/Sek. und Flusslänge in Bits, kann die Kapazität in kbps oder Mbps quantifiziert werden.**

Die dynamischen Parameter des Systems (a, b, w, v, Qeq, Qrq, Qhq) **im Spektrumzugriffsprozess werden wie folgt festgelegt:** a **=** b **= 1,** w **= 2,** v **= 4,** Qeq **= 2,** Qrq **= 1 und** Qhq **= 1. Der proportionale Verteilungskoeffizient beträgt** *ÿ : ÿ* = **1 : 1. Basierend auf den obigen Systemeinstellungen kann der allgemeine Zustand** x **des Systems in Abschnitt 4.1 wie folgt angegeben werden:**

$$(\mathrm{l2}\ \mathbf{l3}\ \mathbf{l4}\ \mathbf{g1}\ \mathbf{g1}\ {}_{6\,6}^{3\ 2\ 1}\ \ \mathrm{lpu}\ \ \mathrm{Qeq}\ \ \mathrm{Qrq}\ \ \mathrm{Qhq})^{6\ \ 2\ \ 1\ \ 1} \tag{15}$$

wobei li (i = **2, 3,** 4) SUe **darstellt ,** die i **Kanäle aggregiert ,** und ga, gb SUr **mit** a **aggregierten Kanälen oder** SUh **mit** b **aggregierten Kanälen** (a = b = 1) **darstellen ; der hochgestellte Index jeder Position gibt die Obergrenze der Anzahl der Benutzer an der entsprechenden Position, wie l3, was bedeutet, dass es höchstens zwei** SUe gibt , **die drei Kanäle gleichzeitig im System,** lpu, **was bedeutet, dass höchstens sechs PUs im System vorhanden sind System gleichzeitig und** Qeq, **was bedeutet, dass die maximale Kapazität der Warteschlange** Le **2 ist.**

Die numerische Simulation basiert auf dem parametrisierten allgemeinen Zustand. Das Verfahren zur Dimensionsreduktion hochdimensionaler Markow-Ketten wird wie folgt dargestellt.
Da der in diesem Dokument angenommene allgemeine Zustand hochdimensional ist, ist es schwierig, die Übergangsratenmatrix Q direkt zu erzeugen. Daher ziehen wir eine Herabstufung der hochdimensionalen Markierungen in Erwägung. Theoretisch beträgt die Größe der Matrix Q **2592 × 2592, die durch das Produkt der hochgestellten Ziffern von Gleichung (15) erzeugt wird. Nehmen wir einen beliebigen allgemeinen Zustand** l : **(1 0 0 1 1 1 0 0** 0), **der im System existiert, um den Herabstufungsprozess wie folgt darzustellen. Der Zustand** l : **(1 0 0 1 1 1 0 0** 0) **gibt an, dass das System eine laufende** SUe **mit zwei aggregierten Kanälen, eine laufende** SUr **mit einem aggregierten Kanal, eine laufende** SUh **mit einem aggregierten Kanal und eine laufende PU hat. Daher existieren in diesem Zustand fünf Flows im System. Dementsprechend kann Zustand** l : **(1 0 0 1 1 1 0 0** 0) als Zustand **l5, ÿ gekennzeichnet werden , wobei die erste Stelle des Index die Gesamtzahl der koexistierenden Flows im System darstellt, die zweite Stelle die Zählreihenfolge und die Obergrenze die Summe der möglichen Kombinationszustände mit fünf Flows im System ist .** Wenn in diesem Zustand ein Benutzer im System ankommt, etwa ein PU-Ankunft, dann **ist der entsprechende Zustandsübergang (1 0 0 1 1 1 0 0** 0) **ÿ (1 0 0 1 1 2 0 0** 0). **In diesem Fall existieren sechs Flüsse gleichzeitig im System und der Übergangsprozess wird als l5, ÿ ÿ l6, ÿ** gekennzeichnet .

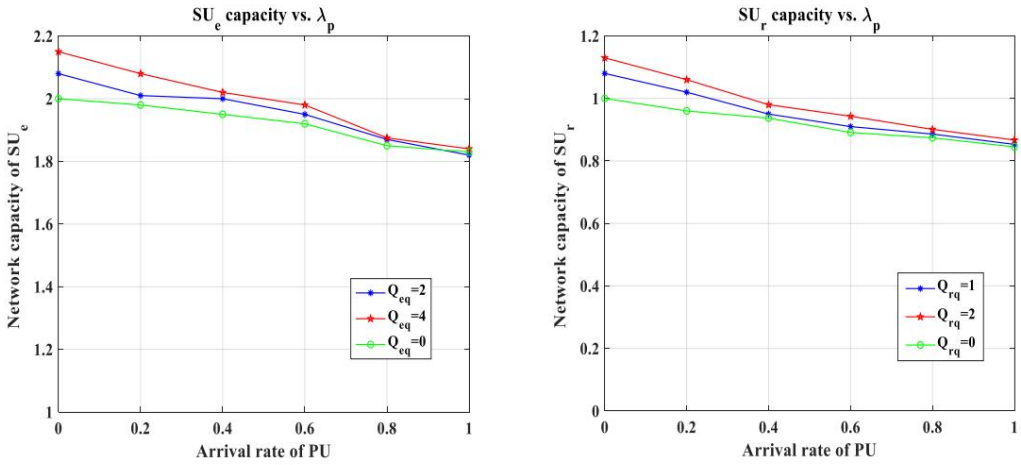
Wenn ein Benutzer abreist in

Machine Translated by Google

das System in diesem Zustand, wie z. B. ein SUAbfahrt, dann der entsprechende Zustandsübergang
ist (1 0 0 1 1 1 0 0 0) ÿ (0 1 0 1 0 1 0 0 0). In diesem Fall gibt es vier Flüsse, die in der
System und der Übergangsprozess wird als I5, ÿ ÿ I4, ÿ bezeichnet . Gemäß den Abbildungen 4–11
Die möglichen Übergangsereignisse und Übergangsraten des Systems in jedem allgemeinen Zustand können
erhalten werden, und dann kann die Übergangsratenmatrix Q entwickelt werden. Darüber hinaus ist die
Steady-State-Wahrscheinlichkeit ÿ(x) kann gelöst werden und die Leistungsmetriken der sekundären
System einschließlich seiner Netzkapazität, Frequenznutzung, Sperrwahrscheinlichkeit und Zwangs
Die Abbruchwahrscheinlichkeit kann weiter simuliert werden. Wir verwenden MATLAB, um die Markov-
Modell zu erstellen und Simulationen durchzuführen. Auf dieser Grundlage werden vier Fallstudien vorgestellt und der Kanal
Die Montageleistung basierend auf der vorgeschlagenen Ps-rq-Strategie wird analysiert.

Fall 1: Kapazität des sekundären Netzwerks vs. Ankunftsrate von PUs

Abbildung 13 zeigt die Beziehung zwischen der Netzkapazität von SUe und SUr und der
Ankunftsrate von PUs. Wie aus der Abbildung ersichtlich, ist die Kapazität des sekundären Netzwerks
sinkt mit der Zunahme der Ankunftsrate von PUs. Wenn die Anzahl der PUs, die in die
System erhöht, reduziert die absolute Datenzugriffspriorität die Gesamtzahl der verfügbaren
verbleibende Kanäle im System, so dass die Anzahl der SUs, die die Dienste in der
wird sich entsprechend verringern. Die Netzkapazität von SUe und SUr wird daher
nimmt mit der Zunahme der Ankunftsrate von Dekubituserkrankungen ab.



(A) (B)

Abbildung 13. (a) Netzkapazität von SUe; (b) Netzkapazität von SUr.

Abbildung 13. (a) Netzkapazität von SUe; (b) Netzkapazität von .

Aus Abbildung 13 ist jedoch ersichtlich, dass eine Erhöhung der Gesamtkapazität der Warteschlange
von SUs erhöht die Kapazität des sekundären Netzes. Bei ÿp = 0,2 ist die Kapazität des
Warteschlange von SUe , Qeq, steigt von 0 auf 4, und die sekundäre Netzwerkkapazität steigt um
Kapazität der Warteschlange von SUr , Qrq, von 0 auf 4 steigt, was einer Steigerung von 9,43 % entspricht.
Netzwerkkapazität. Nachdem die Warteschlangenkapazität erhöht wurde, ist die Anzahl der SU
werden die Flüsse, die im System koexistieren können, zugenommen, was die Wahrscheinlichkeit erhöht, dass
wird in der Lage sein, den Dienst anzunehmen. Durch den Vergleich von Abbildung 13 ist es nicht schwer zu erkennen, dass die
genutzt wird, ist die Netzwerkkapazität von SUr geringer als die von SUe . Dies liegt daran, dass
hat eine höhere Serviceprioritätals SUe (mit Ausnahme der unterbrochenen SUh ,), so dass es einfacher ist,
kann, starten Sie den Dienst für den neu eingetroffenen SUr statt für SUe .

Fall 2: Spektrumnutzung vs. Ankunftsrate von SUe

Die Beziehung zwischen Spektrumnutzung und Ankunftsrate von SUs ist in
Abbildung 14 mit der Ankunftsrate von SUe als x-Achse. Aus der Abbildung ist ersichtlich, dass die
Die Spektrumnutzung steigt mit der Zunahme der Ankunftsrate. Wenn Qrq = 2, Qeq = 4,
Die Frequenznutzungsrate beträgt bis zu 92,3 %. Mit der Zunahme der Ankunftsrate von SUs,
das ungenutzte Spektrum wird voll ausgenutzt. Wenn die Spektrumauslastung gesättigt ist
kommt allmählich gesättigt, der Gradient wird abnehmen. Darüber hinaus ist es nicht schwierig
dass durch Erhöhen der Warteschlangenkapazität die sekundäre Netzwerkkapazität erhöht werden
Die Tankauslastung wird entsprechend verbessert.

Machine Translated by Google

ase: Spektrumnutzung vs. Ankunftsrate o *e*

Die Beziehung zwischen der Spektrumnutzung und der Ankunftsrate von SUs ist in Abbildung 14 dargestellt, wobei die Ankunftsrate als x-Achse dient. Aus der Abbildung ist ersichtlich, dass die Spektrumnutzung mit der Zunahme der Ankunftsrate zunimmt. Wenn = = , Die Spektrumnutzungsrate beträgt bis zu 92,3 %. Mit der Zunahme der Ankunftsrate von SUs wird das ungenutzte Spektrum vollständig genutzt. Wenn die Spektrumnutzung allmählich gesättigt wird, nimmt der Gradient ab. Darüber hinaus ist es nicht schwer zu erkennen, dass der Gradient allmählich abnimmt. Darüber hinaus ist es nicht schwer zu erkennen, dass eine Erhöhung der Warteschlangenkapazität die sekundäre Netzwerkkapazität erhöhen kann, sodass die Spektrumauslastung verbessert. Die Tankauslastung wird entsprechend

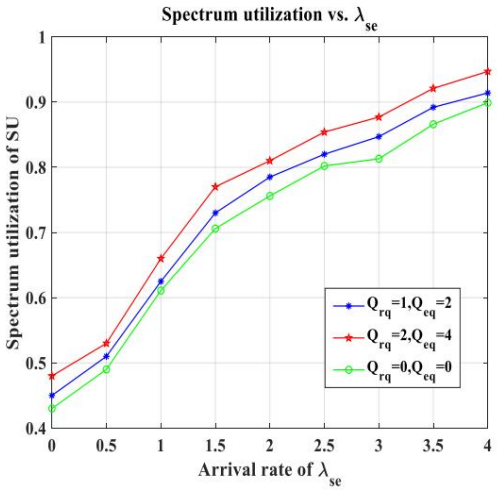


Abbildung 14. Spektrumnutzung vs. SUs.

Abbildung 14. Spektrumnutzung vs. SUs.

Fall 3: Blockierungswahrscheinlichkeit vs. Ankunftsrate von Dekubitus

Fall 3: Blockierungswahrscheinlichkeit vs. Ankunftsrate von Dekubitus

Abbildung 15 zeigt die Beziehung zwischen der Blockierungswahrscheinlichkeit von SUs und der Ankunftsrate von PUs. Im Allgemeinen steigt die Blockierungswahrscheinlichkeit von SUs mit der Rate der PUs. Die freien Kanäle werden voll ausgelastet, wenn die Anzahl der PUs im System steigt aufgrund der absoluten Datenzugriffspriorität, was zu einer Erhöhung der Blockierungswahrscheinlichkeit von SUs führt. Da SUr eine höhere Priorität als SUe hat , ist die Wahrscheinlichkeit von SUs, die Blockierungswahrscheinlichkeit von Da eine höhere Priorität als SUe hat, ist die Blockierungswahrscheinlichkeit des Datenzugriffs geringer als bei SUe . des Datenzugriffs ist geringer als der .

Abbildung 15 zeigt die Beziehung zwischen der Blockierungswahrscheinlichkeit von SUs und der Ankunftsrate von PUs. Im Allgemeinen steigt die Blockierungswahrscheinlichkeit von SUs mit der Rate der PUs. Die freien Kanäle werden voll ausgelastet, wenn die Anzahl der PUs im System steigt aufgrund der absoluten Datenzugriffspriorität, was zu einer Erhöhung der Blockierungswahrscheinlichkeit von SUs führt. Da SUr eine höhere Priorität als SUe hat, ist die Wahrscheinlichkeit von SUs, die Blockierungswahrscheinlichkeit von Da eine höhere Priorität als SUe hat, ist die Blockierungswahrscheinlichkeit des Datenzugriffs geringer als bei SUe . des Datenzugriffs ist geringer als der .

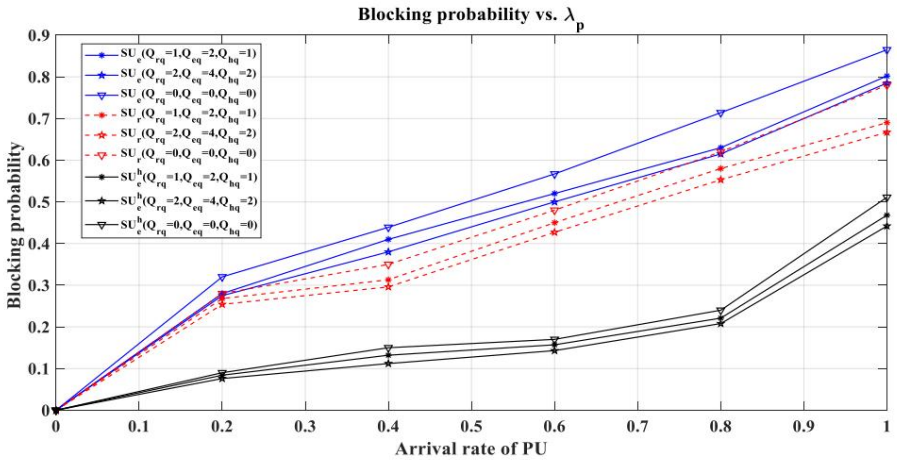


Abbildung 15. Blockierungswahrscheinlichkeit vs. SUs.

Abbildung 15. Blockierungswahrscheinlichkeit vs. SUs.

Es scheint, dass beide Kurven relativ stabil sind und der Unterschied in der Blockierungswahrscheinlichkeit nicht mehr als 0,1 beträgt . Dies ist ein Beispiel dafür, dass die vorgeschlagene Ps-rq-Strategie Fähigkeit beträgt nicht mehr als 0,1. Dies veranschaulicht die Tatsache, dass die vorgeschlagene Ps-rq-Strategie dazu führen kann, dass SUe und SUr eine stabile Leistung aufweisen, indem die Blockierungswahrscheinlichkeit verringert wird .

fähigkeit von Benutzern mit hoher Priorität SUh

& Die Einführung differenzierter Dienste bietet eine ermöglicht eine gerechte Kompromissplanung, ohne dass Benutzer mit niedriger Priorität überlastet werd Wahrscheinlichkeit von Benutzern mit hoher Priorität. Die Einführung differenzierter Dienste, Nie gerechte Kompromissplanung, ohne dass es zu einer Überlastung kommt Abhängigkeit von der Zunahme der Ankunftsrate von PUs. Darüber hinaus verdeutlicht dieses Diagramm auch, dass das Spektrum im Sekund Zunahme der Ankunftsrate von PUs. Darüber hinaus verdeutlicht dieses Diagramm auch, dass die Blockierungswahrscheinlichkeit von SU SUe nähert sich allmählich 1.

steigt mit der Zunahme der Ankunftsrate von PUs, was zu einer Blockierungswahrscheinlichkeit von führt.

Durch Vergleichen der Blockierungswahrscheinlichkeit von SUe und SUr bei unterschiedlichen Warteschlangenkapazitäten und

schrittweiser Annäherung an 1 können wir feststellen, dass eine größere Kapazität die Reduzierung der Blockierungsw

Und

Durch Vergleichen der Blockierungswahrscheinlichkeit bei unterschiedlichen Warteschlangenkapazitäten

können wir feststellen, dass eine größere Kapazität die Blockierungswahrscheinlichkeit verbessert und verringert.

Beispielsweise verringert sich die Blockierungswahrscheinlichkeit um 10,82 %, wenn die , wenn die Ankunftsrate von *PU* ist ü = Warteschlangenkapazität = = 4 im Vergleich zu

Machine Translated by Google

Beispielsweise wird für SUr , **wenn die Ankunftsrate von** PU ÿp = 0,6 **ist , die Blockierungswahrscheinlichkeit um 10,82 % reduziert, wenn die Warteschlangenkapazität** Qrq = **2,** Qeq = **4 ist, im Vergleich zu** Qrq = **0,** Qeq = **0. Und für** SUe , **wenn** ÿp = **0,6, verringert sich die Blockierwahrscheinlichkeit um 11,04 % unter dem gleichen Wert. Dies erklärt, dass das System Dienstleistungen für mehr bereitstellen kann Ströme mit der Erhöhung der Warteschlangenkapazität, was mehr Möglichkeiten für die spekulativer Zugriff auf SUs. Im Vergleich dazu ist zu beobachten, dass der Anstieg der Warteschlange** Kapazität kann die Leistung von SUe **deutlich verbessern . Der Grund dafür ist, dass eine Rückkopplung Schleife wird für den Dienst von** SUe **in der vorgeschlagenen Ps-rq-Strategie und einer größeren Warteschlangenkapazität** eingestellt **wird im System-Setup installiert.**

Fall 4: Wahrscheinlichkeit einer erzwungenen Beendigung vs. Ankunftsrate von PUs

Abbildung 16 zeigt die Beziehung zwischen der erzwungenen Beendigungswahrscheinlichkeit von SUs und die Ankunftsrate der PUs. Wie in der Abbildung dargestellt, ist die Wahrscheinlichkeit einer erzwungenen Beendigung von SUs steigt mit der Ankunftsrate von PUs, die auch durch den Datenzugriff bestimmt wird. Priorität der PUs. Da für die unterbrochene SUh eine reservierte Warteschlange eingerichtet ist , **die erzwungene Kündigung Wahrscheinlichkeit von** SUh **ist der niedrigste Wert in Abbildung 16, also der maximale Differenzwert,** ÿp = **0,8, wenn** Qrq = **1,** Qeq = **2,** Qhq = **1, die erzwungene Beendigungswahrscheinlichkeit von** SUh **ist 29,23 % niedriger als bei** SUe **und 42,5 % niedriger als bei** SUr . **Aufgrund der Einstellung der Rückkopplungsschleife für** SUe , SUr **hat eine höhere Wahrscheinlichkeit für eine erzwungene Beendigung als** SUe . **Daher kann die unterbrochene** SUe **in die Warteschlange zurückkehren und wird nicht gezwungen, beendet, bis die Warteschlange voll ist.**

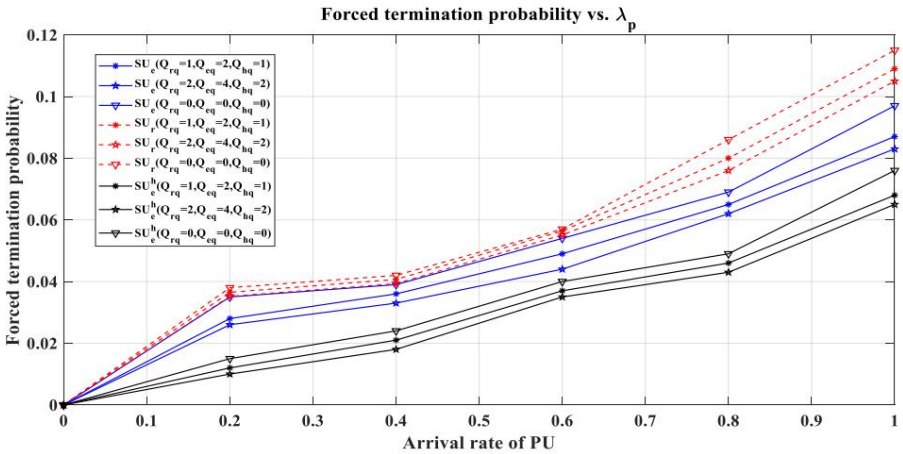


Abbildung 16. Wahrscheinlichkeit der erzwungenen Beendigung von SUs.

Echtzeitdienste reagieren empfindlicher auf Verzögerungen und Paketverluste durch leichtes Jitter verursachen erhebliche Störungen der Echtzeitsprachqualität. Daher wird für SUs keine Zittern beeinträchtigen die Echtzeit-Sprach-/Bildqualität nicht ernsthaft. Daher gibt es keine Rückkopplungen Rückkopplungsschleife verwendet . **Die Gesamtkanalbelegungsrate kann auf diese Weise reduziert werden, aber Die Rückschleife wird für verwendet. Die Gesamtkanalbelegungsrate kann dabei auf Kosten einer Erhöhung der Wahrscheinlichkeit einer erzwungenen Beendigung von** SUr reduziert werden . **Insgesamt betrachtet ist Weg, aber auf Kosten der Erhöhung der Wahrscheinlichkeit der erzwungenen Beendigung von . Die vorgeschlagene Ps-rq-Strategie kann die Serviceprioritäten heterogener Dienste differenzieren differenzieren und innerhalb einer fairen Planung bieten. Auf dieser Grundlage kann die Servicequalität der hochpriorisInsgesamt kann die vorgeschlagene Ps-rq-Strategie die Serviceprioritäten heterogener Dienste** Dienstes SUh , **die durch PUs unterbrochen wird, kann verbessert werden. Dienstleistungen und sorgen für eine faire Planung. Auf dieser Grundlage wird die Servicequalität des elastischen Der hochpriorisierte Dienst der unterbrochenen SUs kann ebenfalls verbessert werden.**

Um die Überlegenheit der vorgeschlagenen Strategie bei der Verbesserung der vorgeschlagenen Strategie bei der Verbesserung der Ps-rq-Strategie hinsichtlich der Wahrscheinlichkeit Leistung von SUs, ein Vergleich zwischen der Ps-rq-Strategie und der Nicht-Ps- Vergleichsergebnisse rq-Strategie hinsichtlich der Wahrscheinlichkeit einer erzwungenen Beendigung, wie in Abbildung 17 einer erzwungenen Beendigung weiter zu demonstrieren, wie in Abbildung 17 dargestellt. Die dargestellt. Die Kombination aus SUe **und** SUr ist aus der Abbildung ersichtlich . **Wenn die Strategie nicht angewendet w Vergleichsergebnisse traten alle unter den Anfangsbedingungen = 2 und = 1 auf. Die Wahrscheinlichkeit einer erzwungenen Beendigung für SUs steigt erheblich. Die Leistungsanalyse** Aus der Abbildung ist ersichtlich, ob für den Fall, dass die Strategie nicht angenommen wird, dies spezifisch **wurde alle unter den Anfangsbedingungen** Qeq = **2 und** Qrq = **1 durchgeführt. Wie** von SUh **für die in diesem Papier vorgeschlagene Strategie ist, gibt es keinen Vergleich von** Die Wahrscheinlichkeit einer erzwungenen Beendigung für SUs steigt deutlich an. Die Leistungsanalyse zeigt, dass der **die erzwungene Beendigungswahrscheinlichkeit von SUs, wenn Ps-rq nicht angewendet wird. Die Abbildung zeigt** Anstieg der Wahrscheinlichkeit einer Zwangsbeendigung für SUe größer ist als für SUr . **von ist spezifisch für die in diesem Papier vorgeschlagene Strategie, daher gibt es keinen Vergleich von in Abwesenheit von Ps-rq. Insbesondere wenn die Ankunftsrate von PU gering ist,** hat SUe **mehr die Wahrscheinlichkeit einer erzwungenen Beendigung bei Nichteinführung von Ps-rq. Die Abbildung zeigt, dass der Anstieg der Wahrscheinlichkeit einer erzwungenen Beendigung bei größer ist als bei**

in Abwesenheit von Ps-rq. Insbesondere wenn die Ankunftsrate von PU gering ist, hat

mehr Möglichkeiten, auf den Kanal zuzugreifen. Aus der vorgeschlagenen Strategie geht hervor, dass das hervorragende Ergebnis auf die verbesserte Leistung von zurückzuführen ist .

Machine Translated by Google

Die Wahrscheinlichkeit einer Zwangsbeendigung für SUs steigt deutlich an. Die Leistungsanalyse von ist spezifisch für die in diesem Dokument vorgeschlagene Strategie; daher gibt es keinen Vergleich der Wahrscheinlichkeit einer Zwangsbeendigung von ohne Ps-rq. Die Abbildung zeigt, dass der Anstieg der Wahrscheinlichkeit einer Zwangsbeendigung für größer ist als für in Abwesenheit von Ps-rq. Insbesondere wenn die Ankunftsrate von PU gering ist, hat mehr Möglichkeiten, auf den Kanal zuzugreifen. Aus der vorgeschlagenen Strategie geht hervor, dass die Möglichkeiten, auf den Kanal zuzugreifen. Aus der vorgeschlagenen Strategie geht hervor, dass das hervorragende Ergebnis auf die verbesserte Leistung von SUn **zurückzuführen ist.** e^{*}.

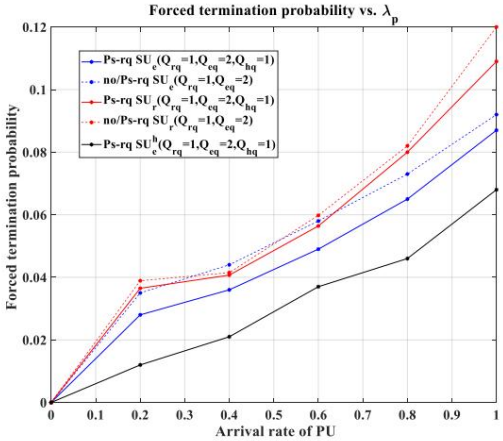


Abbildung 17. Vergleich der Wahrscheinlichkeit einer erzwungenen Beendigung mit und ohne Ps-rq.

Abbildung 17. Vergleich der Wahrscheinlichkeit einer erzwungenen Beendigung mit und ohne Ps-rq.

6. Schlussfolgerungen

Diese Studie entwirft und schlägt eine Prioritätsplanungsstrategie mit reservierter Warteschlange (Ps-rq-Strategie) für CRNs vor. Diese Strategie verwendet ein Dienstklassifizierungsschema, das statische Faktoren (Nachrichtentyp) und dynamische Faktoren (Informationsgültigkeit, Nachrichtenkorrelation und Nachrichtengröße) kombiniert, wodurch die elastischen Dienste mit hoher Priorität auf der Grundlage der heterogenen Dienstklassifizierung in sekundären Diensten unterschieden werden können. Durch die Einrichtung einer neuen reservierten Warteschlange für wichtige elastische Dienste kann die vorgeschlagene Strategie die Dienstpriorität hochprioritärer elastischer Dienste verbessern, die durch PUs unterbrochen werden. Darüber hinaus wird die Takagi-Sugeno-Fuzzy-Steuerung verwendet, um elastische Dienste zu priorisieren, u Auf dieser Grundlage wird ein proportionales Kanalzuweisungsschema mit Priorität bereitgestellt, das mit Kanalzusammenstellungstechnologie und Spektrumanpassungstechnologie kombiniert wird und auf der Ps-rq-Strategie basiert. Durch die Aufteilung der Benutzeraktivitäten in CRNs in vier Ereignisse, darunter PU-Ankunft, PU-Abgang, SU-Ankunft und SU-Abgang, wird der Ressourcenfluss im sekundären Netzwerk untersucht und der dynamische Kanalzugriffsprozess basieren Durch die Einrichtung des CTMC wird der Ressourcenflussprozess im sekundären Netzwerk auf den hochdimensionalen Zustandsraum abgebildet und alle möglichen Übergangszustände, Übergangsraten und Übergangsbedingungen, die durch die vier Arten von Ereignissen ausgelöst werden, die von einem allgemeinen Zustand ausgehen, werden abgeleitet. Nach dem Herabstufen der hochdimensionalen Markierungen kann die stationäre Lösung des Systems durch Erfassen der Übergangsratenmatrix erhalten werden und außerdem werden eine Reihe von Leistungsbewertungsindizes wie sekundäre Netzwerkkapazität, Spektrumnutzung, Blockierungswahrscheinlich

Darüber hinaus vergleichen wir die vorgeschlagene Ps-rq-Strategie mit vier Kanalzusammenstellungsschemata, um ihre Eignung für bestimmte Szenarien mit der Übertragung elastischen Datenverkehrs zu demonstrieren. Durch Kombination des Verfahrens zur Dimensionsreduzierung hochdimensionaler Markow-Ketten simulieren wir die analytischen Ergebnisse der Leistungsindizes mithilfe von MATLAB numerisch und stellen vier Fallstudien bereit. Die Simulationsergebnisse zeigen, dass in derselben Dimension mit zunehmender Pufferkapazität die Netzwerkkapazität von SUe um **4,8 % zunimmt und die Blockierungswahrscheinlichkeit um 11,04 % sinkt, während die Netzwerkkapazität von** SUr **um 9,43 % zunimmt und die Blockierungswahrscheinlichkeit um 10,82 % sinkt. Und die Wahrscheinlichkeit einer erzwungenen Beendigung der in diesem Dokument behandelten elastischen Dienste mit hoher Priorität ist 29,23 % niedriger als die von** SUe **und 42,5 % niedriger als die von** SUr . **Die numerischen Ergebnisse zeigen, dass die vorgeschlagene Strategie dazu beitragen kann, die Servicepriorität von elastischen Diensten mit hoher Priorität , die durch PUs in CRNs unterbrochen werden, zu verbessern und die Wahrscheinlichkeit einer erzwungenen Beendigung durch faire Planung zu verringern.**

Machine Translated by Google

Beiträge der Autoren: QX recherchierte die Literatur, konzipierte die Studienkonzepte, stellte die mathematischen Modelle bereit, entwarf den Algorithmus, führte numerische Simulationen durch, übernahm die Vorbereitung des **ursprünglichen Entwurfs und redigierte das Manuskript; SL konzipierte die Studienkonzepte, verbesserte die systematische Forschungs- und Analysemethodik und überwachte die Fertigstellung der Verfeinerung des Papiers; JG und YH gaben wertvolle Vorschläge zur Überarbeitung und überprüften die Formelableitung und die englische Grammatik. Alle Autoren haben die veröffentlichte Version**

Finanzierung: Diese Forschung wurde zum Teil von der National Natural Science Foundation of China (61663024) und dem Doktorandenfondsprojekt für Naturwissenschaften der Xi'an Aeronautical University (2022KY0201

Erklärung zur Datenverfügbarkeit: Die diesem Artikel zugrunde liegenden Daten stammen aus zuvor veröffentlichten Studien und Datensätzen, die zitiert wurden.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Verweise

1. **Federal Communications Commission.** Bericht der Spectrum Policy Task Force. ET Docket; **Nr. 02-135; Federal Communications Auftrag: Washington, DC, USA, 2002.**
2. **Huang, XL; Li, YX; Gao, Y. Q-Learning-basierter Spektrumzugriff für Multimediaübertragung über kognitive Funknetzwerke.** IEEE Trans. Cogn. Commun. Netw. **2021, 7, 110–119. [CrossRef]**
3. **Srikar, D.; Nella, A.; Mamidi, R. Ein neuartiges integriertes UWB-Sensor- und 8-Element-MIMO-Kommunikationssystem mit kognitiver Radioantenne .** Electronics **2023,** 12, **330. [CrossRef]**
4. **Muzaffar, MU; Sharqi, R. Eine Überprüfung der Spektrumerfassung in modernen kognitiven Funknetzwerken.** Telekommunikationssystem-Modell. Anal. Des. Manag. **2024, 2, 85. [CrossRef]**
5. **Ahmad, W.; Radzi, N.; Samidi, FS 5G-Technologie: Auf dem Weg zur dynamischen Frequenzteilung mithilfe kognitiver Funknetzwerke.** IEEE Access **2020, 8, 14460–14488. [CrossRef]**
6. **Azaly, N.; Badran, EF Leistungssteigerung des dynamischen Spektrumzugriffs durch Kanalreservierung für Cognitive Radio Netzwerke.** Wirel. Pers. Commun. **2021,** 118, **2867–2883. [CrossRef]**
7. **Lei, J.** Kanalaggregation und -fragmentierung für Verkehrsflüsse; **Springer: Berlin/Heidelberg, Deutschland, 2020.**
8. **Liu, X.; Jia, M. Intelligente Spektrumressourcenzuweisung basierend auf gemeinsamer Optimierung im heterogenen kognitiven Radio.** IEEE Trans. Emerg. Top. Comput. Intell. **2020, 4, 5–12. [CrossRef]**
9. **Li, SP; Xu, QY; Jaafar, G. Modellierung und Leistungsanalyse der Kanalzusammenstellung basierend auf der Ps-rc-Strategie mit Prioritätswarteschlangen in CRNs.** Wirel. Commun. Mob. Comput. **2022,** 17, **6384261. [CrossRef]**
10. **Ma, L.; Xu, Y.; Fu, Y. Ein Ressourcenplanungsschema für die Spektrumaggregation in heterogenen Systemen auf Basis kognitiven Radios Netzwerke.** China Commun. **2015,** 12, **100–111.**
11. **Jiao, L.; Li, FY; Pla, V. Modellierung und Leistungsanalyse der Kanalzusammenstellung in mehrkanaligen kognitiven Funknetzwerken mit Spektrumanpassung.** IEEE Trans. Veh. Technol. **2012,** 61, **2686–2697. [CrossRef]**
12. **Balapuwaduge, I.; Jiao, L.; Pla, V. Kanalzusammenstellung mit prioritätsbasierten Warteschlangen in kognitiven Funknetzwerken: Strategien und Leistungsbewertung.** IEEE Trans. Wirel. Commun. **2014,** 13, **630–645. [CrossRef]**
13. **Esenogho, E.; Srivastava, VM Zwei heterogene Kanalzusammenstellungsstrategien in kognitiven Funknetzwerken: Eine Performance Analyse.** Int. J. Eng. Technol. Innov. **2017, 7, 98–116.**
14. **Zhang, W.; Sun, Y.; Deng, L. Dynamische Spektrumzuweisung für heterogene kognitive Funknetzwerke mit mehreren Kanälen.** IEEE Syst. J. **2019,** 13, **53–64. [CrossRef]**
15. **Zhu, X. Analyse des kognitiven Funkspektrumzugriffs mit optimaler Kanalreservierung.** IEEE Commun. Lett. **2007,** 11, **304–306. [Querverweis]**
16. **Esenogho, E.; Mambou, EN Integration des Warteschlangenregimes in kognitive Radiokanalaggregationsrichtlinien: Eine Performance Auswertung .** J. Commun. **2017, 1–5.**
17. **Esenogho, E.; Srivastava, VM Kanalzusammenstellungsstrategie in kognitiven Funknetzwerken: Ein Warteschlangen-basierter Ansatz.** Int. J. Commun. Antenna Propag. IRCEAP **2017, 7, 31–47. [CrossRef]**
18. **Xiao, X.; Zeng, F.; Hu, Z. Dynamisches, flussadaptives Spektrum-Leasing mit Kanalaggregation in kognitiven Funknetzwerken.** Sensoren **2020,** 20, **3800. [CrossRef] [PubMed]**
19. **El-Toukhy, AT; Arslan, H. Verbesserung der Leistung von SUs mit niedriger Priorität durch reservierte Kanäle in CRN.** IEEE Wirel. 2020 , **9, 513–517. [CrossRef]**
20. **Bayrakdar, ME; Calhan, A. Verbesserung der Frequenzübergabenutzung für priorisierte Cognitive Radio-Benutzer durch Ausnutzung Kanalbindung mit Hungerminderung.** Int. J. Electron. Commun. **2017,** 71, **181–191. [CrossRef]**
21. **Suganthi, N.; Meenakshi, S. Ein effizienter Planungsalgorithmus mit Warteschlangensystem zur Minimierung des Mangels an nicht in Echtzeit verfügbaren Sekundärnutzer im Cognitive Radio Network.** Clust. Comput. **2018,** 25, **1–12. [CrossRef]**
22. **Iftikhar, M.; Mathkour, H.; Imran, M. Ein neuartiges Framework für ein G/M/1-Warteschlangensystem basierend auf einem Scheduling-cum-Polling-Mechanismus zur Analyse mehrerer Klassen von selbstähnlichem und LRD-Verkehr.** Wirel. Netw. **2016,** 22, **1269–1284. [CrossRef]**

Machine Translated by Google

23. Khan, M.; Ahmad, J.; Saddik, AE Drone-HAT: Hybrid Attention Transformer zur Erkennung komplexer Aktionen in Drohnenüberwachungsvideos . In Proceedings der IEEE/CVF-Konferenz zu Computer Vision and Pattern Recognition (CVPR) Workshops, Seattle, WA, USA, 19.–21. Juni 2024; S. 4713–4722.

24. Taherkhani, N.; Pierre, S. Priorisierung und Planung von Nachrichten zur Überlastungskontrolle in Ad-hoc-Netzwerken für Fahrzeuge. Comput. Netw. **2016,** 108, **15–28.** **[CrossRef]**

25. Han, W.; Li, J.; Tian, Z. Dynamische Sensorstrategien für eine effiziente Spektrumnutzung in kognitiven Funknetzwerken. IEEE Trans. Wirel. Commun. **2011,** 10, **3644–3655.** **[CrossRef]**

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Auf Deep Reinforcement Learning basierendes Multipath Routing für LEO Megakonstellationsnetzwerke

Chi Han [1], Wei Xiong [1,2,*] und Ronghuan Yu [1,2]

[1] Nationales Schlüssellabor für Weltraumzielerkennung, Space Engineering University, Peking 101400, China; hanchi@hgd.edu.cn (CH); yrh1983@163.com (RY)
[2] School of Space Information, Space Engineering University, Peking 101400, China * Korrespondenz: xiongwei@hgd.edu.cn

Zusammenfassung: Der Ausbau von Megakonstellationsnetzwerken (MCN) stellt eine vielversprechende Lösung für die Erreichung einer globalen Internetabdeckung dar. Um der wachsenden Nachfrage nach Satellitendiensten gerecht zu werden, ermöglicht Multipath-Routing die gleichzeitige Einrichtung mehrerer Übertragungspfade, wodurch die parallele Übertragung von Datenströmen ermöglicht wird. Dennoch stellen die Mobilität von Satelliten und zeitabhängige Verbindungszustände eine Herausforderung für die Ermittlung optimaler Pfade und die Verkehrsplanung beim Multipath-Routing dar. Angesichts der mangelnden Flexibilität traditioneller statischer Routing-Algorithmen auf Basis von Deep Reinforcement Learning (DRL) im Umgang mit zeitabhängigen Konstellationstopologien wird DRL-basiertes Multipath-Routing (DMR) auf Basis eines Graph Neural Network (GNN) als Mittel zur Verbesserung der Übertragungsleistung von MCNs vorgeschlagen. DMR entkoppelt das stochastische Optimierungsproblem des Multipath-Routings unter Verkehrs- und Bandbreitenbeschränkungen in zwei Teilprobleme: die Ermittlung des Multipath-Routings und die Planung des Multipath-Verkehrs. Zunächst wird der auf der Mindestanzahl von Hops basierende Multipath-Routenermittlungsalgo Zweitens wird das GNN-basierte Multipath Traffic Scheduling Scheme (GMTS) als Mittel zur dynamischen Planung des Verkehrs auf jedem verfügbaren Pfad für jeden Datenstrom vorgeschlagen, basierend auf den Statusinformationen der ISLs und der Verkehrsnachfrage. Simulationsergebnisse zeigen, dass das vorgeschlagene Schema ohne wiederholtes Training auf Konstellationen mit unterschiedlichen Konfigurationen skaliert werden kann und den Durchsatz, die Abschlussrate und die Verzögerung im Vergleich **zum Shortest Path First-Algorithmus (SPF) um 42,64 %, 17,39 % bzw. 3,66 % verbessert**

Schlüsselwörter: Satellitennetzwerk; Mehrwege-Routing; Deep Reinforcement Learning; Verkehrsplanung; Hop-Anzahl

check for updates

Zitat: Han, C.; Xiong, W.; Yu, R. Auf Deep Reinforcement Learning basierendes Multipath Routing für LEO-Megakonstellationsnetzwerke. Electronics **2024,** 13, **3054. https:// doi.org/ 10.3390/electronics13153054**

Wissenschaftlicher Herausgeber: Stefano Scanzio

Empfangen: 12. Juli 2024
Überarbeitet: 25. Juli 2024
Akzeptiert: 31. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.

Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

Die Eigenschaften von Megakonstellationsnetzwerken (MCNs), wie z. B. große Flächenabdeckung , geringe Latenz und hohe Bandbreite, verleihen ihnen einzigartige Vorteile bei der Kommunikation in abgelegenen Gebieten und bei verzögerungsempfindlichen Diensten (z. B. Notfallkommunikation und Katastrophenüberwachung) [1]. Die oben genannten Vorteile haben zu einer beschleunigten Entwicklung von MCNs geführt, von denen erwartet wird, dass sie eine Vielzahl von Diensten erfüllen, darunter Kommunikation, Fernerkundung, Navigation und Positionierung. Darüber hinaus sind MCNs im Begriff, ein zentraler Bestandteil der Infrastruktur des integrierten Weltraum-Luft-Boden-Netzwerks (Space-Air-Ground Integrated Network, SAGIN) zu werden

Um der wachsenden Nachfrage nach Satellitenkommunikationsdiensten gerecht zu werden, ist Multipath-Routing ein vielversprechender Ansatz. Multipath-Routing ermöglicht die gleichzeitige Einrichtung mehrerer Übertragungspfade zwischen den Quell- und Zielknoten für jeden Datenstrom , was den Netzwerkdurchsatz erhöhen kann [5]. Im Gegensatz zum Single-Path-Routing ermöglicht Multipath-Routing die optimale Nutzung der verfügbaren Bandbreite, wodurch Überlastungen verringert, der Durchsatz verbessert und die Fehlertoleranz erhöht werden. Dies wird erreicht, indem ein einzelner Dienst über mehrere Pfade

Machine Translated by Google

Multipath-Routing erhöht die Widerstandsfähigkeit ressourcenbeschränkter Satellitennetzwerke gegenüber Knoten- oder Verbindungsausfällen. Die hohe Dynamik von Satellitennetzwerktopologien und Übertragungsverbindungszuständen stellt jedoch Herausforderungen für die Multipath-Routing-Erkennung und die Multipath-Verkehrsplanung dar **[6–8].**

Die Erkennung von Mehrwege-Routing ist die erste Herausforderung, die sich aus der zeitlichen Veränderung der Satellitentopologie und des Satellitenverkehrs ergibt. Faktoren wie Änderungen der Satellitenlage und komplexe elektromagnetische Umgebungen können dazu führen, dass die Übertragungsraten von Laserverbindungen abnehmen oder sogar ausfallen. Infolgedessen können Verbindungsfehler in MCNs häufig auftreten und nicht wie in terrestrischen Netzwerken ein Ausnahmeereignis darstellen. Das häufige Auftreten von Verbindungsüberschreitungen oder -ausfällen beeinträchtigt die Fähigkeit des Übertragungsmechanismus, Routen rechtzeitig zu erkennen und aufrechtzuerhalten, wie dies in terrestrischen Netzwerken der Fall ist. Dies führt zu einem Anstieg des mit der Übertragung und Routensteuerung verbundenen Overheads . Folglich sind topologiebasierte Ad-hoc-Routingprotokolle wie ADOV [9] in MCNs nicht optimal. Darüber hinaus werden aufgrund der Mobilitäts- und Energiebeschränkungen von LEO-Satelliten bestehende Verbindungen häufig nicht stabil aufrechterhalten. Topologische Änderungen im Pfaderkennungsprozess können dazu führen, dass erkannte Pfade schnell ablaufen und die entsprechenden Pfade zum Zeitpunkt der Datenübertragung nicht mehr existieren. Die oben genannten Faktoren erfordern eine Multipath-Routing-Erkennung, um Verbindungsfehler rechtzeitig erkennen und mehrere alternative Pfade sinnvoll festlegen zu können. Einige Systeme setzen derzeit die Pfadberechnung im Bodennetzkontrollzentrum (NCC) ein [10]. Das NCC verbessert die Übertragungsleistung des Netzwerks, indem es die Verteilung des terrestrischen Benutzerverkehrs und die Verbindungsauslastung umfassend berücksichtigt und die Verbindungsredundanz, die durch die Maschentopologie des Satellitennetzwerks entsteht, voll ausnutzt, um mehrere parallele Übertragungspfade für jeden Datenfluss gleichzeitig zu planen.

Die zweite Herausforderung besteht in der Mehrwegeverkehrsplanung in MCNs. Komplexe elektromagnetische Umgebungen, Änderungen der Satellitenlage und andere Faktoren führen zu Schwankungen der Verbindungsqualität . Daher ist die traditionelle Mehrwegeverkehrsplanung auf Basis statischer Routen in MCNs nicht mehr anwendbar. Darüber hinaus machen die zeitlich variierende Natur des terrestrischen Benutzerverkehrs und die Schwankungen der Verbindungsqualität das Satellitennetzwerk unfähig, Verkehrsmuster genau zu identifizieren, und einer feinkörnigen Verkehrsplanung zwischen mehreren Pfaden fehlen häufig wirksame Informationen. Deep Reinforcement Learning (DRL) hat erhebliche Vorteile bei der Beschreibung räumlich-zeitlicher Merkmale, Flussentscheidungen und anderen verwandten Bereichen gezeigt. Eine beträchtliche Anzahl von Wissenschaftlern hat sich für die Einführung von DRL zum Zweck der adaptiven Mehrwegeverkehrsplanung entschieden, wie [11–14] zeigt. Die zeitlich variierenden Satellitennetzwerke und die schwankende Qualität der Inter-Satelliten-Verbindungen (ISLs) erfordern jedoch die Aggregation mehrerer Zeitscheiben für die Entscheidungsfindung in Modellen, die DRL verwenden. Mit zunehmender Konstellationsgröße vergrößert sich die Dimension des Modells dramatisch, was zu einem deutlichen Anstieg des Rechenaufwands und einer Verringerung der Modellskalierbarkeit führt . Graph-Neural-Networks (GNNs) sind in der Lage, Merkmale und Muster in graphisch strukturierten Daten zu erkennen, relationale Schlussfolgerungen und kombinatorische Generalisierungen auf graphisch strukturierten Daten durchzuführen **[15,16] und sie könne**

In jüngster Zeit hat DRL erhebliches Potenzial beim Erlernen zeitlicher und räumlicher Verkehrsmerkmale zum Zweck der Entscheidungsfindung im Bereich der Routenführung gezeigt. Forscher haben DRL zum Zweck der adaptiven Verkehrszuweisung eingesetzt, mit dem Ziel, die statische Verkehrssegmentierung zu ersetzen. DRL-basierte Modelle erfordern jedoch die Fähigkeit, kontinuierliche topologische Schnappschüsse zusammenzufassen, um Entscheidungen in einer dynamischen Satellitenumgebung treffen zu können . Satellitennetzwerke haben hochdimensionale und spärliche Zustandsräume, und die Dimensionalität von DRL-Modellen nimmt mit der Größe der Konstellation und dem Niveau der Verkehrsnachfrage dramatisch zu, was wiederum zu einer schlechten Modellskalierbarkeit führt. Graph- Neural-Networks (GNN) erleichtern die Inferenz relationaler Daten und die Verallgemeinerung kombinatorischer Strukturen innerhalb graphisch strukturierter Informationen. Dies ermöglicht die Anwendung des DRL-Modells auf Satellitenkonstellationen unterschiedlicher Größe ohne die Notwendigkeit zusätzlicher Modifikationen. Im Lichte der oben genannten Analyse schlagen wir DRL-

Machine Translated by Google

Das Multipath-Routing-Problem wird zunächst als stochastisches Optimierungsproblem modelliert, mit dem Ziel, die Netzwerkleistung unter Verkehrs- und Bandbreitenbeschränkungen zu maximieren . Um die optimale Lösung zu erhalten, wird das ursprüngliche Problem in zwei Teilprobleme aufgespalten: Multipath-Routing-Erkennung und Multipath-Verkehrsplanung. In Bezug auf das anfängliche Teilproblem wird vorgeschlagen, die redundante Topologie des Satellitennetzwerks durch die Einführung eines auf minimaler Hop-Anzahl basierenden Multipath-Routing **-Erkennungsalgorithmus (MHMRD) voll auszunutzen .** MHMRD ist darauf ausgelegt, eine Anzahl verfügbarer **Pfade zwischen allen Quell- und Zielknoten zu berechnen. Auf dieser Grundlage wird das zweite Teilproblem durch die Entwicklung einer GNN-basierten Multipath-Verkehrsplanung (GMTS) angegangen. GMTS modelliert den Multipath-Verkehrsplanungsprozess als Markov-Entscheidungsprozess (MDP), der den Verkehr auf jedem verfügbaren Pfad für jeden Datenstrom dynamisch plant, basierend auf Informationen zum Verbindungsstatus des Satellitennetzwerks und der Verkehrsnachfrage. Simulationsergebnisse auf Iridium und OneWeb zeigen, dass das vorgeschlagene Schema nicht auf Umschulung für Konstellationen unterschiedlicher Größe angewiesen ist. Darüber hinaus übertrafen Netzwerkdurchsatz, durchschnittliche Verzögerung und Flussabschlussrate die Werte des Basisschemas.**

Die wichtigsten Beiträge dieses Papiers lassen sich wie folgt zusammenfassen:

- **Das Multipath-Routing-Problem in Satellitennetzwerken wird als stochastisches Optimierungsproblem unter Verkehrs- und Bandbreitenbeschränkungen modelliert. Ziel ist es, die Netzwerkleistung zu maximieren. Das Problem wird in ein Teilproblem der Multipath-Routing-Erkennung und ein Teilproblem der Multipath-Verkehrsplanung aufgespalten.**
- **Um das Teilproblem der Multipath-Routing-Erkennung zu lösen, wurde der MHMRD-Algorithmus als Mittel zur Planung mehrerer Pfade mit minimaler Hop-Anzahl für jeden Datenfluss zwischen jedem Knotenpaar im Netzwerk basierend auf den Link- State-Informationen vorgeschlagen.**
- **Das Teilproblem der Multipath-Verkehrsplanung wird als MDP modelliert, wobei das GMTS- Schema vorgeschlagen wird, um den Verkehrsanteil auf jedem verfügbaren Pfad für jeden Datenstrom dynamisch zu planen. GMTS ist eine skalierbare Lösung, die auf verschiedene Konstellationen angewendet werden kann.**

Der Rest dieses Dokuments ist wie folgt gegliedert. Abschnitt 2 enthält einige damit zusammenhängende Arbeiten. In Abschnitt 3 stellen wir das Systemmodell und die Problemformulierung vor. In Abschnitt 4 wird DRL-basiertes Multipath-Routing beschrieben. Abschnitt 5 enthält die Simulation und die wichtigsten Ergebnisse. Abschließend fassen wir die Ergebnisse in Abschnitt 6 zusammen.

2. Verwandte Arbeiten

2.1. Multipath-Routing **In**

Satellitennetzwerken gibt es mehrere Pfade mit einer gleichen Anzahl von Hops von Ende zu Ende. Um die Zuverlässigkeit der Datenübertragung zu verbessern und die Netzwerklastausgleichsfähigkeit zu steigern , haben Forscher die Verwendung von Multipath-Routing in Satellitennetzwerken vorgeschlagen, bei dem jeder Fluss in mehrere Unterflüsse für die parallele Übertragung entlang verschiedener Pfade aufgeteilt wird. Um den Verkehr auf mehrere Pfade zu verteilen, führt die traditionelle Equal-Cost-Multipath-Methode (ECMP) eine statische Verkehrsaufteilung basierend auf Nachrichteninformationen durch.
Der ECMP-Ansatz berücksichtigt jedoch nicht die Einschränkungen der Netzwerkparameter wie Bandbreite und Verzögerung und ist bei hoher Netzwerklast anfällig für Überlastungen.
Um ein dynamisches Flussmanagement für Multipath-Routing zu erreichen, wurde eine auf Software-Defined Network (SDN) basierende Satellitennetzwerkarchitektur zur zentralen Topologiesteuerung und Verkehrskontrolle vorgeschlagen **[18–21]. In der SDN-basierten Netzwerkarchitektur sind die Datenebene und die Steuerebene getrennt, der SDN-Controller bestimmt das Routing basierend auf den Netzwerkparametern und QoS-Anforderungen und der Satellit ist für die Datenweiterleitung verantwortlich. In [22] versuchten die Autoren, Netzwerkcodierung und Multipath-Routing zu integrieren, um die Datenübertragungseffizienz von Satellitennetzwerken zu verbessern. Der zentralisierte Steuerungsansatz ist jedoch anfällig für Probleme mit hohen Latenzen. In [23]** verwendeten die Autoren einen verteilten Ansatz zur Routing-Steuerung, **um die Steuerungslatenz zu reduzieren und die Übertragungsstabilität zu verbessern . In [24] schlugen die Autoren netzwerkcodierungsbasiertes Multipath Cooperative Routing (NCMCR) vor. Um die Übertragungskapazität zu optimieren, muss jeder Datenfluss**

Machine Translated by Google

in NCMCR wird über mehrere disjunkte Verbindungen übertragen. Die oben genannten Mehrwege-Routing-Schemata berücksichtigen jedoch nicht die Verbindungsparameter der verschiedenen Weiterleitungspfade und können daher keine angemessene Verkehrsaufteilung adaptiv durchführen. In [1] entwickelten die Autoren ein adaptives Verkehrsausgleichsschema für mehrere zeitliche Pfade, indem sie spärliche und redundante Netzwerkcodierungsmechanismen kombinierten. Dieser Ansatz ermöglicht die Erreichung deterministischer Verzögerungsgarantien bei begrenzten Knotenressourcen und Mehrbenutzer- Wettbewerbsszenarien. Allerdings verwenden alle oben genannten Routing-Schemata statische Topologien, was sie daran hindert, sich an Satellitennetzwerke mit dynamisch wechselnden Topologien anzupassen.

2.2. Intelligentes Routing

Die zeitabhängige Satellitennetzwerkumgebung und die dynamischen interstellaren Verbindungen stellen Herausforderungen für das Satellitennetzwerk-Routing dar. Um die Servicequalität des Netzwerks unter den Bedingungen zunehmender Benutzer- und Konstellationsgrößen sicherzustellen, erfordert Multipath-Routing eine effizientere Pfadplanung und feinere Verkehrssteuerungsstrategien. Reinforcement Learning ist ein intelligentes Werkzeug, das die Entscheidungsfindung durch Interaktion mit der Umgebung unterstützt [25–27], und Wissenschaftler haben zahlreiche intelligente Routing-Schemata auf der Grundlage von Reinforcement Learning vorgeschlagen. In [25] entwickelten die Autoren ein überwachtes Deep-Learning- System für die Erstellung von Routing-Tabellen. In [28] verwendeten die Autoren graphenbasiertes Deep Learning in Satellitennetzwerken durch eine neuronale Netzwerkarchitektur namens Graph- Query Neural Network. Solche auf überwachtem Lernen basierenden Routing-Schemata sind jedoch durch ihre Unfähigkeit zur Generalisierung und Anpassung an verschiedene Konstellationen eingeschränkt. Um dieses Problem zu lösen, schlugen die Autoren von [29] ein Markov-Entscheidungsprozessmodell (MDP) für das Satellitennetzwerk-Routing vor, das Multi-Agent-Deep-Reinforcement-Learning (DRL) verwendet, um verschiedene Anforderungen an die Dienstgüte (QoS) zu erfüllen. Die feste Zielpolitik macht DRL-basierte Routing-Algorithmen jedoch unflexibel, da sie sich nicht an dynamisch ändernde Netzwerke anpassen können . Es ist offensichtlich, das Um dieses Problem zu lösen, integrierten die Autoren von [17,30] ein Graph Neural Network (GNN) in DRL-Intelligenzen, das eine zeitnahe Anpassung der Netzwerkleistungsmetriken ermöglicht, indem es den Trend der Optimierungsziele von Routing-Algorithmen vorhersagt. Dadurch können die Intelligenzen optimale Pfade erlernen, die sich an unterschiedliche Umgebungsänderungen anpassen können . In [31,32] schlugen die Autoren verschiedene DRL-Architekturen vor, die alle die Link-Gewichte dynamisch entsprechend der Belastung der Schlüsselknot

In [33] erweiterten die Autoren die TCP-Optionen, um die relevanten Steuerinformationen mitzunehmen und die Kommunikation zwischen den Subflows der Transportschicht und dem SDN-Controller flexibel zu unterstützen. Es wurde eine SDN-kooperierte MPTCP-Architektur (scMPTCP) vorgeschlagen, die Routen für neue Subflows basierend auf der verfügbaren Bandbreite jeder Route auswählt.

und vermeidet den Engpass anderer Subflows. Außerdem kann es sich an Änderungen der Netzwerklast anpassen. In [34] schlugen die Autoren einen DQN-Controller vor, der ein Pfadverlustmodell basierend auf dem Markov-Entscheidungsprozessmodell zur Netzwerkauswahl und adaptiven Ressourcenzuweisung in heterogenen Netzwerken berücksichtigt.

Eine Untersuchung der oben genannten Forschungsergebnisse zeigt jedoch, dass bestehende DRL-basierte Studien aufgrund ihrer inhärenten Kopplung an die Eingangszustände und der begrenzten Topologie, die mit den Trainingsdaten übereinstimmt, nicht ohne weiteres skalierbar sind. Um den oben genannten Widerspruch aufzulösen, kann das vorgeschlagene DRL-basierte Multipath-Routing, das in ein GNN eingebettet ist, die optimale Aktion innerhalb eines kontinuierlichen Aktionsraums leicht identifizieren. DMR entkoppelt das Multipath-Routing-Problem in einzelne Komponenten – Multipath-Erkennung und Multipath-Verkehrsplanung –, um die Komplexität der Lösung zu reduzieren. Gleichzeitig kann die von DMR übernommene Zielrichtlinie entsprechend den vorherrschenden Umgebungsbedingungen geändert werden, wodurch die Flexibilität des Systems in Bezug auf unterschiedliche Topologien verbe

Machine Translated by Google

3. Systemmodell und Problemformulierung 3.1.

Mehrwegeszenario **Das**

Mehrwege-Routing-Szenario im MCN ist in Abbildung 1 dargestellt, wobei das Satellitennetzwerk über das SDN in voneinander unabhängige Datenebenen und Steuerebenen aufgeteilt wird.
Im terrestrischen Netzwerk werden die Daten im Sendepuffer über mehrere Hops durch Router zum Sender übertragen und in mehrere Subflows aggregiert. Jeder Subflow kann je nach aktuellem Netzwerkzustand den optimalen Übertragungspfad auswählen und dann beim Empfänger zusammenlaufen.

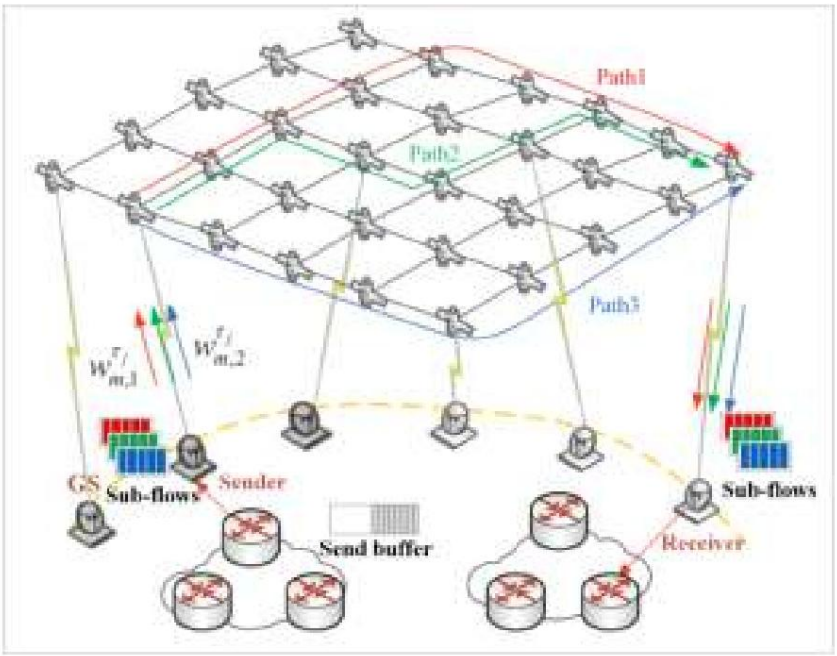


Abbildung 1. Multipath-Routing im Megakonstellationsnetzwerk.

Die Datenebene besteht aus der Walker-Delta-Konstellation NSNP/NP/F, **die gleichmäßig und symmetrisch verteilt ist, und der Bodenstation. Die** NP- **Orbitalebenen sind gleichmäßig entlang des Äquators verteilt, während die** NS- **Satelliten gleichmäßig auf jeder Ebene verteilt sind. Jeder Satellit kann an vier Inter-Satellite- Links (ISLs) angeschlossen werden, die aus zwei Intra-Plane-ISLs und zwei Inter-Plane-ISLs bestehen. In der Kontrollebene übernimmt das Netzwerkkontrollzentrum die Funktionen der einheitlichen Topologieverwaltung, der Verkehrsplanung und der Pfadsteuerung. Aufgrund der Mobilität von LEO-Megakonstellationsnetzwerken unterliegen Inter-Plane-ISLs häufigen Störungen und Umbauten, und die Topologie befindet sich in einem Zustand ständigen Wandels. Folglich muss das NCC kontinuierlich aktualisiert werden, um der sich entwickelnden Topologie gerecht zu werden, was eine erhebliche Herausforderung für die Mehrwege-Routingplanung und die V**

3.2. Multipath-Routing und Verkehrsmodell

Das MCN wird als räumlich-zeitlicher Graph G = {(V, **E, T, N)} modelliert, wobei** V = S ÿ G **der Knoten ist, der aus allen Satelliten und terrestrischen Gateways besteht, und** E = ES ÿ EG **die Menge der ISLs** ES **und die Menge der terrestrischen Links** EG ist, **während** T **der Zeitschlitzvektor einer Länge N ist. Wie in Abbildung 1 dargestellt, ermöglicht die Maschentopologie des MCN die Existenz mehrerer potenzieller Routen zwischen jedem beliebigen Knotenpaar. Nehmen wir an, dass die Anzahl der Paketflüsse im Zeitschlitz** t M **ist (d. h.** Pt = **{ (si ,** di)|i = **1, 2, · · ·, M}). Mit anderen Worten stellt die Variable** M **die Anzahl der Unterflüsse dar, wie in Abbildung 1 gezeigt, wo drei Unterflüsse im Sender auf die Übertragung über mehrere Pfade warten. Die Anzahl der verfügbaren Pfade zwischen den**

Quellknoten si **und Zielknoten** di **ist** Lsi ,di , **bezeichnet als** Rt = p $^{T}_{1,1}$, p $^{t}_{1,L}$, · · ·, p $^{t}_{M,L}$.

Hier, p $^{T}_{m,l}$ **ist der l-te Pfad zwischen dem m-ten Paar von Quell- und Zielknoten, bestehend aus**

Machine Translated by Google

aller ISLs auf diesem Pfad (d. h. p **m,l** ^{T} = *ÿ1,ÿ2, · · ·,ÿj , · · ·,ÿ*nm,l , **wobei nm,l der gesamte Hop ist -Anzahl der aktuellen Verbindung).**

Sei q $^{T}_{ÿj}$ **sei die Warteschlangenverzögerung der Verbindung** $^{t}_{ÿj}$ **sei die Paketverarbeitungsverzögerung und** *ÿ* **sei** *r ÿj ÿj* **die Verbindungsdistanz. Dann kann die Verzögerung des l-ten Übertragungspfads des m-ten Quell- Zielknotenpaars** ^{T} **wie folgt ausgedrückt werden:**

$$ÿ^{T}_{m,x} = \sum_{j=1}^{nm,x} ÿ_{ÿj}{}^{T}\ ÿj\ t + r + q\ ÿj\ \frac{t\ ÿ}{C} \quad (1)$$

Dabei ist c **die Vakuumlichtgeschwindigkeit und nm,l bezeichnet die Hop-Anzahl der aktuellen Verbindung.**

Jeder Quellknoten kann den Verkehr in Lsi ,di- **Unterflüsse aufteilen, die parallel** durch Lsi ,di- Pfade geleitet werden. Lassen Sie ÿt **das Verkehrsverhältnis jedes Unterflusses sein. Dann sollte die Summe von** Lsi ,di zwischen jedem Paar von Quell- und Zielknoten eins sein, was **wie folgt** **Untergeordnete Flüsse** $_{m,l}$ **ausgedrückt werden kann :**

$$\sum_{l=1}^{M} ÿ_{t\ ÿ\ m,l} = 1,\ ÿm = 1, 2, \cdots, M \quad (2)$$

Daher kann der Verkehr auf der Verbindung ÿj **zum Zeitpunkt** t **wie folgt ausgedrückt werden:**

$$f^{T}_{ÿj} = \sum_{m=1}^{M}\sum_{l=1}^{M} ÿ_{tt\ b\ ÿ\ mÿ\ m,l} \quad (3)$$

wobei b t $_{M}$ **ist der Bandbreitenbedarf des m-ten Paares zu den Quell- und Zielknoten ist die Paketverlustrate und** *ÿ* **m,l**^{T} **des Pfades. Da die maximale Bandbreite der interstellaren Verbindung** BW ist, **sollte der ISL-Verkehr die maximale Bandbreitengrenze nicht überschreiten, die wie folgt ausgedrückt werden kann:**

$$f^{T}_{ÿj}\ ÿ\ BW,\ ÿÿj\ ÿ\ p^{T}_{m,l} \quad (4)$$

Folglich kann der Durchsatz zwischen dem m-ten Knotenpaar wie folgt ausgedrückt werden:

$$f = \mathbf{max}\ f_{m,l}\ ÿj^{T}, ÿÿj\ ÿ\ p^{T}_{m,l} \quad (5)$$

3.3. Problemformulierung

Um die Verzögerung und den Durchsatz von ISLs zu messen, wird die Nutzenfunktion des MCN ist definiert als

$$U\ \bar{ft}, \bar{dt} = ÿ1\ \log\ \bar{ft}\ ÿ\ ÿ2\ \log\ \bar{dt} \quad (6)$$

wobei ÿ1 und ÿ2 die Wichtigkeitskoeffizienten des Durchsatzes bzw. der Verzögerung sind, ft **und** $\bar{dt}$ **der durchschnittliche Durchsatz und die** bzw. **der aktuelle Zeitschlitz sind, was wie folgt durchschnittliche Verzögerung von ÿ1 und ÿ2** sind, die die Bedingung **ÿ1** + **ÿ2** = **1 erfüllen , ausgedrückt werden kann:**

$$\bar{Fuß} = \sum_{m=1}^{M}\sum_{l=1}^{|fm|}\sum_{y=1}^{|pk} ÿ\ \frac{ypak}{T}, ÿt\ ÿ\ [1, T] \quad (7)$$

$$\bar{dt} = \frac{\sum_{l=1}^{M}\sum_{y=1}^{|fm|\ |pk} ÿ\ ÿ_{m,l}\ ÿy\ d\ t}{ÿt\ ÿ\ [1, T]} \Big/ \sum_{m=1}^{M}\sum_{l=1}^{|fm|\ |pk} ÿ\ ÿ\ ÿy \quad (8)$$

wobei | fm| **die Anzahl der Datenflüsse zwischen dem m-ten Knotenpaar,** |pk **| die Anzahl der Pakete für den k-ten Datenfluss,** ypak **die Paketgröße und** ÿy **ÿ {0, 1} eine binäre Variable ist , die angibt, ob das y-te Paket erfolgreich gesendet wurde.**

Machine Translated by Google

Folglich kann die Mehrwege-Routing-Optimierung des Satellitennetzwerks ausgedrückt werden als als folgendes Optimierungsproblem:

$$
\begin{aligned}
&\textbf{P0 : max } U\,f, \quad \bar{D}\ \textbf{,}\ \ddot{y}t\ \ddot{y}\ \textbf{[1,}\ T] \\
&\textbf{C1 :}\ \bar{d} = \frac{1}{N}\sum_{ich=1}^{N} \ddot{y}\ \bar{dt} \\
&\textbf{C2 :}\ \bar{f} = \frac{1}{N}\sum_{ich=1}^{N} \ddot{y}_{Fuß} \\
&\text{st } \textbf{C3 :}\ f\,\ddot{y}j \le L^{T}\ \ddot{y}\ BW,\ \ddot{y}\ddot{y}j\ \ddot{y}\ p^{T}_{\textbf{m,l}} \\
&\textbf{C4:}\ \ddot{y}^{\ddot{y}t}_{l=1} = 1,\ \ddot{y}m\ \ddot{y}\ \textbf{[1,}\ M]\ \textbf{m,l} \\
&\textbf{C5 :}\ d\,t\ \ddot{y}\ d\,t_{\textbf{m,l}} \le \textbf{Mein}, \ddot{y}\textbf{1}\ \ddot{y}\ l\ \ddot{y}\ y\ \ddot{y}\ \textbf{L,}\ \ddot{y}m\ \ddot{y}\ \textbf{[1,}\ M]
\end{aligned}
\tag{9}
$$

wobei die Beschränkungen C1 und C2 die durchschnittliche Verzögerung bzw. den Durchsatz des Satellitennetzwerks bezeichnen, Beschränkung C3 vorschreibt, dass der Durchsatz der Verbindung $\ddot{y}j$ **zu einem beliebigen Zeitpunkt** t **die Bandbreitenbeschränkung nicht überschreiten kann, Beschränkung C4 bedeutet, dass die gesamte Kommunikation zwischen jedem Knotenpaar das kumulative Ergebnis mehrerer Unterströme ist, die mehrere Pfade durchlaufen, und schließlich beschreibt Beschränkung C5 die Verzögerungsbeschränkungen aller Kandidatenpfade zwischen jedem Knotenpaar. Man erkennt, dass P0 ein NP-schweres Problem mit einem robusten Lösungsraum ist. P0 besteht aus zwei Teilen – Mehrwege-Routing-Erkennung und Mehrwege-Verkehrsplanung – die langsam konvergieren und in lokale Optima fallen können, wenn sie durch stochastische Optimierung gelöst werden. Um die Komplexität der Lösung zu reduzieren, entkoppelt dieses Dokument das Problem in zwei Unterprobleme, Routing-Erkennung und Verkehrsplanung**

4. DRL-basiertes Multipath-Routing In

diesem Dokument schlagen wir DRL-basiertes Multipath-Routing (DMR) vor, wie in Abbildung 2 dargestellt. Dabei wird zunächst der auf der Mindestanzahl von Hops basierende Multipath-Erkennungsalgorithmus verwendet, um mehrere Pfade mit akzeptablen Verzögerungen zu ermitteln, und anschließend der GMTS-Algorithmus eingesetzt, um das Verkehrsplanungsschema basierend auf dem Zustand des Ko

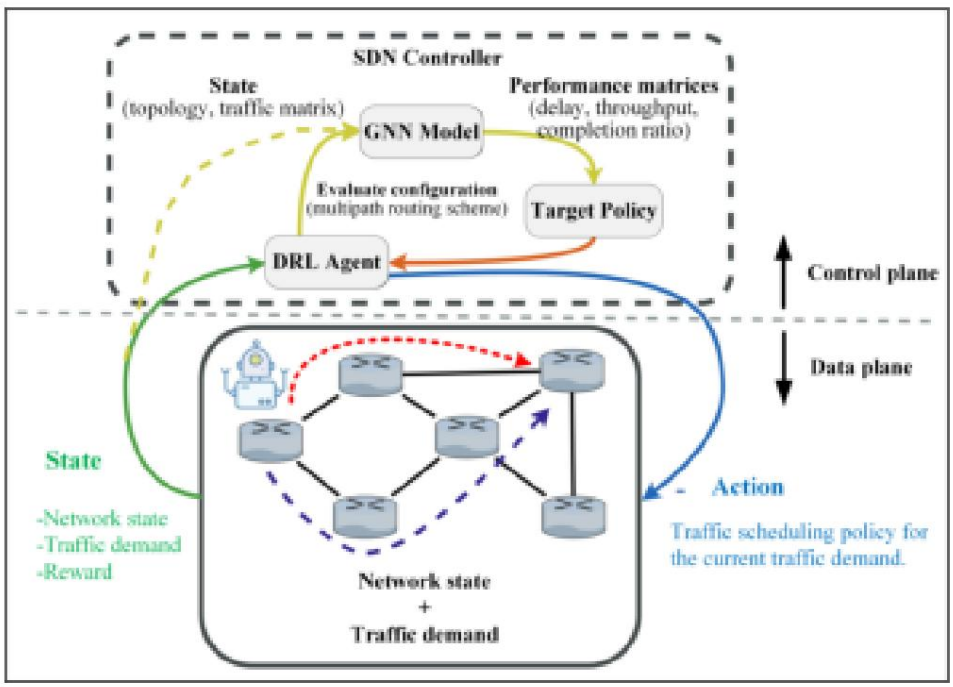


Abbildung 2. Die Gesamtstruktur des DRL-basierten Multipath-Routings.

Machine Translated by Google

4.1. Multipath-Routing-Erkennung

Das Ziel des Teilproblems der Multipath-Routing-Erkennung besteht darin, den optimalen und den suboptimalen Verzögerungspfad zu identifizieren. In MCNs stellt die Ausbreitungsverzögerung, die von der Anzahl der Pfadsprünge abhängt, eine wesentliche Komponente der End-to-End- Verzögerung dar. Folglich werden Multipath-Pfade in erster Linie für die sprungoptimalen und suboptimalen Pfade innerhalb des Netzwerks berücksichtigt. Das Teilproblem der Multipath-Routing-Erkennung kann wie folgt ausgedrückt werden:

$$\begin{aligned} \textbf{P1 : min } & d\,t_{\textbf{m,l}}, \quad \ddot{y}t\ \ddot{y}\ \textbf{T,}\ m\ \ddot{y}\ \textbf{[1,}\ M],\ l\ \ddot{y}\ \textbf{[1,}\ L] \\ st\ & d\,t\ \ddot{y}\ d\,t_{\textbf{m,l}} \qquad _{\textbf{Mein}}\textbf{, }\ddot{y}\textbf{1}\ \ddot{y}\ l\ \ddot{y}\ y\ \ddot{y}\ L \end{aligned} \tag{10}$$

Zunächst wird der Pfad mit der minimalen Hop-Anzahl zwischen jedem Knotenpaar bestimmt. Auf dieser Grundlage bestimmt das NCC den suboptimalen Übertragungspfad (d. h. den Backup-Pfad) für jeden Fluss in absteigender Verkehrsreihenfolge. Dies liegt daran, dass zentralisierter Verkehr eher Engpässe bei der Verbindungsüberlastung verursacht als dezentralisierter Kleinverkehr. In der Wakler-Delta- Konstellation NSNP/NP/F, **wie in Abbildung 3 dargestellt, beträgt die Neigung ÿ und die Differenz der Rektaszension des aufsteigenden Knotens (RAAN) benachbarter Umlaufbahnen beträgt ÿÿ** = 2ÿ/NP. **Die Phasendifferenz zwischen benachbarten Satelliten in derselben Umlaufbahn beträgt ÿÿ** = 2ÿ/NS, **während die Phasendifferenz zwischen benachbarten Satelliten in benachbarten Umlaufbahnen** *ÿÿ* = 2ÿF/ NSNP beträgt, **während u1 und u2 die Argumente der Breitengrade (oder Phasenwinkel) der Quell- bzw. Zielsatelliten sind . Das Argument des Breitengrads ist der Winkel zwischen dem aufsteigenden Knoten und dem Satelliten und definiert grundsätzlich die Position des Satelliten in der Umlaufbahn, wobei** ÿ(u1) **und** ÿ(u2) jeweils **den Längengradunterschied des Quell- und Zielsatelliten zu den entsprechenden aufsteigenden Knoten angeben. In der Walker-Delta-Konstellation u**

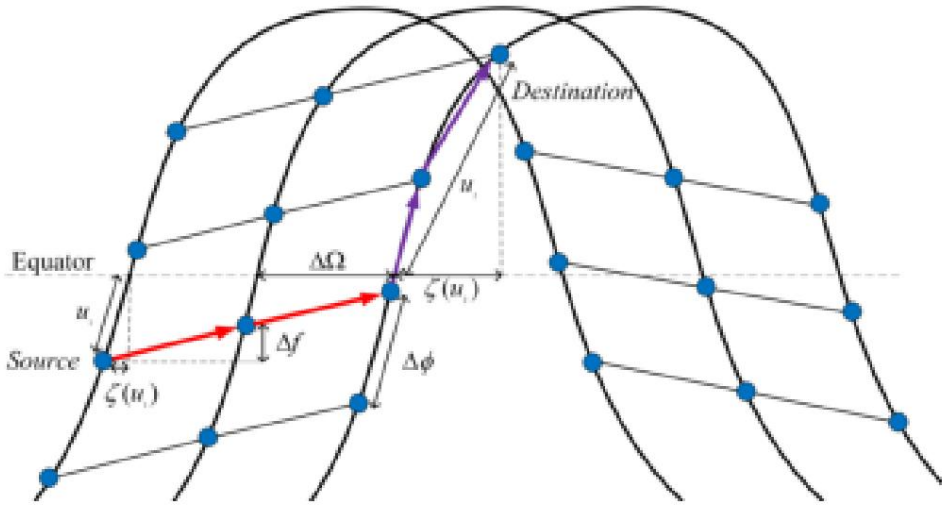


Abbildung 3. Spur des Subsatellitenpunkts von Megakonstellationsnetzwerken.

4.1.1. Zwischenebenensprünge

Hh **Die Anzahl der Zwischenebenensprünge hängt von der RAAN-Differenz ÿÿ0 der Umlaufbahnen** ab , **in denen sich die Quell- und Zielknoten befinden. Bei einem gegebenen Quell- und Zielsatellitenpaar (s,** d) wird **ÿÿ0 wie folgt berechnet:**

$$\ddot{y}\ddot{y}\textbf{0} = \textbf{(}\ddot{y}\textbf{2}\ \ddot{y}\ \ddot{y}\textbf{1) mod } 2\ddot{y}\ \ddot{y}\ \textbf{[0,}\ 2\ddot{y}\textbf{]} \tag{11}$$

wobei ÿ1 und ÿ2 die RAAN-Werte von s **bzw. d sind. Befindet sich der Zielsatellit westlich des Quellsatelliten, dann beträgt die RAAN-Differenz** *2ÿ* **ÿ ÿÿ0. Da die RAAN**

Machine Translated by Google

Da der Unterschied zwischen benachbarten Umlaufbahnen konstant ist (d. h. ÿÿ = 2ÿ/NP), können **die Sprünge zwischen den Ebenen in West- und Ostrichtung wie folgt ausgedrückt werden:**

$$H_{h}^{ÿ} = \frac{2ÿ\ ÿ\ ÿÿ0\ ÿÿ}{} \quad \textbf{(12)}$$

$$H_{h}^{ÿ} = \frac{ÿÿ0}{ÿÿ} \quad \textbf{(13)}$$

wobei ÿxÿ = sgn(x)ÿx + **1/2ÿ** die Ganzzahl bezeichnet, die **x am nächsten kommt ,Hÿ die Anzahl der Sprünge zwischen den Orbitalebenen im Westen und Hÿ die Anzahl der Sprünge zwischen den Ebenen im Osten ist.**

4.1.2. Sprünge innerhalb der

Ebene Hv **Die Sprünge innerhalb der Ebene hängen von der Phasenwinkeldifferenz** ÿu **des Satelliten ab. Jeder Sprung innerhalb der Ebene erhöht den Phasenwinkel um ÿÿ, und jeder Sprung zwischen den Ebenen führt zu einer Phasenwinkelerhöhung von** ÿf . **Daher kann der Phasenwinkel des Zielsatelliten wie folgt ausgedrückt werden:**

$$\textbf{u2} = \textbf{u1} + ÿf \cdot H_{H}^{ÿ} + \frac{ÿÿ \cdot H}{ÿÿu}\ ÿ \quad \textbf{(14)}$$

wobei H **die Anzahl der Hops in Ostrichtung innerhalb der Ebene ist. Um** Hv zu berechnen, **muss** zunächst die durch Hh **verursachte Phasendifferenz** ÿf **· Hÿ aus der Phasendifferenz** ÿu = **u2 ÿ u1 eliminiert werden, wobei u1 und u2 jeweils die Argumente der Breitengrade der Quell- und Zielsatelliten sind. Um zwischen der Phasendifferenz aufgrund der Ausbreitung in Ost- und Westrichtung zu unterscheiden,** kann ÿu **separat wie folgt ausgedrückt werden:**

$$ÿÿu = (u2\ \textbf{ÿ u1 ÿ}\ H_{H}^{ÿ} \cdot ÿf)\ \textbf{mod}\ 2ÿ \quad \textbf{(15)}$$

$$\textbf{ÿ}\ ^{ÿ}u = (u2\ \textbf{ÿ u1} + H_{H}^{ÿ} \cdot ÿf)\ \textbf{mod}\ 2ÿ \quad \textbf{(16)}$$

wobei ÿf **die Phasenwinkeländerung aufgrund von Sprüngen zwischen den Ebenen ist. Wie in Abbildung 3 dargestellt, ist die Spur des Subsatellitenpunkts in ein aufsteigendes Segment (von Südwesten nach Nordosten) und ein absteigendes Segment (von Nordwesten nach Südosten) unterteilt ,** wodurch die **Datenausbreitung in beide Richtungen möglich ist .** Daher **werden in diesem Dokument die Sprünge innerhalb der Ebenen in den vier Richtungen berechnet, die wie folgt ausgedrückt w**

$$H_{\ }^{ÿ} = \frac{ÿ\ u^{ÿ}}{ÿÿ} \quad \textbf{(17)}$$

$$H_{\ }^{ÿ} = \frac{ÿÿu}{ÿÿ} \quad \textbf{(18)}$$

$$H_{\ }^{ÿ} = \frac{2ÿ\ ÿ\ ÿ\ u^{ÿ}}{ÿÿ} \quad \textbf{(19)}$$

$$H_{\ }^{ÿ} = \frac{2ÿ\ ÿ\ ÿÿu}{ÿÿ} \quad \textbf{(20)}$$

wobei ÿ u **und** ÿÿu **bezeichnen die Phasendifferenz aufgrund der Ausbreitung nach Westen und Osten zwischen benachbarten Umlaufbahnen. Gleichzeitig** H **und** H **bezeichnen die -Phasenunterschied aufgrund von Sprüngen in der Umlaufbahn nach Nordwesten, Nordosten, Südwesten bzw. Südosten.**

Machine Translated by Google

Die Hops zwischen den Ebenen Hh **und die Hops innerhalb der Ebenen** Hv **wurden bereits berechnet. Folglich kann die minimale Hop-Anzahl von Ende zu Ende in der geneigten Umlaufbahnkonstellation wie folgt berechnet wie folgt ausgedrückt:**

$$H = \min \begin{cases} \ddot{y}\ H\ddot{y}_H + H\ddot{y}_{\cdot\cdot\cdot} \ddot{y} \\ H\ddot{y}_H + H\ddot{y}_{\cdot\cdot\cdot} \\ H\ddot{y}_H + H\ddot{y}_{\cdot\cdot\cdot} \\ H\ddot{y}_H + H\ddot{y}_{\cdot\cdot\cdot} \end{cases} \quad (21)$$

Auf der Grundlage des kürzesten Hop-Count-Pfades werden mehrere verfügbare Pfade generiert. Nimmt man die suboptimale Pfadberechnung zwischen Knotenpaaren (s, d) **als Beispielsweise zunächst die Belegungshäufigkeit** $F^{T}_{a,b}$ **jeder Verbindung und die Mindestanzahl der Hops Pfadmenge** $R^{T}_{a,b}$ **berechnet. Der Grenzwert für die Verbindungsbelegungshäufigkeit** *ÿÿ* **beträgt definiert. Das Gewicht** mt $_{a,b}$ **ÿ [r1,r2] der Verbindung** (a, b) **ist gesetzt, wenn** $F^{T}_{a,b} \ddot{y} \ddot{y}$; **andernfalls** mt $_{a,b}$ **ÿ** [r2,r3], **wobei r1 < r2 < r3 Zufallszahlen sind. Um die Verschaltung von unterschiedliche Pfade, wird das NCC die Links in** R beseitigen $^{T}_{a,b}$ **aus dem Netzwerk. Anschließend Das NCC ermittelt die optimalen realisierbaren Pfade für** (s, d) **unter den verbleibenden Verbindungen. unter Verwendung des Prinzips der kürzesten End-to-End-Hop-Anzahl. Der oben genannte Prozess wird kontinuierlich wiederholt, bis** L **mögliche Pfade zwischen zwei beliebigen Knoten im Netzwerk identifiziert. Der spezifische Prozess der MHMRD ist in Algorithmus 1** dargestellt.

Die Rechenkomplexität von Algorithmus 1 beträgt O(N **log** N).

Algorithmus 1: Multipath-Routenerkennung basierend auf der Mindestanzahl an Hops (MHMRD).

Eingabe: {(V, **E, T, N)}, Quell-Zielknotenpaar** P

Ausgabe: Kandidatenpfade R

Initialisierung: R = **ÿ**
für i = 1 : N **tun**
  Berechnen Sie den minimalen Hop-Count-Pfad R ^{T} **über Gleichung (21);**
  Berechnen Sie andere verfügbare Pfade.
  für (s, d) in P **do**
    für l = **1 :** L **ÿ 1 do**
      Berechnen Sie die Belegungshäufigkeit $F^{T}_{a,b}$ **jeder ISL;**
      für (a, b) in **R** t **tun**
        $F^{T}_{a,b}$ **ÿ** $F^{T}_{a,b}$ + **1;**
      Ende
      für (a, b) in **E** ^{T} **Tun**
        wenn $F^{T}_{a,b}$ **ÿ** *ÿÿ* **dann**
          mt $_{a,b}$ **ÿ** [r1,r2];
        Ende
        wenn $F^{T}_{a,b}$ > *ÿÿ* **dann**
          mt $_{a,b}$ **ÿ** [r2,r3];
        Ende
      Ende
      für e in **R** $^{T}_{s,d}$ **Tun**
        Entfernen Sie die Kante von E ^{T}**;**
      Ende
      Berechnen Sie die möglichen Pfade R über den Pfad mit der minimalen Hop-Anzahl.
    Ende
Ende
Ende

Machine Translated by Google

4.2. Planung des Mehrwegeverkehrs

Basierend auf dem Abschluss der Multipath-Pfaderkennung ist das zweite Teilproblem des Multipath-Routings für MCNs die Multipath-Verkehrsplanung, die wie folgt ausgedrückt werden kann

$$\mathbf{P2:}\ \max\ U\left(f, \bar{D}\right),\ \ddot{y}t\ \ddot{y}\ T$$
$$\text{st}\quad \sum_{l=1}^{M} \ddot{y}_{m,l} = \mathbf{1},\ \ddot{y}m = \mathbf{1}, \cdots, M \qquad \mathbf{(22)}$$
$$f_{\ddot{y}j}\ \ddot{y}\ BW,\ \ddot{y}\ddot{y}j\ \ddot{y}\ p^{T}_{m,l}$$

Da die Entscheidung über die Verkehrszuweisung in Satellitennetzwerken unabhängig vom historischen Zustand vom aktuellen Zustand des Knotens und der Zustände der umgebenden Knoten im Netzwerk abhängt, kann das Teilproblem der Mehrwege-Verkehrsplanung als ein Markow-Entscheidungsprozess (MD Das Auftreten von Fehlern in ISLs oder Knoten hat zur Folge, dass die Funktionalität des Satellitennetzwerks beeinträchtigt wird, was zu häufigen Änderungen der Netzwerktopologie führt. Da das GNN eine überlegene Generalisierungsfähigkeit für Topologien unterschiedlicher Größe hat, schlägt dieses Dokument die Einführung von GNN-basiertem Multipath Traffic Scheduling (GMTS) vor, wie in Abbildung 4 dargestellt. Der DRL-Agent verwendet einen proximalen Policy Optimization (PPO)-Algorithmus, der ein Akteur-Kritiker-Algorithmus ist. In jeder Phase des MDP wählt das Akteurmodell eine Verkehrsplanungsentscheidung aus , und das Kritikermodell weist dieser Entscheidung basierend auf dem Feedback zur Umweltbelohnung eine Punktzahl zu. Der Prozess wird kontinuierlich wiederholt, mit dem Ziel, die Flussplanungsaufgabe zu optimieren, um die kumulative Belohnung zu maximieren.

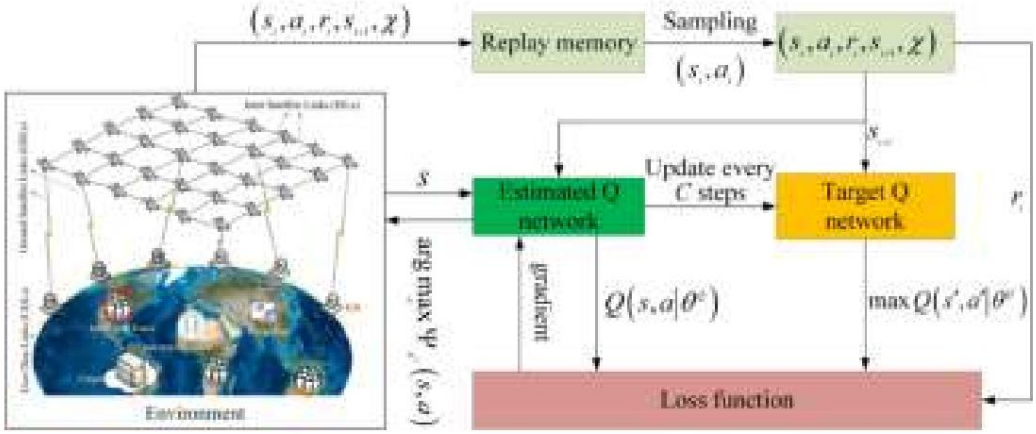


Abbildung 4. Der Rahmen des vorgeschlagenen GNN-basierten GMTS-Algorithmus (Multipath Traffic Scheduling).

4.2.1. Staat

Durch Aufteilung der Zeit in mehrere Zeitschlitze kann die Verkehrsnachfragematrix innerhalb jedes Zeitschlitzes als TR = **[b1, b2, · · ·,** bM] bezeichnet werden, **wobei** bi **die Bandbreite zwischen dem i-ten Knotenpaar bezeichnet. Gemäß Gleichung (5) ist der Durchsatz der j-ten Verbindung** fj .
Folglich kann die Restbandbreite der i-ten Verbindung, bezeichnet mit ci , **durch die Verbindungsbandbreite** BW **ausgedrückt werden :**

$$ci = BW\ \ddot{y}\ fi, \qquad i\ \ddot{y}\ [1, L] \qquad \mathbf{(23)}$$

Unter der Annahme, dass die maximale Anzahl von Hops zwischen Knotenpaaren Hmax beträgt, kann **der j-te Pfad zwischen dem i-ten Knotenpaar als pi,j** = **{e1,e2,** · · ·,eHmax } **bezeichnet werden und die entsprechende Pfadmatrix als** Pt = **[p1,1,** · · ·, **p1,L,** · · ·, pM,L]. **Daher kann der Zustandsvektor zum Zeitpunkt** t wie folgt ausgedrückt werden:

$$st = [Ct, TRt, Gt, Pt],\ st\ \ddot{y}\ S. \qquad \mathbf{(24)}$$

wobei S **der Zustandsraum des Satellitennetzwerks ist.**

Machine Translated by Google

4.2.2. Aktion

Basierend auf den Informationen im Zustandsraum teilt NCC den Netzwerkverkehr zwischen jedem Knotenpaar auf L Kandidatenpfade auf, die vorab berechnet werden. Daher wird die Aktion als Vektor definiert , der **aus dem Anteil des Verkehrs auf den Kandidatenpfaden besteht und wie folgt ausgedrückt werden kann:**

$$\text{ttt at} = \ddot{y} \cdots, \ddot{y} \cdots, \ddot{y}\ \mathbf{1,1,}\ \mathbf{1,L\ ,}\quad \mathbf{M,L} \tag{25}$$

wobei der Strömungsanteil auf jedem Pfad Gleichung (2) genügen sollte.

4.2.3. Belohnung

Nach der Ausführung der Aktion im **Zustand** st **zum Zeitpunkt t erhält der Agent Feedback, um die Wirksamkeit der Aktion zu bewerten. Bei der Mehrwegeverkehrsplanung ist die Belohnung die Zielfunktion des Teilproblems P2, die wie folgt ausgedrückt werden kann:**

$$\text{rt} = \sum_{m=1}^{M} \sum_{l=1}^{M} \cup\ \bar{\mathbf{fm,l}}\ ,\ \bar{}\text{dm,l} \tag{26}$$

wobei $\bar{\mathbf{fm,l}}$ **der durchschnittliche Durchsatz auf dem l-ten Pfad zwischen dem m-ten Knotenpaar und** $\bar{}$dm,l **die durchschnittliche Verzögerung auf dem l-ten Pfad zwischen dem m-ten Knotenpaar ist. Der Agent versucht, die Planung des Mehrwegeverkehrs mit einem höheren Durchsatz und geringerer Latenz abzuschließen. Da die Entscheidungsfindung in mehrere Phasen aufgeteilt ist, kann die kumulative Belohnung** Rt , **basierend auf einer definierten einzelnen momentanen Belohnung, wie folgt ausgedrückt werden:**

$$\text{Rt} = \sum_{k=0}^{\ddot{y}} \ddot{\mathbf{y}}\ \text{k}\,\ddot{y}\ \text{rt+k+1} \tag{27}$$

wobei $\ddot{y}$ **ÿ [0,** 1] **der Diskontierungsfaktor ist. Da es in jedem Zustand mehrere mögliche Aktionen gibt, können der durchschnittliche Gewinn für jeden Zustand und der durchschnittliche Gewinn für jede Zustandsaktion wie folgt ausgedrückt werden:**

$$\ddot{y}\mu(\text{s}) = E\mu(\ Gt|\mathbf{st} = \text{s}) \tag{28}$$

$$\ddot{y}\mu(\text{s, a}) = E\mu(\ Gt|\mathbf{st} = \mathbf{s,}\ \text{at} = \text{a}) \tag{29}$$

Hier wird ÿμ(s, a) **maximiert, wenn der Agent des DRL die optimale Strategie** μ findet , $^{\ddot{y}}$ **, indem er den Prozess der Aktualisierung** von ÿμ(s, a) **mit der Anzahl der Iterationen in die Bellman- Form umwandelt; mit anderen Worten, wir haben**

$$\ddot{y}\mu(\text{s, a}) = E\mu\ \text{rt} + \ddot{\mathbf{y}}\ddot{\mathbf{y}}\boldsymbol{\mu}\mathbf{(st+1,}\ \text{at+1})\ \text{st} = \mathbf{s,}\ \text{at} = \text{a} \tag{30}$$

Nachdem die DRL auf ein stabiles Niveau konvergiert, kann die entsprechende optimale Aktion $a\mu\ddot{y}$ **wie folgt ausgedrückt werden:**

$$a\mu\ddot{y} = \mathbf{arg\ max}^{\wedge}\quad \ddot{y}\mu\ddot{y}\ (\text{s, a}) \tag{31}$$

4.3. Trainingsprozess von GMTS

Der Pseudocode für den GMTS-Trainingsprozess wird in Algorithmus 2 mit der Rechenkomplexität von O(L) gezeigt. Zuerst ist **die Initialisierung von** μ(s|ÿ μ) **und** Q **s, a|ÿ** Q **erforderlich, wobei** $\ddot{y}$ μ **und** $\ddot{y}$ Q **Hyperparameter des Akteurnetzwerks bzw. des kritischen Netzwerks sind. Während der Trainingsphase wird der Experience Replay Buffer verwendet, um Daten zur Interaktion mit der Umgebung zu speichern, darunter Zustände, Aktionen und Belohnungen. Um eine Überanpassung zu verhindern, werden die generierten Samples** (st , at ,rt ,st+1, ÿ) **bei jedem Entscheidungsschritt im Experience Replay Buffer gespeichert. An diesem Punkt** wird $\ddot{y}$ **ÿ {0, 1} verwendet, um zu bestimmen, ob das Training des Modells in der aktuellen Phase abgeschlossen ist. Das Sample** (si , ai ,ri ,si+1 , ÿ) **wird in festgelegten Intervallen in**

Machine Translated by Google

entsprechend dem Wiedergabepufferschwellenwert Bmax mit dem Ziel, das Netzwerk zu trainieren.
An diesem Punkt kann der Ziel-Q-Wert wie folgt ausgedrückt werden:

$$Q\ yi = ri + ÿ(1\ ÿ\ ÿ)\ Q\ si+1\ ,\ ai+1|ÿ \quad (32)$$

Algorithmus 2: Trainingsprozess der GNN-basierten Multipath Traffic Scheduling (GMTS).

Eingabe: Bmax, Kritiker-Netzwerk, Akteur-Netzwerk und Ziel-Netzwerk
Ausgabe: *aµÿ*

```
für i = 1 : L do
    Setzen Sie den Anfangszustand
    als s0 für t = 1
        zurück: T do Berechnen Sie die Aktion bei , die Belohnung rt und
        den neuen Zustand st+1 Speichern Sie (st , bei , rt ,
        st+1, ÿ) im Wiedergabepuffer Aktualisieren Sie den
            Zustand st auf st+1
                während |B| ÿ Bmax do für j = 1 : NB do Berechnen Sie pi gemäß Gleichung (34)
                Berechnen Sie ÿi nach Gleichung (36)
                Berechnen Sie den Ziel- Q- Wert yi gemäß Gleichung (32)
                Aktualisieren Sie die Wahrscheinlichkeit des Abtastübergangs gemäß
                  Gleichung (35)
            end

        Berechnen Sie den Verlust des Kritiker-Netzwerks und des Akteur-Netzwerks nach
          den Gleichungen (38) und (39)
        Aktualisieren Sie die Parameter des Akteurnetzwerks und des Kritikernetzwerks
          basierend auf dem
Gradientenabstiegsansatz. Ende
Ende
```

Im GMTS-Algorithmus wird der zeitliche Differenzfehler (TD-Fehler) als Mittel zur Messung der Signifikanz von Übergängen verwendet. Der TD-Fehler zwischen dem Ziel-Q-Wert und dem aktuellen Q-Wert kann wie folgt ausgedrückt werden:

$$Qÿi = yi\ ÿ\ \mathbf{Q}\ si\ ,\ ai\ |\ ÿ \quad (33)$$

Die entsprechende proportionale Priorisierung ist

$$pi = ÿi + ÿ \quad (34)$$

wobei *ÿ* **einen extrem kleinen Betrag bezeichnet und** *ÿ* **hinzugefügt wird, damit die Stichproben auch dann abgetastet werden können, wenn der TD-Fehler Null ist. Es wird die stochastische Stichprobenmethode angewendet und die Wahrscheinlichkeit des Stichprobenübergangs** i **ist**

$$P(i) = \frac{ich^{ÿ}}{ÿk\ p^{ÿ}_{k}} \quad (35)$$

wobei *ÿ* **bestimmt, wie stark priorisiert wird. Wenn** *ÿ* **= 0, dann wird die Stichprobenmethode auf Zufallsstichproben reduziert. Die Gewichte der Wichtigkeitsstichproben (IS) von** (si , ai , ri , si+1 , ÿ) **sind**

$$ÿi = \left(\frac{Pi)}{\mathbf{Pmin}}\right)^{ÿ} \quad (36)$$

Machine Translated by Google

wobei *ÿ* **für die voreingenommene und unverzerrte Steuerung verwendet wird. Bei gegebener Lernrate** *ÿ* **Q kann das Akteurnetzwerk ÿ** *Q* folgendermaßen **aktualisiert werden:**

$$ÿ\ Q\ ÿ\ ÿ\ Q + ÿ\ \text{Qÿiÿÿ}Q\ Q\ \text{s, a}ÿ \qquad Q \tag{37}$$

Daher kann die Verlustfunktion des Kritikernetzwerks und des Akteurnetzwerks wie folgt ausgedrückt werden:

$$Lq = \frac{1}{\text{Achtung}} \sum_{ich=1}^{Bmax} ÿ\ ÿiÿ\ i^{2} \tag{38}$$

$$L\mu = ÿQ\ \text{s, a}ÿ \qquad Q \tag{39}$$

wobei NB **die Größe des Wiedergabepuffers B** ist.

4.4. Arbeitsablauf von

DMR **Das in diesem Dokument vorgeschlagene DRL-basierte Multipath-Routing (DMR) kann in zwei Phasen unterteilt werden: Multipath-Pfaderkennung auf Basis der minimalen Hop-Anzahl (MHMRD) und Multipath-Verkehrsplanung auf Basis von GNN (GMTS), wie in Abbildung 5 dargestellt. In der ersten Phase plant das Netzwerkkontrollzentrum zunächst mehrere optimale Pfade zwischen beliebigen Knotenpaaren basierend auf dem Prinzip der minimalen Hop-Anzahl entsprechend der Netzwerktopologie. Insbesondere wird das Netzwerk basierend auf der Periodizität und Vorhersagbarkeit der Satellitenumlaufbahnen entsprechend den Zeitschlitzintervallen in mehrere Zeitscheiben unterteilt, und die Netzwerktopologie bleibt in jeder diskreten Zeitscheibe stabil. Innerhalb jedes Zeitschlitzes berechnet MHMRD** L verfügbare Pfade **mit minimaler Hop** -Anzahl für jeden Dienst basierend auf der vom MHMRD **basierend auf dem End-to-End-Dienstvolumen definierten Prioritätsreihenfolge .** Nach Abschluss der Berechnung für den aktuellen **Zeitschlitz wird zum nächsten Zeitschlitz übergegangen. In der zweiten Phase generiert das Netzwerkkontrollzentrum auf Grundlage des aus dem GMTS-Training gewonnenen Modells die Aufteilungsverhältnisse für jeden Dienst auf verschiedenen Pfaden und sendet sie zum entsprechenden Zeitpunkt vor Beginn jedes Zeitschlitzes an die entsprechenden Satelliten.**

Einerseits verwendet dieser Artikel die Prioritätswiedergabe von Erfahrungen, um die Konvergenz des Modells zu beschleunigen, indem die Proben im Pool der Erfahrungen nach einer bestimmten Priorität abgetastet werden. Andererseits sind die Netzwerktopologie und die Verkehrsmatrizen aufgrund der dynamischen Natur von Satellitennetzwerken zeitvariabel, und herkömmliche DNNs können solche sich ändernden Eingaben nicht effektiv verarbeiten. Daher verwenden wir in diesem Artikel GNN für Graphstrukturen mit variabler Größe und Routenkonfigurationen entsprechend. Da GNNs über bessere Generalisierungsfunktionen für dynamische Netzwerkstrukturen in verschiedenen Maßstäben verfügen, können sie verwendet werden, um elementare Merkmale zu aggregieren, ohne die Eingabedimensionen anzugeben.

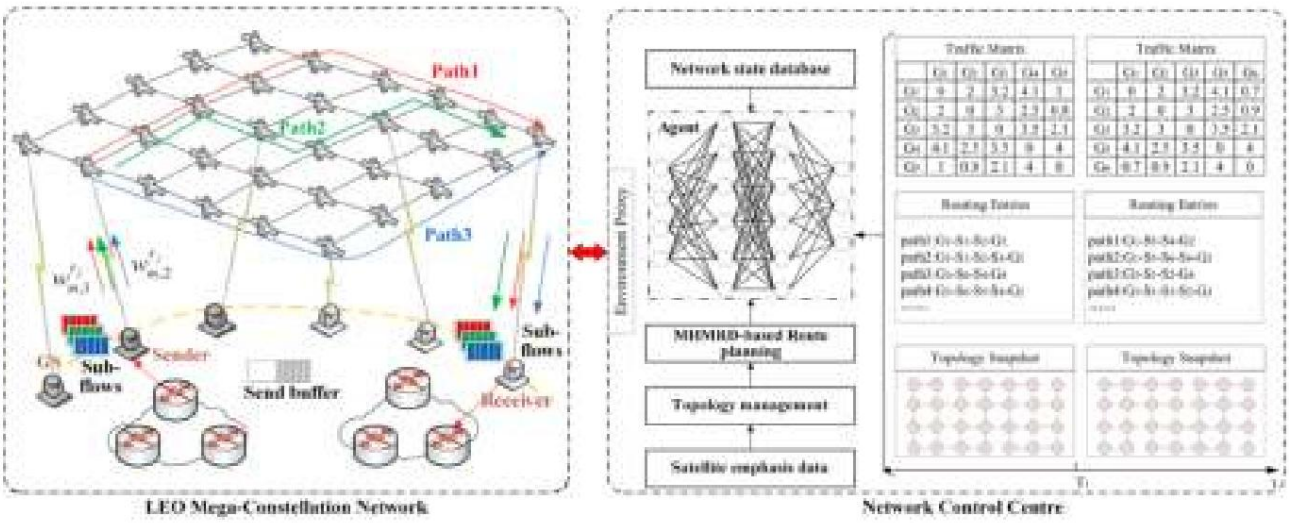


Abbildung 5. DRL-basiertes Multipath-Routing für LEO-Megakonstellationsnetzwerke.

Machine Translated by Google

5. Leistungsbewertung

5.1. Aufbau der Simulation

In diesem Beitrag implementieren wir Routing-Berechnungen und Multipath-Verkehrsplanung für LEO-Megakonstellationen in NS3, um die Leistung des vorgeschlagenen DMR-Algorithmus zu analysieren. Um die Skalierbarkeit von Multipath-Routing auf verschiedenen Konstellationen zu analysieren, In diesem Artikel werden Tests an zwei unterschiedlich großen Konstellationen durchgeführt. Die erste Konstellation war die Iridium-Konstellation [35] mit 66 polarumlaufenden Satelliten, und die zweite war die OneWeb-Konstellation [36] mit 648 Satelliten, die insgesamt 18 Orbitalebenen umfasst mit 36 Satelliten, die gleichmäßig auf jeder Orbitalebene verteilt sind. Die Parameter der beiden Konstellationen sind in Tabelle 1 dargestellt. Um die Leistung von Mehrwege- Um die Routenführung bei unterschiedlichen Verkehrsanforderungen zu optimieren, wurde ein Verkehrsdatensatz erstellt, der auf die Bodenverkehrsdichte, wobei die Anzahl der Quell-Ziel-Knotenpaare 50 betrug. Der vorgeschlagene GMTS-Algorithmus wurde basierend auf Python 3.9 und Pytorch 1.14 implementiert und die entsprechenden Parameter sind in Tabelle 2 aufgeführt.

Tabelle 1. Konstellationsparameter.

Konstellation	NP	ES	NG	ý	H	E
Iridium	6	11	16	90ÿ	780 km	20ÿ
OneWeb	18	36	16	53ÿ	550 km	20ÿ

Tabelle 2. Trainingsparameter.

	Parameter	Wert
Netzwerkparameter	Ng	24
	Lq	100
	ÿpak	1 KB
Lernparameter	Achtung Bmax	32
		100
	ý	0,6
	ÿ	0,5
	ÿ	0,8

Die Leistung des vorgeschlagenen MHMRD-Algorithmus (Minimum Hop Count Multipath Path Discovery), der den End-to-End-Verkehr gleichmäßig verteilt zu vorab berechneten Mehrfachpfaden, wurde zuerst getestet. Darauf aufbauend wurde der GMTS-Algorithmus wurde mit Hilfe der geneigten Orbitkonstellationstopologie und Verkehrsmatrizen trainiert [37], die wurden nach der Schulung im NCC eingesetzt und auf Schlüsselmetriken wie den durchschnittlichen Durchsatz, Latenz und Flussabschlussrate des Satellitennetzwerks. Die folgenden In diesem Artikel wurden Vergleichsalgorithmen verwendet:

- **Shortest Path First Algorithmus (SPF): SPF verwendet die durchschnittliche Übertragungsverzögerung als Link-Metrik mit dem Ziel, die Gesamtverzögerung zu minimieren.**
- **Netzwerkcodierungs-basiertes Multipath Cooperative Routing (NCMCR) [24]: NCMCR ist Entwickelt, um die Routing-Herausforderungen einer häufig wechselnden Topologie zu bewältigen und möglicherweise spärliche und intermittierende Konnektivität. Der NCMCR-Algorithmus nimmt Vorteil der Vorhersagbarkeit der Relativbewegung der Satelliten, wobei die zeitabhängige Topologie als räumlich-zeitliche Karte modelliert wird.**
- **Ant-basiertes Multipath-Backbone-Routing für Lastenausgleich (AMBRLB) [38]: Um die Probleme von Datenverkehr-Überlauf und Overhead zu überwinden, wurde ein auf Ant basierendes Load Balancing Multipath Backbone Routing Algorithmus in MANET wurde vorgeschlagen. Bei der Einführung der Übertragung von einem Quellknoten zu einem Zielknoten, die Ameisenkolonieoptimierung (ACO)-Algorithmus wird verwendet, um mehrere Pfade mit der höchsten Wahrscheinlichkeit zu identifizieren des Erfolgs.**

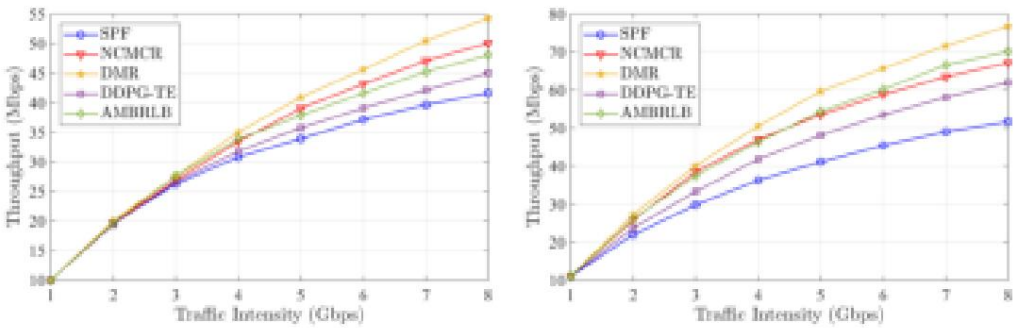
Machine Translated by Google

- **Deep deterministic policy gradient traffic engineering (DDPG-TE) [39]: Diese Methode verwendet den DDPG-Algorithmus, um die Verkehrsanteile verschiedener Übertragungswege dynamisch zu verteilen. Dabei ist zu beachten, dass das Modell nicht auf beliebige Satellitenkonstellationen anwendbar ist, da die Dimensionen der Zustands- und Aktionsräume durch die Größe der Eingangstopologie und der Verkehrsmatrizen beschränkt**

5.2. Ergebnisse und Analyse

5.2.1.

Durchsatz Der durchschnittliche Durchsatz des MCN über die Iridium-Konstellation und die OneWeb- Konstellation ist in Abbildung 6 dargestellt, wobei jeder Datenpunkt den Durchschnittswert des Durchsatzes aus fünf aufeinanderfolgenden topologischen Schnappschüssen darstellt. Es ist zu beobachten, dass im Kontext des Multipath-Routings mit zunehmendem Eingangsverkehrsvolumen ein größerer Anteil des Verkehrs auf Grundlage unterschiedlicher Richtlinien mehreren Pfaden zur Weiterleitung zugewiesen wurde. Dies führte zu einer allgemeinen Steigerung des durchschnittlichen Netzwerkdurchsatzes. Allerdings verlangsamte sich der Durchsatzanstieg allmählich. Der vorgeschlagene DMR-Algorithmus zeigte bei unterschiedlichen Benutzergrößen einen besseren Durchsatz. DMR steigerte den durchschnittlichen Durchsatz im Vergleich zu SPF und NCMCR um 30,26 % bzw. 8,45 %, als das gesamte Servicevolumen seine Obergrenze erreichte. Abbildung 6b veranschaulicht den Durchsatz des Multipath-Routings auf der OneWeb-Konstellation. Es lässt sich beobachten, dass DMR dasselbe Modell verwendet hat wie das für die Iridium-Konstellation, während die anderen Algorithmen basierend auf den Eigenschaften des neuen Netzwerks neu trainiert wurden. Es lässt sich beobachten, dass der DMR-Algorithmus im Kontext von Änderungen in der Konstellationstopologie und -größe weiterhin eine bessere Leistung im Vergleich zu den anderen Benchmark-Algorithmen zeigte , was die verbesserte **Generalisierungsfähigkeit von DMR widerspiegelt. Bei einer Verkehrsintensität von 8 Gbit/s zeigte DM**



(A) **(B)**

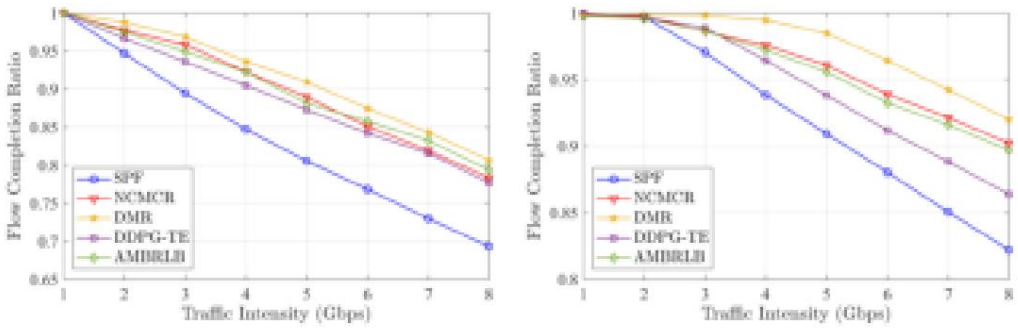
Abbildung 6. Durchschnittlicher Durchsatz bei verschiedenen Konstellationen mit unterschiedlicher Verkehrsintensität. (a) Durchschnittlicher Durchsatz für Iridium. (b) Durchschnittlicher Durchsatz für OneWeb.

5.2.2. Durchschnittliche Durchfluss-Abschlussrate

Abbildung 7 zeigt die durchschnittliche Flussabschlussrate bei unterschiedlichen Verkehrsintensitäten auf den Iridium- und OneWeb-Konstellationen. Mit zunehmender Verkehrsintensität wurden einige Knoten im Netzwerk allmählich überlastet. Dies wird durch einen abnehmenden Trend bei den durchschnittlichen Flussabschlussraten aller fünf Routing-Schemata belegt. Bei einer Verkehrsintensität von 8 Gbit/s wurde die durchschnittliche Flussabschlussrate des vorgeschlagenen DMR-Schemas **auf den Iridium- und OneWeb-Konstellationen im In der Zwischenzeit verbesserte sich auf der OneWeb-Konstellation die durchschnittliche Stream-Abschlussrate des DMR- Algorithmus um 1,98 %, 2,58 % bzw. 6,49 % im Vergleich zu NCMCR, AMBRLB und DDPG-TE. Folglich ist ersichtlich, dass die durchschnittliche Flussabschlussrate auf der größeren OneWeb-Konstellation durchweg höher war als auf der kleineren Iridium- Konstellation, selbst bei unterschiedlicher Verkehrsintensität. Dies zeigt den Vorteil großer Konstellationen bei der Handhabung hoher Lasten. Man kann daher den Schluss zie**

Machine Translated by Google

Das vorgeschlagene DMR-Schema weist eine überlegene Generalisierungsfähigkeit über verschiedene Maßstabskonstellationen hinweg auf und weist eine höhere Flussvervollständigungsrate als das Basisschema auf.



(A) **(B)**

Abbildung 7. Durchschnittliche Flussabschlussrate verschiedener Konstellationen mit unterschiedlicher Verkehrsintensität. (a) Durchschnittliche Flussabschlussrate für die Iridium-Konstellation. (b) Durchschnittliche Flussabschlussrate für die OneWeb-Konstellation.

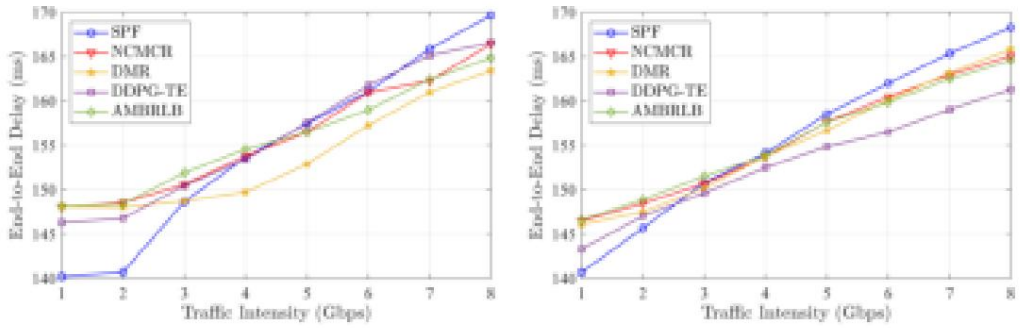
5.2.3. Durchschnittliche End-to-End-

Verzögerung Abbildung 8 zeigt die durchschnittliche End-to-End-Verzögerung als Funktion der Verkehrsintensität in den Iridium- und OneWeb-Konstellationen. Mit zunehmender Netzwerklast nahm auch die Warteschlangenverzögerung zu, da mehr Datenströme über die Kandidatenpfade mit höheren Hop-Zahle Folglich nahm die durchschnittliche End-to-End-Verzögerung bei allen verglichenen Schemata allmählich zu. In der Iridium-Konstellation erreichte SPF die niedrigste End-to-End-Verzögerung, wenn die Netzwerklast gering war. Dies liegt daran, dass das Schema die Pakete immer auf dem kürzesten Weg überträgt, wenn sich alle Links im Netzwerk in einem normalen Zustand befinden. Wenn die Netzwerklast zunimmt und das gesamte Servicevolumen die Obergrenze erreicht, kann SPF den Pfad nicht dynamisch ändern, um überlastete Pfade zu vermeiden, die durch die erhöhte Last verursacht werden. Folglich reduzierte das DMR-Schema bei der Iridium-Konstellation die durchschnittliche End-to-End-Verzögerung um 3,67 %, 1,78 %, 1,86 % und 0,86 % im Vergleich zu SPF, NCMCR, DDPG- TE und AMBRLB. Bei der Ausweitung des Multipath-Routings auf größere OneWeb- Konstellationen näherte sich die Verzögerung für DMR der von NCMCR, AMBRLB und DDPG-TE an, als die Verkehrsintensität ihre Obergrenze erreichte. Dies liegt daran, dass sowohl das DMR- Schema als auch das DDPG-TE-Schema die dynamische Verteilung des Datenverkehrs **auf verschiedene Pfade berücksichtigen .** Im Allgemeinen **lag die mittlere End-to-End-Verzögerung von DMR bei hoher Last zwischen denen von SPF und DDPG-TE. Das Hauptziel des Multipath-Routings besteht darin, den Durchsatz und die Belastbarkeit des Netzwerks zu verbessern, wenn die Topologie gestört wird. In kleinen Konstellationen bietet Multipath-Routing zusätzliche Routing-Optionen, mit denen überlastete Knoten rechtzeitig umgangen und so die Verzögerung reduziert werden kann.**

Allerdings gibt es in groß angelegten Konstellationen immer genügend Alternativrouten und die Verbesserung der Ende-zu-Ende-Verzögerungsmetriken ist nicht wesentlich.

Die experimentellen Ergebnisse zeigen, dass das SPF-Schema mit dem kürzesten Pfad in Szenarien mit geringer Belastung eine geringere durchschnittliche Verzögerung aufwies. Allerdings wies SPF den niedrigsten durchschnittlichen Durchsatz und die niedrigste durchschnittliche Flussabschlussrate auf, und die Leistung von SPF verschlechterte sich unter Bedingungen mit hoher Belastung rapide. Im Gegensatz dazu wies das vorgeschlagene DMR-Schema unter unterschiedlichen Belastungen einen besseren durchschnittlichen Durchsatz und eine bessere Flussabschlussrate auf und zeigte eine bessere Generalisierungsfähigkeit über Konstellationen unterschiedlicher Größe und Topologie ohne wiederholtes Training.

Machine Translated by Google



(A) **(B)**

Abbildung 8. Durchschnittliche End-to-End-Verzögerung für verschiedene Konstellationen mit unterschiedlicher Verkehrsintensität. (a) Durchschnittliche End-to-End-Verzögerung für die Iridium-Konstellation. (b) Durchschnittliche End-to-End-Verzögerung für die OneWeb-Kons

6. Schlussfolgerungen

In diesem Artikel wird eine Multipath-Routing- Lösung (DMR) auf der Basis von Deep Reinforcement Learning (DRL) für Satellitennetzwerke vorgeschlagen. Ziel ist die Lösung des Multipath-Routing- Übertragungsproblems in LEO-Megakonstellationsnetzwerken. Der vorgeschlagene DMR-Ansatz teilt das Multipath-Routing in zwei Teilprobleme auf: Multipath-Routenauswahl und Multipath-Verkehrsplanung. Im Zusammenhang mit dem Teilproblem der Multipath-Pfadauswahl wird in diesem Artikel ein zentralisiertes Routing-Schema, MHMRD, vorgeschlagen, das eine Reihe verfügbarer Pfade für End-to-End-Knoten generiert. Dies wird erreicht, indem Knotenstatusinformationen basierend auf der aktuellen End-to-End-Mindest-Hop-Anzahl und der Verbindungsauslastung gesammelt werden. Im Teilproblem der Multipath-Verkehrsplanung Jeder Routing-Knoten wird von einem DQN-Agenten gesteuert und das GNN wird verwendet, um das Problem topologisch unterschiedlicher Satellitennetzwerktopologien zu lösen. Das optimale Mehrwege-Verkehrsplanungsmodell wurde durch iteratives Training entwickelt. Das vorgeschlagene DMR-Schema wurde schließlich an zwei unterschiedlich großen Megakonstellationen in der NS3-Simulation Die Simulationsergebnisse zeigen, dass das vorgeschlagene DMR-Schema im Vergleich zum Basisschema eine bessere **Leistung in Bezug auf den durchschnittlichen Netzwerkdurchsatz und die durchschnittliche Flussabschlussrate** zeigte. Darüber hinaus bewältigte das DMR-Schema effektiv **das Satellitennetzwerk-Routing unter Lastausgleich, während die Verzögerungskosten nur minimal anstiegen. Darüber hinaus zeigte das in diesem Dokument entwickelte DMR-Schema eine günstige Skalierbarkeit und Anpassungsfähigkeit im Kontext sich entwickelnder Konstellationstopologien. Es kann auch auf Satellitennetzwerke mit sich ändernden Topologien ausgeweitet werden, ohne dass zusätzliche Schulungen erforderlich sind.**

Es ist zu beachten, dass mit der weiteren Vergrößerung der Konstellation auch die Rechenleistungsanforderungen des DQN-basierten Online-Routings steigen. In Zukunft möchten wir das Potenzial von Lösungen wie der Bereitstellung mehrerer Controller ausloten, um die Reaktionsfähigkeit des Netzwerks zu verbessern und seine Wirksamkeit in realen Szenarien weiter zu validi

Beiträge der Autoren: Konzeptualisierung, CH; Methodik, CH; Software, RY; Validierung, WX; formale Analyse, CH; Untersuchung, WX; Ressourcen, WX; Datenkuratierung, CH; Schreiben – Vorbereitung des Originalentwurfs, CH; Schreiben – Überprüfung und Bearbeitung, CH; Visualisierung, CH; Überwachung, WX; Projektverwaltung, WX und RY; Mittelbeschaffung, WX und RY. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Forschung wurde vom National Key Laboratory Fund Project unter der Zuschussnummer 614201003022207 und dem Aerospace Discipline Education New Engineering Project unter der Zuschussnummer 145AXL250004000X finanziert.

Erklärung zur Datenverfügbarkeit: Die relevanten Daten und der Quellcode dieses Dokuments können durch Kontaktaufnahme mit hanchi@hgd.edu.cn abgerufen werden.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Machine Translated by Google

Abkürzungen

In diesem Manuskript werden die folgenden Abkürzungen verwendet:

Tagfahrlicht	Tiefes bestärkendes Lernen
DMR	DRL-basiertes Multipath-Routing
MHMRD	Auf der Mindestanzahl von Hops basierender Multipath-Routing-Erkennungsalgorithmus
GNN	Graph neuronales Netzwerk
GMTS	GNN-basierte Mehrwege-Verkehrsplanung
MCN	Megakonstellationsnetzwerk
SAGIN	Integriertes Weltraum-Luft-Boden-Netzwerk
ISLs	Inter-Satelliten-Verbindungen
SDN	Softwaredefiniertes Netzwerk
RAAN	Rektaszension des aufsteigenden Knotens
NCC	Netzwerkkontrollzentrum
LÖWE	Niedrige Erdumlaufbahn
DQN	Deep Q-Netzwerk

Verweise

1. Jiang, X.; Huang, Y.; Li, J.; He, H.; Chen, S.; Yang, F.; Yang, J. Spatio-Temporal Routing, redundante Kodierung und Multipath Scheduling für deterministische Satellitennetzwerkübertragung. IEEE Trans. Commun. **2023,** 71, **2860–2875.** [CrossRef]
2. Kodheli, O.; Lagunas, E.; Maturo, N.; Sharma, SK; Shankar, B.; Montoya, JFM; Duncan, JCM; Spano, D.; Chatzinotas, S.; Kisseleff, S.; et al. Satellitenkommunikation im neuen Weltraumzeitalter: Eine Übersicht und zukünftige Herausforderungen. IEEE Commun. Surv. Tutor. **2021,** 23, **70–109.** [CrossRef]
3. Rakhmanov, A.; Wiseman, Y. Komprimierung von GNSS-Daten mit dem Ziel, die Kommunikation mit autonomen Fahrzeugen zu beschleunigen. Remote Sens. **2023,** 15, **2165.** [CrossRef]
4. Correia, SD; Perez, R.; Matos-Carvalho, J.; Leithardt, VRQ µJSON, ein leichtes Komprimierungsschema für eingebettete GNSS- Datenübertragung auf IoT-Knoten. In Proceedings der 5. Konferenz 2022 zu Cloud und Internet der Dinge (CIoT), Marrakesch, Marokko, 28.–30. März 2022; S. 232–238.
5. Xue, C.; Li, W.; Yu, L.; Shang, J.; Chen, X.; Lu, S. SERO: Ein modellbasiertes Seamless Roaming Framework für Wireless Mesh Netzwerk mit Multipath TCP. IEEE Trans. Commun. **2019,** 67, **1284–1296.** [CrossRef]
6. Liu, X.; Ma, T.; Qin, X.; Zhou, H.; Zhao, L. Ein DRL-gestütztes Multipath Cooperative Routing für ultradichte LEO-Satellitennetzwerke . In Proceedings of the GLOBECOM 2023, Kuala Lumpur, Malaysia, 4.–8. Dezember 2023; S. 5961–5966.
7. Li, R.; Zhang, J.; Zheng, S.; Wang, K.; Wang, P.; Zhang, X. Routing-Algorithmus für LEO-Megakonstellationen basierend auf Gebietssegmentierung . In Proceedings der 2023 IEEE Wireless Communications and Networking Conference (WCNC), Glasgow, Großbritannien, 26.–29. März 2023; S. 1–6.
8. Li, Y.; Zhang, Q.; Yao, H.; Gao, R.; Xin, X.; Yu, FR Stigmergie und hierarchisches Lernen zur Routing-Optimierung in Multi-Domain- kollaborativen Satellitennetzwerken. IEEE J. Select. Areas Commun. **2024,** 42, **1188–1203.** [CrossRef]
9. Perkins, CE; Royer, EM Ad-Hoc on-Demand Distance Vector Routing. In Proceedings of the WMCSA'99. Zweiter IEEE- Workshop zu mobilen Computersystemen und -anwendungen, New Orleans, LA, USA, 25.–26. Februar 1999; Seiten 90–100.
10. Jiang, F.; Zhang, Q.; Yang, Z.; Yuan, P. Ein auf Raum-Zeit-Diagrammen basierendes Mehrwege-Routing in der störungstoleranten Erdbeobachtung Satellitennetzwerke. IEEE Trans. Aerosp. Electron. Syst. **2019,** 55, **2592–2603.** [CrossRef]
11. Huang, Y.; Yang, D.; Feng, B.; Tian, A.; Dong, P.; Yu, S.; Zhang, H. Ein GNN-fähiger Multipath-Routing-Algorithmus für räumliche Zeitlich variierende LEO-Satellitennetzwerke. IEEE Trans. Veh. Technol. **2024,** 73, **5454–5468.** [CrossRef]
12. Huang, Y.; Jiang, X.; Chen, S.; Yang, F.; Yang, J. Pheromon-Incentivierter intelligenter Multipath-Verkehrsplanungsansatz für LEO-Satellitennetzwerke. IEEE Trans. Wirel. Commun. **2022,** 21, **5889–5902.** [CrossRef]
13. Tian, A.; Feng, B.; Zhou, H.; Huang, Y.; Sood, K.; Yu, S.; Zhang, H. Effizientes föderiertes DRL-basiertes kooperatives Caching für Mobilgeräte Edge-Netzwerke. IEEE Trans. Netw. Serv. Manag. **2023,** 20, **246–260.** [CrossRef]
14. Huang, X.; Yuan, T.; Qiao, G.; Ren, Y. Deep Reinforcement Learning für die Multimedia-Verkehrssteuerung in Software Defined Vernetzung. IEEE Netw. **2018,** 32, **35–41.** [CrossRef]
15. Lei, K.; Guo, P.; Wang, Y.; Wu, X.; Zhao, W. Lösen Sie Routing-Probleme mit einem Residual Edge-Graph Attention Neural Network. Neurocomputing **2022,** 508, **79–98.** [CrossRef]
16. Zheng, X.; Huang, W.; Li, H.; Li, G. Forschung zur verallgemeinerten intelligenten Routing-Technologie basierend auf Graph Neural Network. Elektronik **2022,** 11, **2952.** [CrossRef]
17. Wei, H.; Zhao, Y.; Xu, K. G-Routing: Flexibles Online-Routing auf Basis von Graph-Neural-Networks. IEEE Netw. **2023,** 37, **90–96.** [Querverweis]
18. Rusek, K.; Suarez-Varela, J.; Almasan, P.; Barlet-Ros, P.; Cabellos-Aparicio, A. RouteNet: Nutzung von Graph Neural Networks für Netzwerkmodellierung und -optimierung in SDN. IEEE J. Select. Areas Commun. **2020,** 38, **2260–2270.** [CrossRef]
19. Bi, Y.; Han, G.; Xu, S.; Wang, X.; Lin, C.; Yu, Z.; Sun, P. Softwaredefinierte raum-terrestrische integrierte Netzwerke: Architektur, Herausforderungen und Lösungen. IEEE Netw. **2019,** 33, **22–28.** [CrossRef]

Machine Translated by Google

20. Hu, M.; Xiao, M.; Hu, Y.; Cai, C.; Deng, T.; Peng, K. Softwaredefiniertes Multicast mit Segmentrouting in LEO-Satellitennetzwerken. IEEE Trans. Mob. Comput. **2024,** 23, **835–849. [CrossRef]**

21. Han, Z.; Xu, C.; Zhao, G.; Wang, S.; Cheng, K.; Yu, S. Zeitvariables Topologiemodell für dynamisches Routing in LEO-Satellitenkonstellationsnetzwerken . IEEE Trans. Veh. Technol. **2023,** 72, **3440–3454. [CrossRef]**

22. Giambene, G.; Luong, DK; De Cola, T.; Le, VA; Muhammad, M. Analyse eines Paket-Level-Block-Codierungsansatzes für Terrestrische Satelliten-Mobilsysteme. IEEE Trans. Veh. Technol. **2019,** 68, **8117–8132. [CrossRef]**

23. Chen, Z.; Zhou, W.; Wu, S.; Cheng, L. Ein adaptives On-Demand-Multipath-Routing-Protokoll mit QoS-Unterstützung für High-Speed- MANET. IEEE Access **2020, 8, 44760–44773. [CrossRef]**

24. Tang, F.; Zhang, H.; Yang, LT Multipath Cooperative Routing mit effizienter Bestätigung für LEO-Satellitennetzwerke. IEEE Trans. Mob. Comput. **2019,** 18, **179–192. [CrossRef]**

25. Li, G.; Zhou, H.; Feng, B.; Zhang, Y.; Yu, S. Effiziente Bereitstellung von Service-Funktionsketten in Overlay-Netzwerken durch Verstärkung ment Learning. IEEE Trans. Cloud Comput. **2022,** 10, **383–395. [CrossRef]**

26. Kato, N.; Fadlullah, ZM; Mao, B.; Tang, F.; Akashi, O.; Inoue, T.; Mizutani, K. Die Deep Learning-Vision für die Steuerung des heterogenen Netzwerkverkehrs: Vorschlag, Herausforderungen und Zukunftsperspektive. IEEE Wirel. Commun. **2017,** 24, **146–153. [CrossRef]**

27. Mao, B.; Fadlullah, ZM; Tang, F.; Kato, N.; Akashi, O.; Inoue, T.; Mizutani, K. Routing oder Computing? Der Paradigmenwechsel hin zur intelligenten Paketübertragung in Computernetzwerken auf Basis von Deep Learning. IEEE Trans. Comput. **2017,** 66, **1946–19[Querverweis]**

28. Geyer, F.; Carle, G. Lernen und Generieren verteilter Routing-Protokolle mithilfe von graphenbasiertem Deep Learning. In Proceedings des Workshops 2018 zu Big Data Analytics und maschinellem Lernen für Datenkommunikationsnetzwerke, Budapest, Ungarn, 20. August 2018; S. 40–45.

29. Liu, C.; Xu, M.; Yang, Y.; Geng, N. DRL-OR: Deep Reinforcement Learning-basiertes Online-Routing für Multi-Type-Serviceanforderungen . In Proceedings der IEEE INFOCOM 2021—IEEE Conference on Computer Communications, Vancouver, BC, Kanada, 10.–13. Mai 2021; S. 1–10.

30. Almasan, P.; Suárez-Varela, J.; Rusek, K.; Barlet-Ros, P.; Cabellos-Aparicio, A. Deep Reinforcement Learning trifft auf Graph Neural Networks: Untersuchung eines Anwendungsfalls zur Routing-Optimierung. Comput. Commun. **2022,** 196, **184–194. [CrossRef]**

31. Sun, P.; Guo, Z.; Li, J.; Xu, Y.; Lan, J.; Hu, Y. Ermöglichung skalierbaren Routings in softwaredefinierten Netzwerken mit Deep Reinforcement Learning auf kritischen Knoten. IEEE/ACM Trans. Netw. **2022,** 30, **629–640. [CrossRef]**

32. Kim, G.; Kim, Y.; Lim, H. Routing auf Basis von Deep Reinforcement Learning in softwaredefinierten Netzwerken. IEEE Access **2022,** 10, **18121–18133. [CrossRef]**

33. Jiang, Z.; Wu, Q.; Li, H.; Wu, J. scMPTCP: SDN-kooperierter Mehrwegetransfer für Satellitennetzwerke mit Lasterkennung. IEEE Access **2018, 6, 19823–19832. [CrossRef]**

34. De Santis, E.; Giuseppi, A.; Pietrabissa, A.; Capponi, M.; Delli Priscoli, F. Satellitenintegration in 5G: Tiefgreifende Verstärkung Lernen für die Netzwerkauswahl. Mach. Intell. Res. **2022,** 19, **127–137. [CrossRef]**

35. Iridium-Konstellation. Online verfügbar: https://celestrak.org/NORAD/elements/table.php?GROUP=iridium (aufgerufen am 30. Juli 2024).

36. Oneweb Constellation. Online verfügbar: https://celestrak.org/NORAD/elements/table.php?GROUP=oneweb (aufgerufen am 30. Juli 2024).

37. MAWI-Arbeitsgruppen-Verkehrsarchiv. Online verfügbar: https://mawi.wide.ad.jp/mawi/ (abgerufen am 30. Juli 2024).

38. Francis Antony Selvi, P.; Manikandan, MSK Ant-basiertes Multipath-Backbone-Routing für Lastausgleich in MANET. IET Commun. **2017,** 11, **136–141. [CrossRef]**

39. Chen, B.; Sun, P.; Zhang, P.; Lan, J.; Bu, Y.; Shen, J. Verkehrstechnik basierend auf Deep Reinforcement Learning in hybriden IP/SR- Netzwerken. China Commun. **2021,** 18, **204–213. [CrossRef]**

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Entwurf eines Störungskompensators basierend auf erweitertem LMI für Lineare Parametervariationssysteme

Mustafa ÿInci 1 und Yusuf Altun 2,*

[1] Volta Motor Industry and Trade Inc. Design Center, Cumayeri Factory, 81700 Düzce, Türkiye; mustafa@volta.com.tr

[2] Fakultät für Computertechnik, Universität Düzce, 81620 Düzce, Türkiye * Korrespondenz: yusufaltun@duzce.edu.tr

Zusammenfassung: Dieser Artikel präsentiert eine neue Darstellung der erweiterten linearen Matrixungleichung (LMI) zur Entwicklung eines Zustandsrückkopplungsreglers und eines dynamischen Feedforward-Störungskompensators für lineare Parametervariationssysteme (LPV). Die verbesserten LMIs sind konvex und endlichdimensional ohne jeglichen iterativen Ansatz. Die Entwürfe basieren auf einem neuen vorgeschlagenen äquivalenten beschränkten reellen Lemma (BRL) mittels Matrixdilatation für LPV-Systeme und unsichere lineare Systeme unter zeitabhängigen parametrischen Unsicherheiten (TVPUs). Dieses erweiterte BRL liefert weniger konservative Ergebnisse als bestehende Methoden in Bezug auf robuste Stabilität. Dementsprechend wird zusätzlich zu einem Zustandsrückkopplungsregler ein dynamischer Störungskompensator entwickelt . Dieser Artikel konzentriert sich hauptsächlich auf die Entwicklung von Kompensatoren gegen Störungen zusätzlich zur Entwicklung von Zustandsrückkopplungsreglern. Die dynamischen Matrizen des Kompensators ändern sich mit den zeitabhängigen Parametern des LPV oder unsicheren Systems während des Betriebs, vorausgesetzt, dass die Störungen und die Parameter messbar oder beoba Abschließend werden numerische und Simulationsergebnisse vorgestellt, um sowohl die Wirksamkeit als auch die geringere Konservativität der vorgeschlagenen LMIs zu beweisen.

Schlüsselwörter: LPV-Feedforward; Störungskompensator; Hÿ-Steuerung; erweitertes BRL; robuste Steuerung

check for updates

Zitat: ÿInci, M.; Altun, Y. Entwurf eines Störungskompensators Basierend auf erweitertem LMI für linear Parametervariierende Systeme. Elektronik **2024,** 13, **3055.** **https://doi.org/10.3390/Elektronik13153055**

Wissenschaftlicher Redakteur: Luca Patanè

Erhalten: 5. Juli 2024
Überarbeitet: 26. Juli 2024
Akzeptiert: 29. Juli 2024
Veröffentlicht: 1. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz. Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

Robuste Leistungsanalysen von LPV-Systemen, die eine spezielle Klasse nichtlinearer Systeme darstellen, und linearen Systemen mit zeitvariablen parametrischen Unsicherheiten (TVPUs) haben beträchtliche Aufmerksamkeit auf sich gezogen. In den letzten zwei Jahrzehnten wurden für diese Systeme viele LMIs auf Grundlage der Stabilitätskriterien von Ljapunov vorgeschlagen, da die Effizienz von auf LMIs basierenden Methoden gut bekannt ist, wie z. B. [1–4]. Die meisten von ihnen basieren auf der standardmäßigen quadratischen Ljapunov-Funktion, die als quadratische Stabilität bezeichnet wird, wobei eine einzelne quadratische Ljapunov-Funktion wie in [5] verwendet wird. So wird beispielsweise in [6] ein numerischer Vergleich hinsichtlich der Konservativität der quadratischen Stabilitätsbedingungen angestellt . Viele Studien haben jedoch gezeigt, dass eine parameterabhängige Ljapunov-Matrix (PDLM) die Konservativität, die aus einer einzelnen quadratischen Ljapunov-Funktion für Systeme wie die in [6–13] abgeleitet wird, überwindet. Für LPV-Systeme oder kontinuierliche Systeme (CTS) unter TVPUs ist es ohne zusätzliche Methode jedoch nicht geeignet, da PDLM oder parametrisiertes LMI (PLMI) ein nichtkonvexes Optimierungsproblem verursachen. PDLM verringert jedoch den Konservativismus, der sich aus den quadratischen Stabilitätsbedingungen ergibt. Zur Lösung dieses Problems wurden einige Ansätze vorgeschlagen, wie beispielsweise die Gittermethode [14], das s-Verfahren [15], die Multikonvexitätsmethode [15,16] und die Methoden in [17,18]. Die Gittermethode ist jedoch nur für lokale Betriebsbedingungen gültig, die allen zulässigen Parametertrajektorien w Darüber hinaus beinhalten die anderen Ansätze zusätzliche LMI-Bedingungen, einige Einschränkungen oder Komplexitäten abgesehen von den bestehenden LMIs, um eine konvexe Lösung zu erhalten, wenn das PDLM verwendet wird. Darüber hinaus ermöglichen diese Ansätze nicht die Entwicklung von Mehrzwe

Machine Translated by Google

in [12]. Darüber hinaus hängt ein mit den Methoden entwickelter Zustandsrückkopplungsregler von einer Ljapunow-Funktion ab, die durch die Symmetrieeigenschaft eingeschränkt ist.

Um eine geringere Konservativität für unsichere und LPV-Systeme zu erreichen, wurden in jüngster Zeit von den Parametern abhängige Lyapunov-Matrizen oder PLMIs ohne zusätzliche LMI- Bedingungen erforscht. Dementsprechend wurden „dilatierte“ oder „erweiterte“ LMIs zur Lösung nichtkonvexer Probleme vorgeschlagen und erfreuten sich in den letzten zwei Jahrzehnten großer Beliebtheit. S-Variablen- und Slack-Variable-LMIs (SV-LMIs) sind ebenfalls bekannt, da Slack-Variablen im Allgemeinen für LMIs verwendet werden [19]. Der Vorteil dilatierter LMIs besteht darin, dass für sehr große Parametersätze weniger konservative Ergebnisse erzielt werden, da sie im Gegensatz zu den üblichen **LMIs wie in** [5] **nicht die Multiplikation der Lyapunov-Matrix mit den Zustandsraummatrizen des Systems** enthalten . **In diesem Fall können wir auch die robuste Leistung oder Stabilität überprüfen, und aufgrund der Trennung wird für die Systeme eine konvexe LMI-Formulierung erhalten. Darüber hinaus waren dilatierte LMIs wertvolle Werkzeuge für die Entwicklung von Mehrzielreglern mit geringerer Konservativität, wie in [20]. Kurz gesagt, der Ansatz des dilatierten LMI bietet trotz des PDLM** eine konvexe Lösung **und weniger konservative Ergebnisse für LPV oder unsichere lineare Systeme ohne weitere** Bedingungen, was in gängigen Designs zu Nichtkonvexität oder unendlichen LMIs führt. In **[8,21] wird zum Beispiel eine Stabilitätsbedingung vorgeschlagen , ein Hÿ -Reglerdesign wird jedoch nicht behandelt. In [20] wird ein offensichtlicher Vorteil dilatierter LMIs gegenüber klassischen LMIs für die Entwicklung multikriterieller H2/D-Stabilitätsregler vorgestellt, diese Formulierung lässt sich jedoch nicht auf die Synthese** von **Hÿ-** Regler erweitern. In **[22]** wird **eine H2- Regelung über eine Schlupfvariable vorgeschlagen, die Hÿ- Regelung wird jedoch nicht behandelt. In [23] werden weniger konservative Bedingungen für dilatierte LMIs unter rational variierenden Parametern vorgeschlagen, diese sind jedoch für ein CTS mit TVPU- und LPV-Systemen nicht geeignet. Abgesehen von diesen Studien sind die Verbesserungen der Zustandsrückkopplungsregelung in [24–31]** durch die Verwendung erweiterter LMIs in der Literatur in **Bezug auf Konservativität und/oder Einfachheit bemerkenswerter .** Diese Studien sind weniger konservativ als das **Standarddesign des Zustandsrückkopplungsreglers in [1]. Keine von ihnen ist jedoch für LPV-Systeme und CTS unter TVPU vollständig geeignet, während sie für CTS unter zeitinvarianten parametrischen Unsicherheiten (TIPUs) verbessert sind. Darüber hinaus übertrifft das vorgeschlagene erweiterte LMI TIPU in Bezug auf Konservativität. Dementsprechend ist dieses spezielle Thema in der Literatur insbesondere für CTS mit TVPU- und LPV-Systemen noch offen** Daher werden für LPV-Systeme oder das CTS mit TVPU **gemäß den in der Literatur vorhandenen Ergebnissen weniger konservative Ergebnisse erzielt .**

Feedforward-Kompensatoren oder -Regler werden verwendet, um das Referenzsignal zu verfolgen oder externe oder interne Störungen zu dämpfen, die meist vorhersehbar oder messbar sind, wie in [32,33]. Die Leistung des Systems wird durch ein zusätzlich generiertes Steuersignal entsprechend der gemessenen Störung oder der Referenzwerte verbessert. Insbesondere wurden Störungskompensatoren für viele Systeme realisiert, wie beispielsweise [34]. Ein Störungskompensator wird zusammen mit einem Feedback-Regler verwendet, wie in [35,36], da das Feedback- Design die Anlagenstabilität garantiert. Der Kompensator stabilisiert das instabile System jedoch nicht. Es gibt zwei Hauptmethoden. Die erste besteht darin, dass sie gleichzeitig entworfen werden, wie in [37,38]. Bei der zweiten wird vorab eine bestimmte Leistung mit dem Feedback-Regler erreicht und dann wird der Feedforward-Teil entworfen, wie in [39]. Dieses Dokument basiert auf dem zweiten Ansatz. Viele Feedforward-Kompensatoren oder -Regler wurden für LTI-Systeme wie die in den obigen Studien vorgestellt. Es gibt einige Arbeiten mit der Hÿ- Methode, insbesondere für LPV-Systeme, wie in [40–45]. In [40] wurde **ein Hÿ- Regler reduzierter Ordnung entwickelt, um die Störungen des Schwingungssystems zu berücksichti In [41]** wurde **ein dynamischer Hÿ- Feedforward-Reglerentwurf durch Ermittlung der Systeminverse vorgestellt. In [42] wurde eine geschaltete robuste Feedforward-Steuerung für eine Klasse unsicherer linearer Systeme vorgestellt, die in einer klassischen linearen fraktionalen Transformationsfo Keiner von ihnen ist jedoch für LPV-Systeme geeignet. In [43,44]** werden **statische LPV-Feedforward- Designs verwendet, aber es handelt sich nicht um dynamische Regler. Ein dynamischer Feedforward-Regler wurde für LPV-Systeme in [45] entwickelt, aber viele LMIs existieren, w**

Machine Translated by Google

Bedingungen sind aufgrund der Nichtkonvexität durch die Verwendung des PDLM erforderlich. Darüber hinaus LMIs basieren auf einem gemeinsamen BRL, daher sind LMIs aufgrund der

Multiplikation der Lyapunov-Matrix. Daher wurde der Feedforward-Regler oder **Kompensator basierend auf der L2-** Verstärkung selten diskutiert, insbesondere für LPV-Systeme. Darüber hinaus

ein dynamischer Feedforward-Regler oder Kompensator wurde in den oben genannten Studien, die die Vorteile des erweiterten LMI nutzen. Daher ist der zweite Beitrag

Gegenstand dieses Beitrags ist die Verwendung eines dynamischen LPV-Feedforward-Kompensators auf Basis eines L2-Gain Reglerentwurf, um die Vorteile des erweiterten LMI-Ansatzes zu erreichen, der laut

zur Literaturrecherche wurde bisher nicht in der Literatur berichtet. In [45], das näher an

Zu diesem Papier ist die Studie aufgrund des Auftretens unendlicher LMIs recht komplex.

Darüber hinaus basiert das Design auf dem Standard-BRL, das konservativere Ergebnisse liefert.

Kurz gesagt werden in dieser Studie Lösungen auf der Grundlage nichtiterativer LMIs vorgeschlagen, um ein LPV-Zustandsrückkopplungsregler und ein dynamischer LPV-Feedforward-Kompensator. Sie sind praktisch sowohl für LPV als auch für CTS mit TVPU (lineare Systeme mit parametrischer Unsicherheit).

Die wichtigsten Beiträge des Papiers sind wie folgt:

(1) Für das CTS mit TVPUs wird ein erweiterter BRL auf der Grundlage eines nichtiterativen LMI vorgeschlagen und LPV-System, indem es die Vorteile einer lockeren Variablen und eines erweiterten LMI bietet. Dann Der neue Zustandsrückkopplungsregler wird auf der Grundlage eines neuen äquivalenten BRL entworfen. Mit In Bezug auf seine Überlegenheit hat es das beste Ergebnis unter den jüngsten Studien in [5,24–30] in Ansicht, dass es weniger konservativ ist, wie in Tabelle 1 gezeigt, und eine ungefähr 10% Das beste Ergebnis unter ihnen wurde verbessert.

(2) Alle vorgeschlagenen Lösungen benötigen keine zusätzlichen LMI-Bedingungen wie in [45] , da zum vorgeschlagenen erweiterten BRL, selbst wenn das PDLM verwendet wird, was zu nichtkonvexen oder unendlichen LMIs führt. Sie sind einfacher und weniger konservativ als die üblichen Ansätze in der Literatur und weniger konservativ als ähnliche Studien in der Literatur, wie durch numerische Tests nachgewiesen wurde. Daher bieten die LMIs sowohl Einfachheit als auch geringere Konservativität.

(3) Ein dynamischer LPV-Feedforward-Kompensator auf Basis des erweiterten BRL wird vorgeschlagen. für LPV-Systeme mit messbaren oder beobachtbaren Störungen oder Lärm. Daher Der Kompensator arbeitet online mit den zeitabhängigen parametrischen Unsicherheiten von unsichere Systeme oder die zeitvariablen Parameter für LPV-Systeme.

Tabelle 1. Vergleich der erreichten ÿmin.

	[5]	[24]	[25]	[26]	[27]	[28]	[29]	[30]	Vorgeschlagener Satz 1
ÿmin	1,557	1,498	1,478	1,476	1,251	1,241	1,240	1,241	1.118

2. Vorbemerkungen

In diesem Abschnitt werden einige Definitionen und Lemmata gegeben, um Licht auf die vorgeschlagenen Lemma und Theoreme.

Definition 1. Das in (1) beschriebene System wird als LPV-System bezeichnet. Dabei stellt x(t) **ÿ** ÿnw dar die Zustände, u(t) **ÿ** ÿnu stellt die Steuereingangssignale dar, z(t) **ÿ** ÿnz stellt die gesteuerten Ausgänge, y(t) **ÿ** ÿny stellt die gemessenen Ausgänge dar, ÿ(t) **ÿ** ÿnw stellt die Störungen dar und die gesamten Matrizen haben eine geeignete Größe und sind von den sich mit der Zeit ändernden *ÿ*- Parametern abhängig. Die Parametersätze sind definiert durch (2), wobei ÿR und ÿD die Parametersätze für *ÿ* und seine Ableitung bzw. seine. Zusätzlich ist die Matrix *A(ÿ)* die Zustandsmatrix, die die Beziehung zwischen dem aktuellen Zustand und seiner Änderungsrate, B1(ÿ) ist die Störungseingangsmatrix , B2(ÿ) ist die Steuereingangsmatrix, C1(ÿ) ist die Matrix, die den Zustandsvektor x(t) mit dem geregelten Ausgängen z(t), D11(ÿ) ist die Vorwärtskopplungsmatrix für den Störeingang ÿ(t) direkt Verknüpfung des Störgrößeneingangs ÿ(t) mit dem geregelten Ausgang z(t), D12(ÿ) ist die Vorwärtskopplungsmatrix für den Steuereingang u(t), der den Steuereingang direkt mit dem gesteuerten Ausgang z(t) verbindet, ist C2(ÿ) die gemessene Ausgabematrix, die die Beziehung zwischen dem Zustand und der gemessenen Ausgabe definiert.

Machine Translated by Google

Angenommen, D21(ÿ) ist die Vorwärtskopplungsmatrix für den Störungseingang, die den Störungseingang ÿ(t) direkt mit dem gemessenen Ausgang y(t) verbindet , und D22(ÿ) ist die Vorwärtskopplungsmatrix für den Steuereingang u (t), die den Steuereingang direkt mit dem gemessenen Ausgang y(t) verbindet.

$$\begin{aligned} \dot{x}(t) &= A(\ddot{y}(t))x(t) + B1(\ddot{y}(t))\ddot{y}(t) + B2(\ddot{y}(t))u(t) \\ z(t) &= C1(\ddot{y}(t))x(t) + D11(\ddot{y}(t))\ddot{y}(t) + D12(\ddot{y}(t))u(t) \\ y(t) &= C2(\ddot{y}(t))x(t) + D21(\ddot{y}(t))\ddot{y}(t) + D22(\ddot{y}(t))u(t) \end{aligned} \quad \textbf{(1)}$$

$$\ddot{y}R = \ddot{y}\,\ddot{y}\,\ddot{y}n : \ddot{y}i\,\ddot{y}\,\ddot{y}i\,\ddot{y}\,\ddot{y}i, \ddot{y}i = \mathbf{1, \ldots, \ddot{y}}D = \ddot{y}\,\ddot{y}\,\ddot{y}n : \ddot{y}i, \ddot{y}i = {}^{N}_{1, \ldots, \ddot{y}i\,\ddot{y}\,\ddot{y}i\,\ddot{y}} \quad \textbf{(2)}$$

Darüber hinaus kann die geschlossene Form des Systems wie in (3) dargestellt werden, wobei Z die Ausgänge z **und** y angibt. **Zusätzlich werden die Matrizen A, B,** C **und** D als **Matrizen des geschlossenen Kreislaufs betrachtet .**

$$\begin{aligned} \dot{x}(t) &= A(\ddot{y}(t))x(t) + B(\ddot{y}(t))\ddot{y}(t) \\ Z(t) &= C(\ddot{y}(t))x(t) + D(\ddot{y}(t))\ddot{y}(t) \end{aligned} \quad \textbf{(3)}$$

Dabei wird vorausgesetzt, dass die Parameter online verfügbar (messbar oder berechenbar) sind. Zusätzlich werden die Parameter aus den Zuständen des nichtlinearen Systems abgeleitet, obwohl es normalerweise wie ein unsicheres System aussieht. Wenn die Parameter nicht aus Zuständen abgeleitet werden, im Gegensatz zu denen des LPV-Systems, ist (1) das lineare System unter zeitabhängigen Unsicherheiten wie in Beispiel 1.

Die Hÿ- Norm für das System in (3) kann über die allgemeine LMI-Formulierung berechnet werden , die durch Lemma 1 gegeben ist und in der Literatur als BRL bekannt ist. Dementsprechend für a (0 ÿ ist ein positiver Skalar ÿ) den Leistungsindex durch (4), wobei ÿ(t) **ÿ** L $^{Q}_{2}$ **definieren wir exogene Störung.**

$$J(\ddot{y}, z) = z \quad {}^{T}z\,\ddot{y}\,\ddot{y}\,2\ddot{y}\,T\ddot{y}\,d\ddot{y} \quad \textbf{(4)}$$

Lemma 1. Es sei *G(ÿ)* die stabile Transfermatrix des LPV-Systems in (3) für alle zulässigen I **ÿ** DT (ÿ)D(ÿ) **ÿ** Parametertrajektorien. Genau dann, wenn $\ddot{y}^{2}$ **0,** G(s)ÿ **ÿ** $\dot{\ddot{y}}$ und die folgenden *ÿ* **ÿ** ÿD.
-Aussagen für alle *ÿ(ÿ)* **ÿ** ÿR und **ÿ** äquivalent sind

(ich) Wenn es eine positiv definite differenzierbare symmetrische Matrix *P(ÿ)* **ÿ ÿn×n** gibt , Dann

$$\begin{aligned} &A^{T}(\ddot{y})P(\ddot{y}) + P(\ddot{y})A(\ddot{y}) + \dot{P}(\ddot{y}) + C^{T}(\ddot{y})C(\ddot{y}) \\ &+ B^{T}(\ddot{y})P(\ddot{y}) + DT(\ddot{y})C(\ddot{y}) \quad {}^{T}{}_{2\ddot{y}}\,\text{Ich}\,\ddot{y}\,DT(\ddot{y})D(\ddot{y})^{\ddot{y}1}\,B^{T}(\ddot{y})P(\ddot{y}) + DT(\ddot{y})C(\ddot{y})\,\ddot{y}\,\mathbf{0} \end{aligned} \quad \textbf{(5)}$$

(ii) Wenn es eine positiv definite differenzierbare symmetrische Matrix *P(ÿ)* **ÿ ÿn×n** gibt , Dann

$$\ddot{y}\,{}^{A^{T}(\ddot{y})P(\ddot{y}) + P(\ddot{y})A(\ddot{y}) + B^{T}}_{\ddot{y}\qquad\qquad C(\ddot{y})}\quad \dot{P}(\ddot{y})\;{}^{P(\ddot{y})B(\ddot{y})\;C^{T}(\ddot{y})P(\ddot{y})\;(\ddot{y})}_{2\ddot{y}\ddot{y}\quad D(\ddot{y})\qquad \ddot{y}\text{Ich}}\;\text{Ich DT}\;\ddot{y}\;\ddot{y}\;\ddot{y}\,\mathbf{0} \quad \textbf{(6)}$$

Bemerkung 1. LMIs wie die in (6) sind nicht geeignet, um Regler für lineare Systeme mit polytopischen oder affinen unsicheren Parametern oder LPV-Systeme zu entwerfen, da die Ungleichung die Multiplikation der Lyapunov-Matrix mit den Matrizen der Anlage beinhaltet. Daher führt diese Angelegenheit zu einer nichtkonvexen Optimierung und den unendlichen LMIs. Dennoch wurden einige Methoden vorgeschlagen, um diese Schwierigkeit zu beseitigen, wie die Gittermethode [14], das s-Verfahren [15], die Multikonvexitätsmethode [11] und Methoden in [17,18]. Die Multikonvexitätsmethode und das s-Verfahren führen jedoch zu vielen LMIs, und die Gittermethode ist im Gegensatz zu Trajektorien mit allen zulässigen Parametern nur für lokale Betriebsbedingungen gültig . Darüber hinaus ermöglichen solche Techniken nicht den Entwurf eines Mehrzweckreglers in [12]. Die Verwendung der parameterunabhängigen (festen) Lyapunov-Funktio

Machine Translated by Google

in (7) wie in [5,41,46] ist üblich, führt aber zu sehr konservativen Ergebnissen. Umgekehrt kann man, wenn man die Lyapunov-Funktion von den Anlagenmatrizen isoliert, indem man eine Schlupfvariable verwendet, was aufgrund der zunehmenden Dimension des LMI allgemein als dilatierter oder erweiterter LMI-Ansatz bezeichnet wird, den Konservativismus für Analyse und Design verringern, wie in [23,24,27,31]. Daher ist das Erreichen eines dilatierten Zustands, der dem BRL entspricht, ein effektiver Ansatz, um diese Komplexität in robusten Reglerdesigns wie H2 und **Hÿ zu überwinden.** Dementsprechend wurde in Lemma 3 für das Reglerdesign eine neue Notation für die **Hÿ-** Berechnung einschließlich der TVPU vorgeschlagen.

Für den **L2-Verstärkungsfaktor) für LPV-Systeme wird in der Literatur [5] allgemein der aus Lemma oder Hÿ -Design** (K(ÿ) = *R(ÿ)P* **1 abgeleitete Standard-LMI in (7) verwendet .**

$$\begin{matrix} \ddot{y} & A^{T}(\ddot{y})P + B2R(\ddot{y}) + (\ddot{y}) & & B1(\ddot{y})^{T} & PCT_{1}(\ddot{y}) + R^{T}(\ddot{y})DT_{12} & \ddot{y} \\ \ddot{y} & \ddot{y} & & ich\ \ddot{y}\ddot{y}^{2} & DT\ 11(\ddot{y}) & \ddot{y} \\ \ddot{y} & \ddot{y} & & \ddot{y} & \ddot{y}Ich & \ddot{y} \end{matrix} \quad \ddot{y}\ 0 \qquad (7)$$

Das folgende Lemma 2, das auf dem Projektionslemma basiert, ist nützlich zur Ableitung von die erweiterten LMIs.

Lemma 2. [21]: Wenn es eine symmetrische Matrix **ÿ** und kompatible dimensionale Matrizen **ÿ1, ÿ2** gibt , dann sind die folgenden Bedingungen äquivalent:

i. **ÿ** < **0** und **ÿ** + **ÿ1ÿT** + **ÿ2ÿT** ii. Das folgende $_{2}^{1} < \mathbf{0}$
LMI-Problem ist gemäß Q möglich.

$$\begin{matrix} \ddot{y} & \ddot{y}2 + \ddot{y}1Q \\ \ddot{y}T_{2} + QT\ddot{y}T_{1} & \ddot{y}Q\ \ddot{y}\ QT \end{matrix} \quad \ddot{y}\ 0$$

3. Ein neues dilatiertes beschränktes reelles Lemma

In diesem Abschnitt wird ein erweiterter BRL für LPV und unsichere Systeme vorgeschlagen. Dementsprechend wird eine LMI-basierte Lösung verbessert, um den Hÿ -Zustands-Feedback-Controller zu entwerfen . **Lemma 3 zeigt das erweiterte BRL.**

Lemma 3. Bezeichnet *G(ÿ)* die Transfermatrix des LPV-Systems in (3), dann sind die folgenden beiden Aussagen äquivalent.

ich. Genau dann, wenn die Matrix *A(ÿ)* für alle zulässigen Parametertrajektorien stabil ist und DT (ÿ)D(ÿ) **ÿ** I **ÿ** $_{2}\ddot{y}$ 0, ist die **Hÿ-** Güte für das System in (3) so beschränkt, dass der Index $\ddot{y} > 0$.
Daher *ist ÿGzÿÿÿ* **ÿ** ÿ.

ii. Wenn es eine positiv definite parameterabhängige differenzierbare symmetrische Matrix *P(ÿ)* **ÿ ÿn×n** gibt , eine reelle positive Konstante *µ, die P(ÿ)* **ÿ** 2µP(ÿ) **ÿ 0** erfüllt , und eine kompatible dimensionale Matrix Q, dann gilt das LMI-Problem in (8) für alle *ÿÿ* **ÿ** ÿR **und ÿ** *ÿ* **ÿ** ÿD.

$$\begin{matrix} \ddot{y} & P(\ddot{y})\ \ddot{y}\ 2\mu P(\ddot{y}) & 0 & P(\ddot{y}) + A^{T}(\ddot{y})Q + \mu Q & C^{T}(\ddot{y}) \\ & \ddot{y} & ich\ \ddot{y}\ddot{y}^{2} & B^{T}(\ddot{y})Q & DT\ (\ddot{y}) \quad \ddot{y} \\ & \ddot{y} & \ddot{y} & \ddot{y}Q\ \ddot{y}\ QT & 0 \\ \ddot{y}\ddot{y}\ddot{y} & \ddot{y} & \ddot{y} & \ddot{y} & \ddot{y}Ich\ \ddot{y}\ddot{y}\ddot{y}\ \ddot{y}\ 0 \end{matrix} \qquad (8)$$

Beweis von Lemma 3. Die LPV-Form der Ungleichung in Aussage i **von Lemma 1 ist wie folgt.**

$$\begin{matrix} A^{T}(\ddot{y})P(\ddot{y}) + P(\ddot{y})A(\ddot{y}) + \dot{P}(\ddot{y}) + C^{T}(\ddot{y})C(\ddot{y}) \\ + B^{T}(\ddot{y})P(\ddot{y}) + DT\ (\ddot{y})C(\ddot{y}) \quad {}^{T}{}_{2\ddot{y}}\ Ich\ \ddot{y}\ DT\ (\ddot{y})D(\ddot{y})^{\ddot{y}1}\ B^{T}(\ddot{y})P(\ddot{y}) + DT\ (\ddot{y})C(\ddot{y})\ \ddot{y}\ \mathbf{0} \end{matrix} \qquad (9)$$

Machine Translated by Google

Die Ungleichung in (9) wird wie folgt modifiziert, wenn eine reelle konstante Variable μ in die Ungleichung einbezogen wird .

$$\begin{matrix} A^{T}(\ddot{y}) + \mu I\, P(\ddot{y}) + P(\ddot{y})(A(\ddot{y}) + \mu I) + C^{T}(\ddot{y})C(\ddot{y}) + \dot{P}(\ddot{y})\,\ddot{\mathbf{y}}\, 2\mu P(\ddot{y}) \\ + B^{T}(\ddot{y})P(\ddot{y}) + DT(\ddot{y})C(\ddot{y}) \quad {}_{2\ddot{y}}\,\text{Ich}\,\ddot{\mathbf{y}}\, DT(\ddot{y})D(\ddot{y}) \quad B^{T}(\ddot{y})P(\ddot{y}) + DT(\ddot{y})C(\ddot{y})\,\ddot{\mathbf{y}}\,\mathbf{0} \end{matrix}$$

Durch Verwendung des Schur-Lemmas ergibt sich eine neue Matrixungleichung wie folgt.

$$\begin{matrix} \ddot{y}11(\ddot{y}) & \ddot{y}12(\ddot{y}) \\ \ddot{y}13(\ddot{y}) & \ddot{y}14(\ddot{y}) \end{matrix} \quad \ddot{\mathbf{y}}\,\mathbf{0}$$

Wo

$$\begin{matrix} \ddot{y}11(\ddot{y}) = A^{T}(\ddot{y}) + \mu I\, P(\ddot{y}) + P(\ddot{y})(A(\ddot{y}) + \mu I) + C^{T}(\ddot{y})C(\ddot{y}) + \dot{P}(\ddot{y})\,\ddot{\mathbf{y}}\, 2\mu P(\ddot{y}) \\ \ddot{y}12(\ddot{y}) = P(\ddot{y})B(\ddot{y}) + C(\ddot{y})P(\ddot{y})^{T}(\ddot{y})D(\ddot{y}), \\ \ddot{y}14(\ddot{y}) = DT^{T} + DT(\ddot{y})C(\ddot{y})\ \ddot{y}13(\ddot{y}) = B \\ (\ddot{y})D(\ddot{y})\,\ddot{\mathbf{y}}\,\ddot{y} \quad {}_{2\text{ICH}} \end{matrix}$$

Wir können die Ungleichung wie folgt umschreiben.

$$\begin{matrix} A^{T}(\ddot{y}) + \mu I\, P(\ddot{y}) + \mathbf{(\ddot{y})}(\ddot{y})P(\ddot{y})^{T} \quad P(\ddot{y})B(\ddot{y})\,\mathbf{0} \\ B^{T} \end{matrix} + \begin{matrix} C^{T}(\ddot{y})C(\ddot{y}) + \dot{P}(\ddot{y})\,\ddot{\mathbf{y}}\, 2\mu P(\ddot{y}) \\ DT(\ddot{y})C(\ddot{y}) \end{matrix} \quad \begin{matrix} C^{T}(\ddot{y})D(\ddot{y}) \\ DT(\ddot{y})D(\ddot{y})\,\ddot{\mathbf{y}}\,\ddot{y} \end{matrix} \quad {}_{2\text{ICH}} \quad \ddot{\mathbf{y}}\,\mathbf{0} \qquad \textbf{(10)}$$

Dies wird wie folgt umgeschrieben.

$$\begin{matrix} C^{T}(\ddot{y})C(\ddot{y}) + \dot{P}(\ddot{y})\,\ddot{\mathbf{y}}\, 2\mu P(\ddot{y}) \\ DT(\ddot{y})C(\ddot{y}) \end{matrix} \quad \begin{matrix} C(\ddot{y})\,TD(\ddot{y}) \\ DT(\ddot{y})D(\ddot{y})\,\ddot{\mathbf{y}}\,\ddot{y} \end{matrix} \quad {}_{2\text{ICH}} + \begin{matrix} A^{T}(\ddot{y}) + \mu I \\ B^{T}(\ddot{y}) \end{matrix} \quad P(\ddot{y})\,\mathbf{0} + \quad P(\ddot{y})\,\mathbf{0} \quad \begin{matrix} A(\ddot{y}) \\ +\mu I \end{matrix} \quad B(\ddot{y})\,\ddot{\mathbf{y}}\,\mathbf{0} \textbf{ (11)}$$

Anschließend können wir die Ungleichung (12) erhalten, wenn wir Lemma 2 auf (11) anwenden.

$$\begin{matrix} \ddot{y}11(\ddot{y}) & \ddot{y}12(\ddot{y}) & \ddot{y}13(\ddot{y}) \\ \ddot{\mathbf{y}} & {}_{\ddot{y}} & \ddot{y}22(\ddot{y}) \\ \ddot{\mathbf{y}} & {}_{\ddot{y}} & {}_{\ddot{y}} \end{matrix} \quad \begin{matrix} \ddot{y}23(\ddot{y}) \\ \ddot{\mathbf{y}}33 \end{matrix} \quad \begin{matrix} \ddot{\mathbf{y}} \\ \ddot{\mathbf{y}} \end{matrix} \quad \ddot{\mathbf{y}}\,\mathbf{0} \qquad \textbf{(12)}$$

Wo

$$\begin{matrix} \ddot{y}11(\ddot{y}) = C^{T}(\ddot{y})C(\ddot{y}) + \dot{P}(\ddot{y})\,\ddot{\mathbf{y}}\, 2\mu P(\ddot{y}), \\ \ddot{y}12(\ddot{y}) = C^{T}(\ddot{y})D(\ddot{y}) \\ \ddot{y}13(\ddot{y}) = P(\ddot{y}) + A\ \ddot{y}22(\ddot{y})^{T}(\ddot{y}) + \mu I\,\mathbf{Q}, \\ = DT(\ddot{y})D(\ddot{y})\,\ddot{\mathbf{y}}\,\ddot{y}(\ddot{y})Q,\ \mathbf{\ddot{y}33} = {}_{2\text{ICH}} \\ \ddot{y}23(\ddot{y}) = B^{T}\ddot{y}Q\,\ddot{\mathbf{y}}\,QT \end{matrix}$$

Allerdings muss der Ausdruck auf der linken Seite von (11) gemäß Lemma 2 negativ definit sein. Entsprechend schreiben wir den Ausdruck inklusive $P(\ddot{y})\,\ddot{\mathbf{y}}\,2\mu P(\ddot{y})\,\ddot{\mathbf{y}}\,\mathbf{0}$ wie folgt um.

$$\begin{matrix} \dot{P}(\ddot{y})\,\ddot{\mathbf{y}}\, 2\mu P(\ddot{y}) \\ {}_{\ddot{y}} \end{matrix} \quad \begin{matrix} \mathbf{0} \\ \text{ich}\ {}_{2\ddot{y}\ddot{y}} \end{matrix} + \begin{matrix} C^{T}(\ddot{y}) \\ DT(\ddot{y}) \end{matrix} \quad I\ C(\ddot{y})\ D(\ddot{y})\,\ddot{\mathbf{y}}\,\mathbf{0}$$

Daher erhalten wir die Ungleichung (13) über das Schur-Lemma.

$$\begin{matrix} \dot{P}(\ddot{y})\,\ddot{\mathbf{y}}\, 2\mu P(\ddot{y}) \\ \ddot{\mathbf{y}} \quad {}_{\ddot{y}} \\ \ddot{\mathbf{y}} \quad {}_{\ddot{y}} \end{matrix} \quad \begin{matrix} \mathbf{0}\ C^{T}(\ddot{y}) \\ {}_{2\ddot{y}\ddot{y}}\ \text{Ich}\ DT(\ddot{y}) \\ {}_{\ddot{y}} \quad \ddot{y}\text{Ich} \end{matrix} \quad \begin{matrix} \ddot{\mathbf{y}} \\ \ddot{\mathbf{y}} \end{matrix} \quad \ddot{\mathbf{y}}\,\mathbf{0} \qquad \textbf{(13)}$$

Machine Translated by Google

Schließlich erhalten wir (8), wenn wir das Schur-Lemma anwenden, nachdem wir die Ungleichung (12) wie folgt umgeschrieben haben:

die folgende Form. Darüber hinaus deckt die Bedingung (8) auch die Bedingung (13) ab.

$$\begin{bmatrix} \dot{P}(\ddot{y}) + A^{T}(\ddot{y})Q + \mu Q & C^{T}(\ddot{y}) \\ B^{T}(\ddot{y})Q & D^{T}(\ddot{y}) \end{bmatrix} + \begin{bmatrix} C(\ddot{y}) & D(\ddot{y}) \end{bmatrix} < 0$$

Damit ist der Beweis abgeschlossen. ÿ

Alle Designs können unter Einhaltung einiger Regeln für polytopische oder affine Parameter verwendet werden . In den Beispielen dieses Dokuments werden jedoch affine variierende Parameter verwendet , und daher ändern sich alle zeitabhängigen Matrizen affin, wie beispielsweise die Lyapunov-Matrix in (14). Daher wird im gesamten Dokument affine quadratische Stabilität erreicht, wie in [11].

$$P(\ddot{y}) = P_0 + \sum_{i=1}^{N} \ddot{y}_i P_i, \quad \dot{P}(\ddot{y}) = \sum_{i=1}^{N} \dot{\ddot{y}}_i P_i \tag{14}$$

Theorem 1. Das LPV-System sei in (1) mit dem zeitabhängigen Parametersatz in (2). Wenn es einen positiven reellen Skalar μ gibt , *der* $L(\ddot{y}) > 2\mu L(\ddot{y}) > 0$ erfüllt , die symmetrische positiv definite differenzierbare Matrix $L(\ddot{y})$ mit den kompatiblen Dimensionsmatrizen **R,** $F(\ddot{y})$, dann gilt die LMI in (15) für $\ddot{y}(\ddot{y}) > \ddot{y}R$ und $\dot{\ddot{y}}(\ddot{y}) > \ddot{y}D$. Somit gibt es einen Zustandsrückkopplungs **-Hÿ-** Regler $K(\ddot{y}) = F(\ddot{y})R^{-1}$

$$\begin{bmatrix} \dot{L}(\ddot{y}) + A(\ddot{y})R + B_2F(\ddot{y}) + \mu R & B_1(\ddot{y}) \\ C_1(\ddot{y})R + D_{12}F(\ddot{y}) & D_{11}(\ddot{y}) \end{bmatrix} + \begin{bmatrix} \dot{\ddot{y}}_1 & \dots & \dot{\ddot{y}}_{ich} \\ 0 & & \end{bmatrix}^{T} \begin{bmatrix} \ddot{y}_1 & \dots & \ddot{y}_{ich} \\ 0 & & \end{bmatrix} > 0 \tag{15}$$

Beweis von Theorem 1. Wir haben die Closed-Loop-Matrizen in (16) wie in (17) erhalten, wenn der Regler u = $K(\ddot{y})x$ **im System (1) eingesetzt wird, wobei der Ausgang y(t) eliminiert wird, da er als Zustandsrückkopplung betrachtet wird.**

$$\dot{x}_{cl}(t) = A_{cl}(\ddot{y}(t))x_{cl}(t) + B_{cl}(\ddot{y}(t))\ddot{y}(t), \quad z_{cl}(t) = C_{cl}(\ddot{y}(t))x_{cl}(t) + D_{cl}(\ddot{y}(t))\ddot{y}(t) \tag{16}$$

$$A_{cl}(\ddot{y}) = A(\ddot{y}) + B_2K(\ddot{y}), \; B_{cl}(\ddot{y}) = B_1(\ddot{y}), \quad C_{cl}(\ddot{y}) = C(\ddot{y}) + D_{12}K(\ddot{y}), \; D_{cl}(\ddot{y}) = D_{11}(\ddot{y}) \tag{17}$$

Ersetzt man die Zustandsraummatrizen in (17) durch (8), so erhält man folgende Ungleichung.

$$\begin{bmatrix} \dot{P}(\ddot{y}) + Q^{T}(A(\ddot{y}) + B_2K(\ddot{y})) + \mu Q^{T}B_1(\ddot{y}) & Q^{T}C(\ddot{y}) + D_{12}K(\ddot{y}) \\ D_{11}(\ddot{y}) & 0 \end{bmatrix} + \begin{bmatrix} \dot{\ddot{y}}_1 & \dots & \dot{\ddot{y}}_{ich} \\ 0 & & \end{bmatrix}^{T} \begin{bmatrix} \ddot{y}_1 & \dots & \ddot{y}_{ich} \\ 0 & & \end{bmatrix} < 0$$

Es gibt jedoch Multiplikationen unbekannter Matrizen. Um das LMI zu erhalten, wenden wir eine Kongruenztransformation auf die Ungleichung an, indem wir die Transponierte von (18) bzw. (18) multiplizieren . Somit ergibt sich (15), wobei $\dot{L}(\ddot{y}) = Q^{-1}\dot{P}(\ddot{y})Q^{-1}$, $L(\ddot{y}) = Q^{-1}P(\ddot{y})Q^{-1}$, **Q^-1 ,** $F(\ddot{y}) = R = K(\ddot{y})R$. **Damit ist der Beweis abgeschlossen.** ÿ

$$\text{Diagnose}\left(Q_1^{-1}, \dots, Q_{ich}^{-1}, \dots\right) \tag{18}$$

Bemerkung 2. Die vorgeschlagene Lösung bietet drei wesentliche Vorteile. Erstens kann die Lyapunov-Matrix im Gegensatz zu (7) in Form einer parameterabhängigen Matrix vorliegen, ohne dass zusätzliche Ansätze wie

Machine Translated by Google

wie das S-Verfahren, Gitter und Multikonvexität, was die Konservativität verringert. Zweitens ist die Reglerverstärkung im Gegensatz zum üblichen Design wie in (7) in [5] unabhängig von der Lyapunov-Matrix, die durch die symmetrische Eigenschaft eingeschränkt ist. Stattdessen ändert sie sich mit einer nichtsymmetrischen Matrix R für das vorgeschlagene LMI in Theorem 1. Drittens führt die reale Schlupfvariable μ sogar bei großen Parameterbereichen zu machbaren Ergebnissen, da eine große Vergrößerung des Bereichs den $\ddot{y}$- Leistungswert weniger beeinflusst als ähnliche Werte in Bezug auf Machbarkeit und Leistung, wie die numerischen Ergebnisse belegen.

Wenn wir CTS mit TVPU betrachten, können wir das LMI für einen parameterunabhängigen (statischen oder festen) Reglerentwurf problemlos wie in Theorem 1 ändern, vorausgesetzt, dass die unsicheren Parameter zum Entwurfszeitpunkt unbekannt und nicht durch Online-Messungen bekannt, aber als begrenzt bekannt sind. Um die Form des LMI zu erhalten, werden also einige Matrizen in parameterabhängige Matrizen umgewandelt, während andere in parameterunabhängige Matrizen umgewandelt werden . Zudem kann das Theorem durch bestimmte Vorschriften für zeitabhängige Parameter oder Unsicherheiten vom affinen oder polytopen Typ verwendet werden. In diesem Dokument werden jedoch durchgängig zeitabhängige Parameter vom affinen Typ für die Entwürfe berücksichtigt. Für den vorgeschlagenen Zustandsrückkopplungsregler wurden zwei numerische Beispiele realisiert. Alle numerischen Berechnungen wurden mit dem Yalmip-Parser [47] und dem Sedumi-Solver [48] durchgeführt, die auf einem PC (2 Kerne, 2,8 GHz CPU mit 4 GB RAM) verwendet wurden. Beispiel 1 stellt einen Reglerentwurf für das CTS mit der TIPU dar, während Beispiel 2 eine verstärkungsgesteuerte Regelung für das CTS mit TVPU darstellt.

Beispiel 1. Es wird der Zustandsrückkopplungsregler für den Satelliten in [25,49] behandelt. Das System entsteht aus zwei starren Körpern, die durch eine flexible Verbindung verbunden sind. Das System ist wie in (19). Die unsicheren Parameter werden als Dämpfung f und Feder k modelliert. Der Bereich von k beträgt [0,09 0,4], während der Bereich von f [0,0038 0,04] beträgt. Die Leistungsfähigkeit der Methoden in [5,24–30] und der vorgeschlagenen Lösung in Theorem 1 wird in Bezug auf das Niveau $\ddot{y}$ verglichen , das direkt mit der Konservativität zusammenhängt. Die Ableitungen der Parameter sind Null; das heißt, die Unsicherheiten sind nicht differenzierbar.

$$\ddot{y}\ \dot{\ddot{y}}_1\ \ddot{y}\quad \ddot{y}\ \begin{matrix}0&0&1&0\\0&0&0&1\end{matrix}\quad \text{ykk yffk ykf yf}$$

$$\ddot{y}\ \ddot{y}\ \dot{\ddot{y}}_1\ \ddot{y}\quad \ddot{y}\ \begin{matrix}0\\0\\0\\1\end{matrix}\ \ddot{y}\quad \ddot{y}\ \begin{matrix}0\\0\\1\\0\end{matrix}\ \ddot{y}$$

$$\dot{\ddot{y}}_1\ \dot{\ddot{y}}_2\ \dot{\ddot{y}}_3\ \dot{\ddot{y}}_4\ \ddot{y}\ \ddot{y}\ \ddot{y}\ \ddot{y} = \ddot{y}\ \ddot{y}\ \ddot{y}\ \ddot{y}$$

$$\ddot{y}\ \dot{\ddot{y}}_1\ \dot{\ddot{y}}_2\ \dot{\ddot{y}}_1\ \dot{\ddot{y}}_2\ \ddot{y}\ \ddot{y}\ \ddot{y}\ \ddot{y} + \ddot{y}\ \ddot{y}\ \ddot{y}\ \ddot{y}\ \dot{\ddot{y}} + \ddot{y}\ \ddot{y}\ \ddot{y}\ \ddot{y}\

Machine Translated by Google

die TVPU. Schließlich stellt Beispiel 2 nur einen Vergleich zwischen diesem Papier und dem Standard dar Methode in [5] , **da die Studien in** Tabelle 1 mit Ausnahme des Standarddesigns in [1] nicht vollständig gültig sind für die TVPU. Die Leistung wird getestet, indem der unsichere Satz für die TVPU in Beispiel 2 erweitert wird.

Beispiel 2. In diesem Beispiel wird die Gain Scheduling LPV-Steuerung unter Verwendung des Theorems 1 für das zufällig generierte LPV-System in (20), wobei die Bereiche der zeitabhängigen Parameter ÿ1 **und** ÿ2 **jeweils** [0,01, 2] und [0,5, 1,5] betragen. Zusätzlich sind die Variationsraten für die Parameter **ÿ1** und **ÿ2** sind [0, 3] bzw. [0, 1].

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} \ddot{y}\ddot{y}_1 & \ddot{y}_1 & 0{,}8 \\ 0 & 0{,}2\,\ddot{y}\ddot{y}_2 & \ddot{y}_1 \\ 0{,}5\,\ddot{y}_2 & 0{,}1 & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0{,}1 \\ 0{,}2 \\ 0 \end{bmatrix} \ddot{y}\ddot{y} + \begin{bmatrix} 1 \\ 0 \\ 0{,}1 \end{bmatrix} \ddot{y}u, \quad z = \begin{bmatrix} 100 \\ 0\,0\,0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \ddot{y} + \begin{bmatrix} 0 \\ 0{,}8 \end{bmatrix} \text{Sie} \tag{20}$$

Wie in Tabelle 2 zu sehen, nimmt die Verbesserung erheblich zu, wenn der unsichere Satz größer wird verglichen mit dem gemeinsamen Satz. Beispielsweise verringert die Methode den ÿ **-Pegel mit einer Rate von 78,59 % für ÿ1 = [0,01 50], ÿ2 = [0,1 190].**

Tabelle 2. Die ÿmin- Leistungswerte für TVPU.

	ÿ- **Werte**								
	ÿ1= [0,01 2], ÿ2= [0,5 1,5]			**ÿ1= [0,01 5], ÿ2= [0,1 19]**			**ÿ1= [0,01 50], ÿ2= [0,1 190]**		
[5]	**1,0280**			**6.3151**			**22.745**		
Satz 1	**0,890** (μ **= 3)** **14,42 % verringert**			**2,2466** (μ **= 2,5)** **64,42 % verringert**			**4,869** (μ **= 3,5)** **78,59 % verringert**		
	Slack-Variablensätze für Machbarkeit								
Satz 1	μ **= (0, 9 × 105]**			μ **= (0, 8 × 105]**			μ **= (0, 8 × 105]**		
	Anzahl der Iterationen – Anzahl der Variablen – Solverzeit								
[5]	**15**	**18**	**0,0174 s**	**28**	**18**	**0,126 s**	**49**	**18**	**0,215 s**
Satz 1	**39**	**41**	**9,556 s**	**41**	**41**	**9,904 s**	**56**	**41**	**13,503 s**

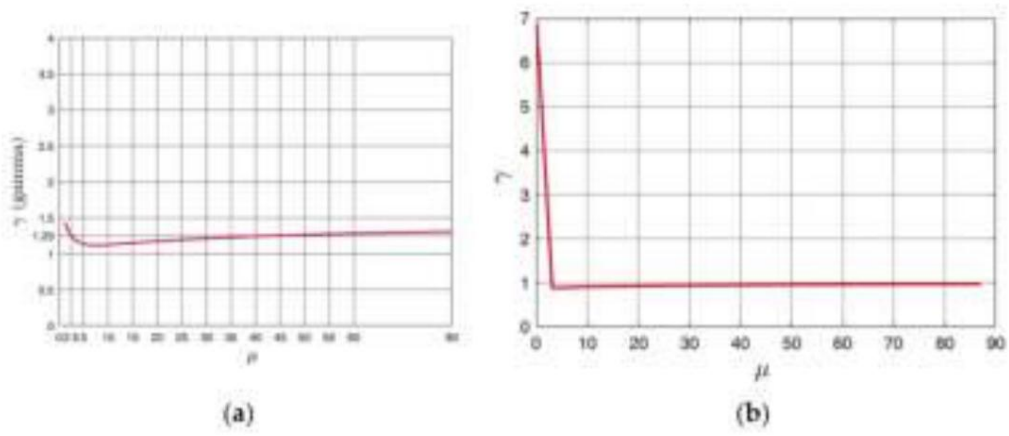


Abbildung 1. Die Beziehungen zwischen ÿ- **Niveau und Slack-Variable: (a) Beispiel 1; (b) Beispiel 2.**

Darüber hinaus ist die Beziehung zwischen μ **und** ÿ **in Abbildung 1b dargestellt . Daher ist** ÿ **Der Pegel wird fast nie durch Änderung** *von* μ **nach einem bestimmten Wert geändert, daher ist seine Bestimmung einfach, sodass die vorgeschlagene Lösung eine bessere Leistung aufweist als die übliche Lösung auch wenn** μ **zwischen 1 und 85 liegt.**

Machine Translated by Google

In Bezug darauf werden die Ergebnisse für die vorgeschlagene Lösung in Theorem 1 und den gängigen Entwurf in [5] erhalten, bei dem die Lyapunov-Matrix fixiert ist. Für ÿ1 = [0,01 2], ÿ2 = [0,5 1,5] ergibt das übliche Design ein ÿ- **Niveau von 1,028, während die vorgeschlagene Methode 0,890 für** μ **= 3 ergibt. Dementsprechend betragen die Rückkopplungsgewinne des Standarddesigns und des Designs in Theorem 1 K0 = 103 [ÿ23,383 170,568 ÿ269,482], K1 = [0,30251 ÿ2,2045 3,4833], K2 = [0,11503 ÿ0,84184 1,33] und K0 = [ÿ12,17 74,915 ÿ102,05], K1 = [0,9135 ÿ3,0938 3,5387], K2 = [5,306 ÿ40,69 47,744]. Die Beziehung zwischen** μ **und** ÿ **ist in Abbildung 1b dargestellt.**

Die Ergebnisse lassen darauf schließen, dass die vorgeschlagene erweiterte LMI-Lösung für das Gain-Scheduling-Design weniger konservativ ist als die übliche LMI in [5]. Allerdings weist sie Nachteile auf, wie eine größere Anzahl von Iterationen, Variablen und einer längeren Lösungszeit. Die Machbarkeit wird für den großen unsicheren Satz in Tabelle 2 ermittelt, und die Schlupfvariable μ **liefert durchführbare Ergebnisse für einen großen Satz (0, 8 × 105). Obwohl die Anzahl der LMI-Variablen entsprechend der Standardvariable in [5] von 18 auf 41 erhöht wurde, wurde eine Verbesserung von bis zu 78,59 % erreicht. Schließlich werden für die TVPU eine gute Verbesserung und weniger konservative Ergebnisse erreicht, selbst wenn die erweiterten LMIs teurer sind als die Standard-LMI. Insbesondere werden die Vergleiche im Beispiel für die parameterunabhängige Ljapunow-Matrix in [5] angestellt. Wie in Anmerkung 1 hervorgehoben, werden bei Verwendung von PLDF Methoden wie die Multikonvexitätsmethode [11] und Methoden in [17,18] verwendet. Daher erhöhen sich die Standard-LMI-Variablen und die Komplexität von [5] erheblich. Beispielsweise muss eine weitere LMI-Bedingung derselben Größe hinzugefügt werden. Daher wird die vorgeschlagene LMI-Be**

Zwei Beispiele haben gezeigt, dass die vorgeschlagene Lösung weniger konservative Ergebnisse liefert. Darüber hinaus kann sie für das LPV-System und das unsichere System unter TVPU vollständig verwendet werden, die anderen jedoch nicht, mit Ausnahme des gemeinsamen LMI in [5]. Andererseits erhöhen zunehmende Grenzen für zeitabhängige Unsicherheiten oder Parameter die Konservativität der LMIs. Das auf dem vorgeschlagenen LMI basierende Design erzielt jedoch **neben einer einfacheren Struktur weniger** konservative und erfolgreiche

4. Entwurf eines dynamischen Störungskompensators für LPV-Systeme

Das Blockdiagramm des Störgrößenkompensators ist in Abbildung 2 dargestellt , **wobei NCF der Normalisierungskoeffizient bzw. Filter ist. Hier ist der dynamische Feedforward-Kompensator wie in (21).**

$$\begin{aligned} \dot{x}c(t) &= Ac(ÿ)xc(t) + Bc(ÿ)ÿ(t) \\ uc(t) &= Cc(ÿ)xc(t) + Dc(ÿ)ÿ(t) \end{aligned} \quad \textbf{(21)}$$

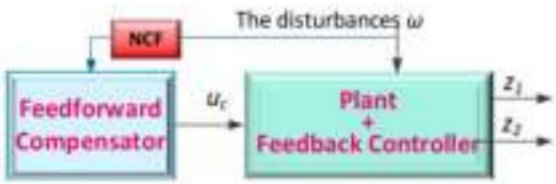


Abbildung 2. Blockdiagramm des Störungskompensators.

Dementsprechend stellt Theorem 2 einen dynamischen Feedforward-Störungskompensatorentwurf gemäß dem erweiterten LMI in Lemma 3 dar.

Theorem 2. Lassen Sie ein stabiles LPV-System in (1) für alle zulässigen Trajektorien von Parametern vorliegen, deren Satz wie in (2) ist. Wenn es einen reellen Skalar $\mu > 0$, die positiv definite differenzierbare symmetrische Matrix *Y(ÿ)* und die kompatiblen Matrizen **M,** *N(ÿ) gibt,* dann gelten die LMIs in (22) und (23) mit den Definitionen in (24) für *ÿ(ÿ)* **ÿ** ÿR und **ÿ** *(ÿ)* **ÿ** ÿD. In einem solchen Fall ergibt sich aus (25) ein dynamischer **Hÿ** -Feedforward- Kompensator in (21).

Machine Translated by Google

$$\begin{matrix} \ddot{y}11(\ddot{y}) & \ddot{y}12(\ddot{y}) & \ddot{y}13(\ddot{y}) & \ddot{y}14(\ddot{y}) & \ddot{y}22(\ddot{y}) & \ddot{y}23(\ddot{y}) \\ \ddot{y} & & \ddot{y}24(\ddot{y}) & \ddot{y}33(\ddot{y}) & \ddot{y}34 & \ddot{y}44 \\ \ddot{y} & & \ddot{y} & & & \\ \ddot{y}\ddot{y}\ddot{y} & \ddot{y} & & \ddot{y} & \ddot{y} & \ddot{y}\ddot{y}\ddot{y}\ddot{y} \ddot{y} 0 \end{matrix} \quad \ddot{y} \quad \ddot{y} \quad (22)$$

$$Y(\ddot{y}) = \begin{matrix} Y1(\ddot{y}) & Y2(\ddot{y}) \\ \ddot{y} & Y3(\ddot{y}) \end{matrix} \quad \ddot{y} \, 0 \quad (23)$$

$$\begin{matrix} \ddot{y}11(\ddot{y}) = & \dot{Y}1(\ddot{y}) \ddot{y} 2\mu Y1(\ddot{y}) & \dot{Y}2(\ddot{y}) \ddot{y} 2\mu Y2(\ddot{y}) \\ \ddot{y} & & \dot{Y}3(\ddot{y}) \ddot{y} 2\mu Y3(\ddot{y}) \\ \ddot{y}12(\ddot{y}) = & 0 \, Y1(\ddot{y}) + R_1^T A^T(\ddot{y}) + NT(\ddot{y}) B_2^T + \mu R_1^T & \\ & 0 \, Y1(\ddot{y}) + R_1^T A^T(\ddot{y}) + NT(\ddot{y}) B_2^T + \mu R_2^T(\ddot{y}) & \\ \ddot{y}13(\ddot{y}) = & Y2(\ddot{y}) + MT(\ddot{y}) R_1^T C_1^T(\ddot{y}) + NT(\ddot{y}) DT_{12} & \\ & Y3(\ddot{y}) + MT(\ddot{y}) R(\ddot{y}) R_1^T C_1^T(\ddot{y}) + NT(\ddot{y}) DT_{12} & \\ \ddot{y}14(\ddot{y}) = & R_1^T C_2^T + NT(\ddot{y}) DT(\ddot{y}) + \ddot{y}_{22} & , \ddot{y}22(\ddot{y}) = \ddot{y}_{2\ddot{y}\ddot{y}} \, I \, B1(\ddot{y}) + DT(\ddot{y}) B_2^T \\ & R_1^T C_2^T NT(\ddot{y}) DT_{22} & \ddot{y}R1 \ddot{y} R_1^T \\ \ddot{y}23(\ddot{y}) = & B_F^T(\ddot{y}) \quad DT \, 11(\ddot{y}) + D_F^T(\ddot{y}) DT_{12} & , \ddot{y}24(\ddot{y}) = DT \, 21(\ddot{y}) + D_F^T(\ddot{y}) DT_{22} \\ & \ddot{y}R1 \ddot{y} R_2^T(\ddot{y}) \, 0 & 0 \\ \ddot{y}33(\ddot{y}) = & \ddot{y}R2(\ddot{y}) \ddot{y} R_2^T(\ddot{y}), \, 0 \quad \ddot{y}lch & , \ddot{y}34 = \begin{matrix} 0 \\ 0 \end{matrix} , \ddot{y}44 = \ddot{y}l \end{matrix} \quad (24)$$

$$\begin{aligned} Ac &= M(\ddot{y}) R^{\ddot{y}1}(\ddot{y}), \; Bc(\ddot{y}) = BF(\ddot{y}), \\ Cc &= N(\ddot{y}) R_2^{2\,\ddot{y}1}(\ddot{y}), \; Dc = DF(\ddot{y}) \end{aligned} \quad (25)$$

Beweis von Theorem 2. Der LMI in (8) wird mit der folgenden Matrix bzw. der Transponierten der Matrix von der linken–rechten Seite multipliziert, wenn wir die Kongruenztransformation anwenden .

$$\begin{matrix} \ddot{y} & Q\ddot{y}1 \, 0 \, 0 \, 0 & \\ & 0 \quad \text{ich} \; 0 \, 0 & \ddot{y} \\ & 0 \quad 0 \, Q\ddot{y}1 \, 0 & \\ \ddot{y}\ddot{y}\ddot{y} & 0 \quad 0 \, 0 \, \text{Ich} & \ddot{y}\ddot{y}\ddot{y} \end{matrix}$$

Daher wird unter den Definitionen in (27) folgender LMI ermittelt:

$$\begin{matrix} \ddot{y} & \dot{Y}(\ddot{y}) \ddot{y} \, 2\mu Y(\ddot{y}) & \ddot{y} & \ddot{y} & \ddot{y} & \\ & 0 \quad \text{ich} \; \ddot{y}\ddot{y}^2 & \ddot{y} & \ddot{y} & \ddot{y} & \\ & Y(\ddot{y}) + A(\ddot{y})R + \mu R \; B(\ddot{y}) \; \ddot{y}R \ddot{y} \, R^T & & & & \\ \ddot{y}\ddot{y}\ddot{y} & C(\ddot{y})R \quad D(\ddot{y}) & 0 & \ddot{y}lch & \ddot{y}\ddot{y}\ddot{y}\ddot{y} \ddot{y} 0 & \end{matrix} \quad (26)$$

$$Y(\ddot{y}) = Q \, \ddot{y}TP(\ddot{y})Q \quad \ddot{y}1 \;, \dot{Y}(\ddot{y}) = Q^{\ddot{y}T} \dot{P}(\ddot{y}) Q^{\ddot{y}1} \;, R = Q \quad \ddot{y}1 \quad (27)$$

Die Zustandsraumform von ÿ(t) **bis** Z(t) **ist wie in (28), wenn der Kompensator in (21) zu (1) realisiert wird. Dabei stellt der Vektor** Z **die Ausgänge** z **und y dar, der Vektor** X stellt **die Zustände** x **in (1) und xc in (21) dar und ÿ** *ist* **das Störsignal.**

$$\begin{matrix} \dot{X}(t) \\ Z(t) \end{matrix} = \begin{matrix} Acl(\ddot{y}) & Bcl(\ddot{y}) \\ Ccl(\ddot{y}) & Dcl(\ddot{y}) \end{matrix} \begin{matrix} x(t) \\ \ddot{y}(t) \end{matrix} \quad (28)$$

Entsprechend ergeben sich mit der Realisierung die Zustandsraummatrizen in (28) wie in (29). Daher wird der optimale Regler durch Minimierung der L2- Verstärkung von *ÿ* **bis Z** erreicht.

$$\begin{matrix} Acl(\ddot{y}) = & A(\ddot{y}) \, B2Cc(\ddot{y}) \, 0 & , Bcl(\ddot{y}) = & B1(\ddot{y}) + B2Dc(\ddot{y}) & , \\ & & & \text{Chr.} & \\ Ccl(\ddot{y}) = & C1(\ddot{y}) \, D12Cc(\ddot{y}) & & D11(\ddot{y}) + D12Dc(\ddot{y}) & \\ & C2(\ddot{y}) \, D22Cc(\ddot{y}) \;, Dcl(\ddot{y}) = & & \mathbf{D21} + D22Dc(\ddot{y}) & \end{matrix} \quad (29)$$

Machine Translated by Google

Definieren wir die Lyapunov-Matrix wie in (23), und die *R(ÿ)* **-Matrix wird wie folgt definiert.**

$$R(\ddot{y}) = \begin{matrix} \mathbf{R1} & \mathbf{R1} \\ R2(\ddot{y}) & R2(\ddot{y}) \end{matrix}$$

Wenn man dementsprechend die Lyapunov-Matrix *Y(ÿ)* **in (23), ihre Ableitung, die** *R(ÿ)* **-Matrix und die Zustandsraummatrizen in (29) in (26) anstelle von** *A(ÿ), B(ÿ), C(ÿ)* **und** *D(ÿ) einsetzt,* erhält **man eine Matrixungleichung. Wenn man einige Ausdrücke wie in (30) definiert, erhält man die LMI in (22) mit den Definitionen in (24) und der Beweis ist abgeschlossen.** ÿ

$$M(\ddot{y}) = Ac(\ddot{y})R2(\ddot{y}),\ N = Cc(\ddot{y})R2(\ddot{y}),$$
$$BF(\ddot{y}) = Bc(\ddot{y}),\ DF(\ddot{y}) = Dc(\ddot{y}) \quad \textbf{(30)}$$

Anmerkung 4. Die vorgeschlagenen LMIs bieten vier wesentliche Vorteile. Einige davon werden in Anmerkung 2 erwähnt. Erstens kann die Lyapunov-Matrix in der LPV-Form bezeichnet werden, was den Konservativismus reduziert, ohne zusätzliche Ansätze wie Gitter, S-Prozedur, Multikonvexität und die Bedingungen in [17,18]. Beispielsweise gibt es in [45] viele LMIs , da aufgrund dieser Produkte Multikonvexitätsbedingungen verwendet werden. Darüber hinaus basieren die LMIs auf der gemeinsamen BRL in (6), sodass die LMIs aufgrund der Multiplikation der Lyapunov-Matrix etwas konservativ sind. Die vorgeschlagene erweiterte BRL ist jedoch die beste der verfügbaren, wie in Tabelle 1 gezeigt. Zweitens treten aufgrund der Produkte von Systemmatrizen mit einer Lyapunov-Matrix wie in (6) unendliche LMIs und nichtkonvexe Optimierung auf, wenn die Lyapunov-Matrix in der LPV-Form vorliegen kann. Ein solcher Fall tritt jedoch für die vorgeschlagene LMI-Formulierung in (8) nicht auf. Somit ist klar, dass die vorgeschlagene Lösung einfacher ist.
Drittens sind die von PDLF wie in [45,50] entworfenen LPV-Regler im Allgemeinen von der Ableitung der Ljapunow-Matrix abhängig (Matrix Ac(ÿ) des Reglers). Die Ableitungsoperation verursacht jedoch in praktischen Echtzeitanwendungen Rauschprobleme. Die vorgeschlagenen Entwürfe sind nicht von der Ableitung der Ljapunow-Matrix abhängig. Schließlich sind die Reglermatrizen nicht von der Ljapunow- Matrix abhängig, die in üblichen Entwürfen durch die Symmetriebedingung eingeschränkt ist. Stattdessen sind die Matrizen in den vorgeschlagenen Entwürfen von einer vollständigen Matrix abhängig. Darüber hinaus hat sie eine einfache Struktur.

Anmerkung 5. Der vorgeschlagene Kompensator enthält keinen Rückkopplungsregler. Es ist bekannt, dass eine Feedforward-Methode angewendet wird, um die Referenzverfolgungsleistung zu verbessern oder Störungen oder Rauschen zu verringern, und dass sie die Stabilität nicht beeinträchtigt, sodass ein Rückkopplungsregler erforderlich ist. Für einige Anlagen, wie beispielsweise viele mechanische und elektrische Systeme, ist jedoch kein Rückkopplungsdesign erforderlich, da es sich um stabile Systeme handelt. Daher kann der Kompensator neben stabilisierten Systemen mit einem Rückkopplungsregler auch für stabile Systeme mit Störungen verwendet werden. Darüber hinaus kann er verwendet werden, um die Referenzverfolgung zu verbessern, indem Referenzsignale anstelle von Störungseingängen für die stabilisierten Systeme mit einem Rückkopplungsregler ausgetauscht werden. Dementsprechend wird für alle Beispiele im gesamten Dokument angenommen, dass die Systeme stabil oder mit einem Rückkop
Die interne Stabilität des LPV-Feedforward-Reglers/Kompensators muss jedoch für alle zeitabhängigen Parameter gewährleistet sein, wie in den obigen Theoremen. Andernfalls wird das stabile System instabil. Diese Stabilität wird durch die Zerlegung der Lyapunov-Matrix erreicht, ohne dass für alle vorgeschlagenen Theoreme eine zusätzliche Stabilitätsmethode erforderlich ist .

5. Simulationsstudien

In diesem Abschnitt wurde ein numerisches Beispiel für die Auslegung eines Kompensators realisiert. Alle numerischen Optimierungsberechnungen wurden mit Yalmip Version R20210331 [47] und Sedumi Version 1.30 [48] durchgeführt, die auf einem PC (2 Kerne 2,8 GHz CPU mit 4 GB RAM) verwendet wurden .

Beispiel 3. Dieses Beispiel zeigt den Entwurf des LPV-Kompensators für das zufällig erzeugte LPV-System in (31). Die zeitabhängigen Parameter ÿ1(t) und ÿ2(t) sind durch (32) begrenzt. Obwohl es normalerweise als unbestimmtes lineares System erscheint, wird es als LPV-System betrachtet, wenn angenommen wird

Machine Translated by Google

dass die messbaren oder berechenbaren Parameter Zustände des nichtlinearen Systems sind. Ein LPV-System wird im Gegensatz zu einem unsicheren linearen System aus einem nichtlinearen System abgeleitet.

$$\dot{x}(t) = x(t) + \begin{bmatrix} 1 + \ddot{y}1(t) & \ddot{y}3 \\ \ddot{y}4 & \ddot{y}2(t) \end{bmatrix} z(t) = \begin{bmatrix} 1 & 0,1 \\ (0,01)\ddot{y}(t) & \end{bmatrix} x(t) + \begin{bmatrix} 0,5 + \ddot{y}1(t) \\ 0,8 \end{bmatrix} \ddot{y}(t) + \begin{bmatrix} 0,2 \\ 0,1 \end{bmatrix} u(t)$$
$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) \quad (31)$$

$$0,1 \ddot{y} \ddot{y}1(t) \ddot{y} 0,7 \text{ mit } \ddot{y} 0,445 \ddot{y} . \ddot{y}1(t) \ddot{y} 1,82$$
$$1 \ddot{y} \ddot{y}2(t) \ddot{y} 5 \text{ mit } \ddot{y} 11,7 \ddot{y} . \ddot{y}2(t) \ddot{y} 2,295 \quad (32)$$

Wenn das System im Hinblick auf zeitabhängige Parameter analysiert wird, erhält man die Ergebnisse in den Abbildungen 3 und 4. Die verschiedenen zeitabhängigen Parameter werden in Abbildung 3 zur Analyse getestet. Einer davon wird jedoch für die Simulation verwendet und befindet sich in Abbildung 3b. Abbildung 3a wird nur für die Analyse verwendet (Bode-Diagramm und Polabbildung), da angenommen wird, dass die Parameteränderung linear ist. Normalerweise ist es unwahrscheinlich, dass beide Parameter linear variieren. Die Parameteränderung in Abbildung 3b tritt während der Simulation auf. Ziel ist es, die Reaktionen auf verschiedene Parameteränderungen zu beobachten. Für die Analyse wird das Bode-Diagramm des Systems in Abbildung **4 mit einer Abtastzeit von Abbildung 4a zeigt das Bode-Diagramm des Systems unter Annahme der zeitabhängigen Parameter in Abbildung 3a, während Abbildung 4b das Bode-Diagramm unter Annahme der zeitabhängigen Parameter in Abbildung 3b zeigt. Die Abbildung zeigt, dass die zeitabhängigen Parameter die Nichtlinearität des Systems gemäß Bode-Diagrammen der Ausgänge von den Eingänge Aus den Bode-Diagrammen geht hervor, dass die Störung bei etwa 3–6 rad/s wirksamer ist. Daher haben die Störtestsignale eine niedrige Frequenz.**

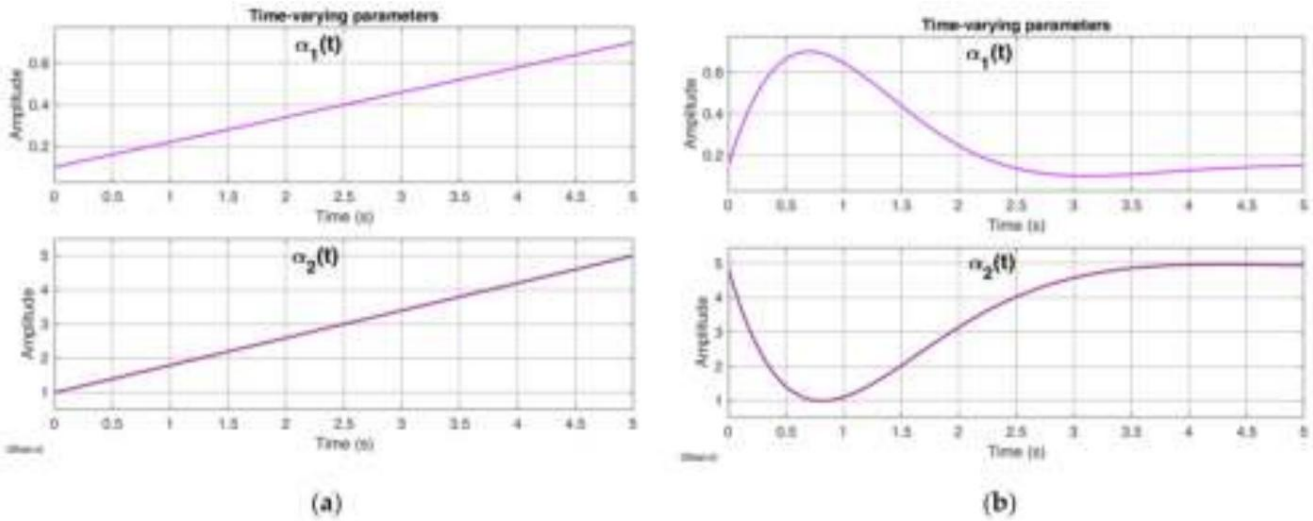


Abbildung 3. Die zeitabhängigen Parameter für: (a) Analyse; (b) Analyse und Simulation.

In ähnlicher Weise zeigt Abbildung 5 die Polkarte des Systems für die verschiedenen zeitabhängigen Parameter in Abbildung 3b, wobei die Abtastzeit 50 ms beträgt. Aus der Polkarte lassen sich zwei Fälle ableiten. Die Pole der gesamten Übertragungsmatrizen des Systems befinden sich in der linken Hälfte der Ebene. Dieser Fall bestätigt die Stabilität des Systems in Übereinstimmung mit allen in Anmerkung 5 genannten zeitabhängigen Parametern. Andernfalls kann der Kompensator nicht angewendet werden, da er die Stabilität nicht beeinflusst. Der andere Fall ist, dass die Parameter das System stark beeinflussen .

Machine Translated by Google

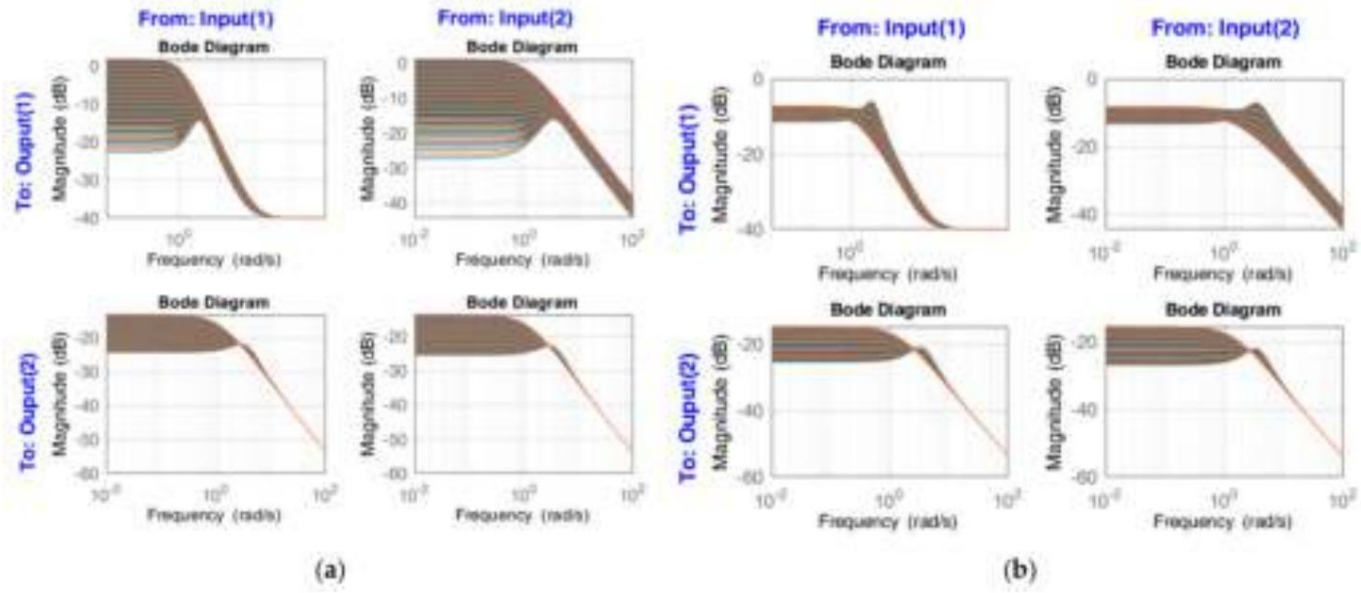


Abbildung 4. Bode-Diagramme des Systems für zeitabhängige Parameter in: (a) Abbildung 3a; (b) Abbildung 3b.

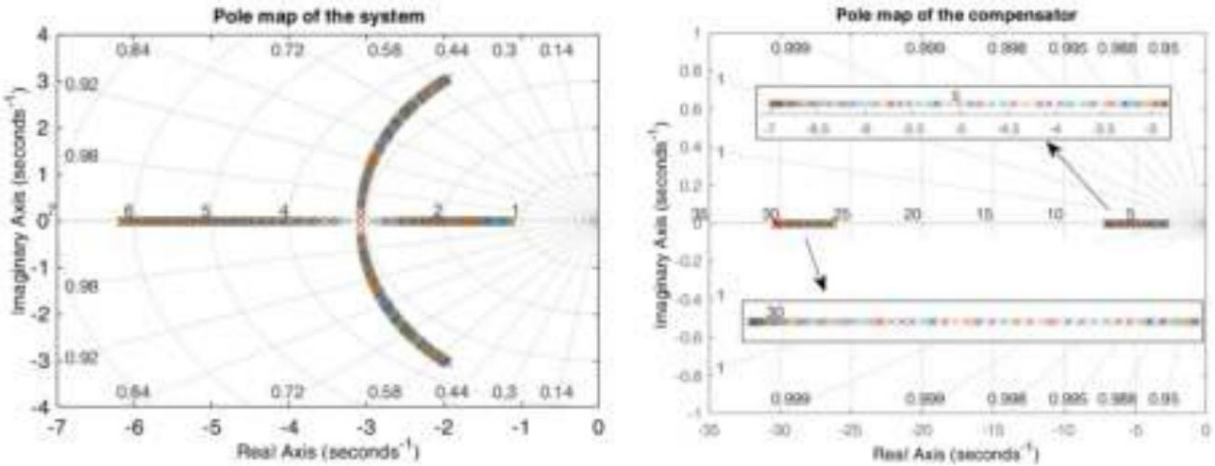


Abbildung 5. Polkarten des Systems und des Kompensators.

Abbildung 6 zeigt ein schematisches Blockdiagramm für den LPV-Kompensatorentwurf, wobei Wu und Wz jeweils die Eingangs- und Ausgangsgewichtungs Für dieses Beispiel wird allerdings ein entsprechender Kompensator ohne diese Filter ausgelegt.
Die Filter können jedoch aus Gründen wie beispielsweise strukturellen Einschränkungen verwendet werden, wie in Anmerkung 6 erwähnt.

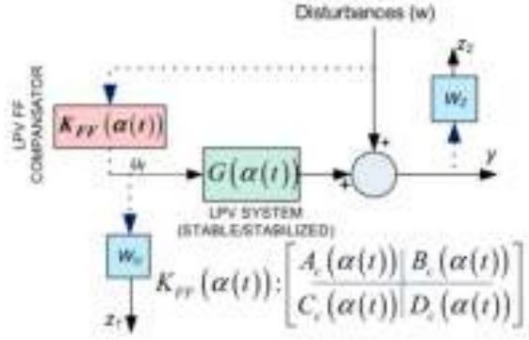


Abbildung 6. Blockdiagramm des LPV-Kompensatordesigns.

Machine Translated by Google

Anmerkung 6. Die Ein- und Ausgänge sind aufgrund der mechanisch/elektrischen Struktur begrenzt. Daher sollten die Beschränkungen je nach Problemstellung des jeweiligen Bereichs berücksichtigt werden. Beispielsweise sollte die Federungsauslenkung die Obergrenze nicht überschreiten, während die Reifenauslenkung bei aktiven Fahrzeugfederungssystemen so weit wie möglich minimiert werden sollte. Es gibt einige LMI-basierte Studien zu den Eingangs- und Ausgangsbeschränkungen, wie beispielsweise [51]. Diese werden mit der NCF geregelt , wie in Abbildung 2 gezeigt . Die NCF kann ein Koeffizient oder ein Filter, wie beispielsweise ein Tiefpassfilter, sein, der je nach Systemverhalten unter Berücksichtigung der Frequenz- und Verstärkungsgrenzen der Ein- und Ausgänge bestimmt wird. Die andere Möglichkeit sind die Eingangs- und Ausgangsgewichtungsfunktionen für den **Hÿ-** Reglerentwurf, wie in Abbildung 6 und dem Blockdiagramm in [45] gezeigt . Daher kann unter Berücksichtigung der Systemstruktur für die NCF entschieden werden. Insbesondere kann bei einer hohen Frequenz ein Filter für die NCF anstelle eines festen NC erforderlich sein. Dies liegt daran, dass der Kompensator ihn benötigt, um die Aktuatorgrenze bei einer h

Die Matrizen werden in (33) erhalten, wenn die LMIs in Theorem 2 gelöst werden, um den ÿ- **Wert gemäß den parameterabhängigen Variablen** Yi , BFi, DFi, Mi , Ni **für** i **= 0, 1, 2 und den parameterunabhängigen Variablen R1 und R2 zu minimieren. Daher werden die Zustandsraummatrizen des LPV-Kompensators** Aci, Bci, Cci **und** Dci **in (21) für** i = 0, 1 und 2 affin aus **den Matrizen in (33) gemäß den zeitabhängigen Parametern** ÿ1(t) **und** ÿ2(t) **in (32) für alle** t während des Online-Betriebs **erhalten . Daher zeigt (34) die Zustandsraummatr**

$$
\begin{aligned}
\mathbf{M0} &= \begin{matrix} \mathbf{ÿ163{,}81\ 12{,}724} \\ \mathbf{ÿ1{,}7999\ ÿ202{,}82} \end{matrix}, \mathbf{M1} = \begin{matrix} \mathbf{5{,}1074\ ÿ22{,}914\ ÿ11{,}194} \\ \mathbf{36{,}474} \end{matrix}, \\
\mathbf{M2} &= \begin{matrix} \mathbf{ÿ10{,}704\ 16{,}627} \\ \mathbf{15{,}009\ ÿ84{,}506} \end{matrix}, \mathbf{R1} = \begin{matrix} \mathbf{7{,}1371\ ÿ3{,}0487\ ÿ20{,}361} \\ \mathbf{92{,}818} \end{matrix}, \\
\mathbf{R2} &= \begin{matrix} \mathbf{7{,}139\ ÿ3{,}0494\ ÿ20{,}361} \\ \mathbf{92{,}826} \end{matrix} \mathbf{N1} = \mathbf{ÿ111{,}95}, \mathbf{N0} = \mathbf{ÿ1175{,}1\ 1477{,}5}, \\
&\mathbf{364{,}76\ ÿ0{,}07727\ 0{,}51139}, \mathbf{N2} = \mathbf{ÿ53{,}522\ 83{,}135\ 0{,}070224}, \\
\mathrm{BC0} &= \begin{matrix} \mathbf{0\ ÿ0{,}46491\ 0} \end{matrix}, \mathrm{BC1} = \begin{matrix} \mathbf{0\ ÿ0{,}46491\ 0} \end{matrix}, \mathrm{BC2} = \begin{matrix} \mathbf{0\ ÿ0{,}46491\ 0} \end{matrix}, \\
&\mathrm{DC0} = \mathbf{ÿ3{,}3213}, \mathrm{DC1} = \mathbf{ÿ4{,}6489}, \mathrm{DC2} = \mathbf{ÿ0{,}0003355}
\end{aligned}
\tag{33}
$$

$$
\begin{aligned}
\mathrm{Ac}(ÿ1, \mathbf{ÿ2}) &= \begin{matrix} \mathbf{ÿ24{,}886} + \mathbf{0{,}012541}ÿ1\ ÿ\ \mathbf{1{,}0907}ÿ2 \\ \mathbf{ÿ7{,}1544}\ ÿ\ \mathbf{0{,}49361}ÿ1\ ÿ\ \mathbf{0{,}54525}ÿ2 \end{matrix} \begin{matrix} ÿ\mathbf{2{,}42} + \mathbf{0{,}14329}ÿ2 \\ ÿ\mathbf{2{,}42} + \mathbf{0{,}37671}ÿ1\ ÿ\ \mathbf{0{,}92828}ÿ2 \end{matrix} \\
\mathit{Bc}(ÿ1) &= \begin{matrix} \mathbf{ÿ0{,}07727} + \mathbf{0{,}070224}ÿ1 \\ \mathbf{0{,}51139}\ ÿ\ \mathbf{0{,}46491}ÿ1 \end{matrix} \\
\mathrm{Cc}(ÿ1, \mathbf{ÿ2}) &= \mathbf{ÿ131{,}53}\ ÿ\ \mathbf{4{,}9366}ÿ1\ ÿ\ \mathbf{5{,}4537}ÿ2\ \mathbf{11{,}596} + \mathbf{3{,}7674}ÿ1 + \mathbf{0{,}71644}ÿ2 \\
\mathrm{Dc}(ÿ1, \mathbf{ÿ2}) &= \mathbf{ÿ2{,}8863}\ ÿ\ \mathbf{4{,}6489}ÿ1\ ÿ\ \mathbf{0{,}0003355}ÿ2
\end{aligned}
\tag{34}
$$

Dementsprechend zeigt Abbildung 7a das Bode-Diagramm des Kompensators für die zeitabhängigen Parameter in Abbildung 3a, während Abbildung 7b es für die zeitabhängigen Parameter in Abbildung 3b bei einer Abtastzeit von 50 ms zeigt. Den Poldiagrammen und Bode- Diagrammen zufolge ändert sich der LPV-Kompensator mit den zeitabhängigen Parametern erh

Der Wert von μ **wird durch Liniensuche nach einer guten Antwort auf 0,42 festgelegt. Dementsprechend** beträgt **der Leistungsindex** $ÿ$ **0,037268. Dieser Wert ändert sich eindeutig je nach System und Eingangs-/Ausgangsfiltern. Er bietet Machbarkeit für sehr große zeitabhängige Parameter, wie in Beispielen für die Gestaltung von Zustandsrückkopplungsreglern bewiesen wurde. Abbildung 8a zeigt auch das Bode-Diagramm des Systems kombiniert mit dem Kompensator vom Störungseingang zu den** z- **und** y **-Ausgängen unter Annahme der zeitabhängigen Parameter in Abbildung 3a, während Abbildung 8b das Bode-Diagramm unter Annahm Abbildung 9 zeigt die für die Simulationen verwendeten Störtestsignale. Dabei handelt es sich um ein Stoßsignal (siehe Abbildung 9a), ein 5-Rad/s-Sinussignal (siehe Abbildung 9b) und ein weißes Rauschsignal (siehe Abbildung 9c).**

Machine Translated by Google

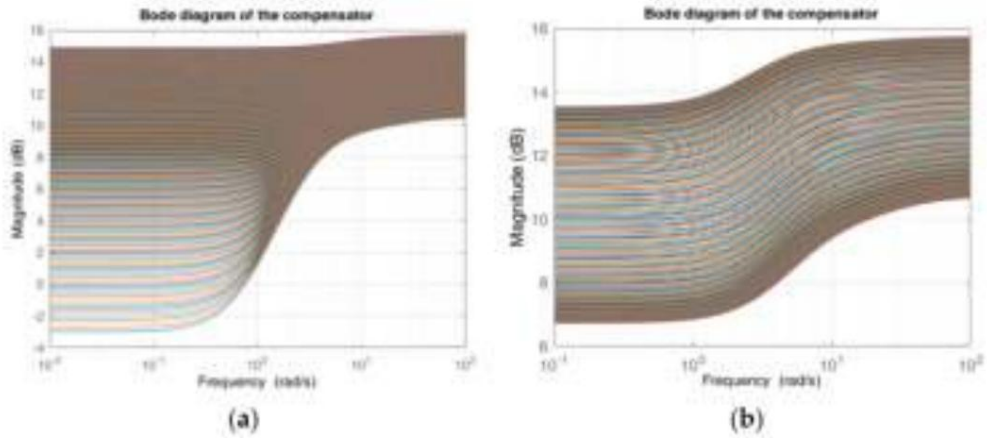


Abbildung 7. Bode-Diagramme des Kompensators für zeitveränderliche Parameter in: (a) Abbildung 3a; (b) Abbildung 3b.

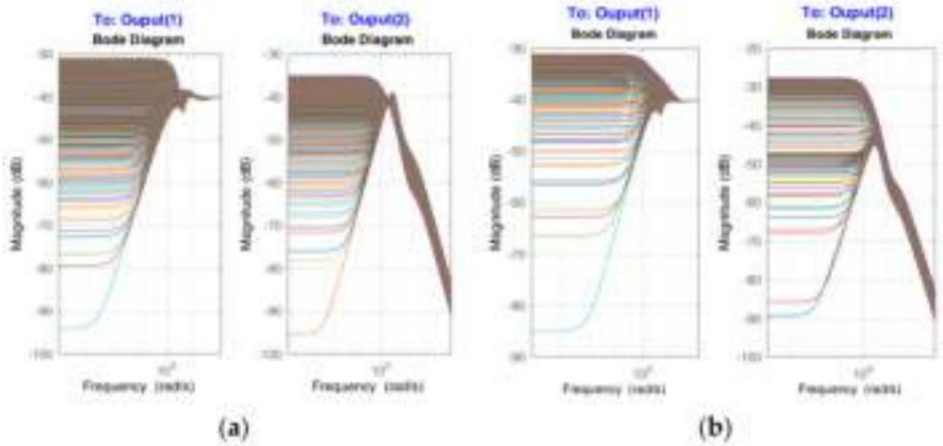


Abbildung 8. Bode-Diagramme des Systems mit dem Kompensator für zeitveränderliche Parameter in: (a) Abbildung 3a; (b) Abbildung 3b.

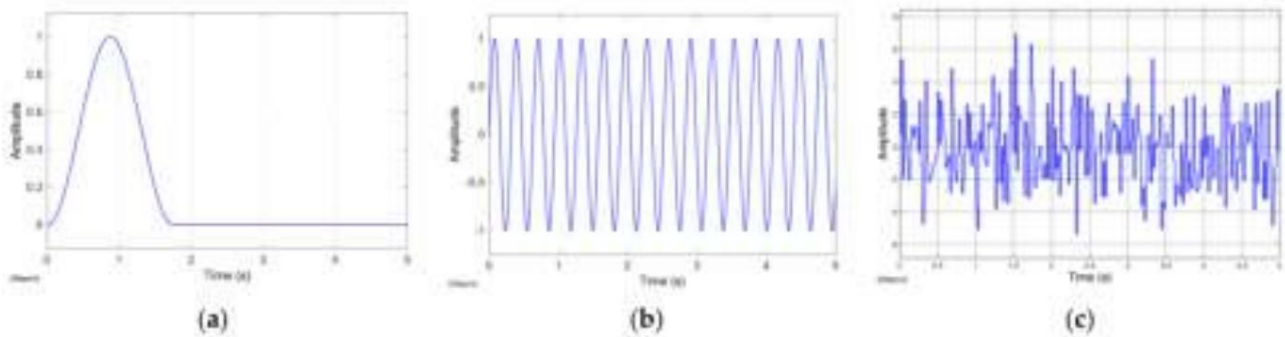


Abbildung 9. Angewandte Teststörungen: (a) eine Beule; (b) ein Sinussignal mit 5 rad/s; (c) ein weißes Rauschsignal.

Was die Ergebnisse betrifft, zeigt Abbildung 10 die Ausgänge z und y für das externe Störungsstoßsignal in Abbildung 9a. Wie aus den Ausgängen ersichtlich, dämpft der entworfene LPV-Kompensator die Störung effektiv mit einer Rate von ungefähr 100 % im Vergleich zum Fall ohne den Kompensator.

Machine Translated by Google

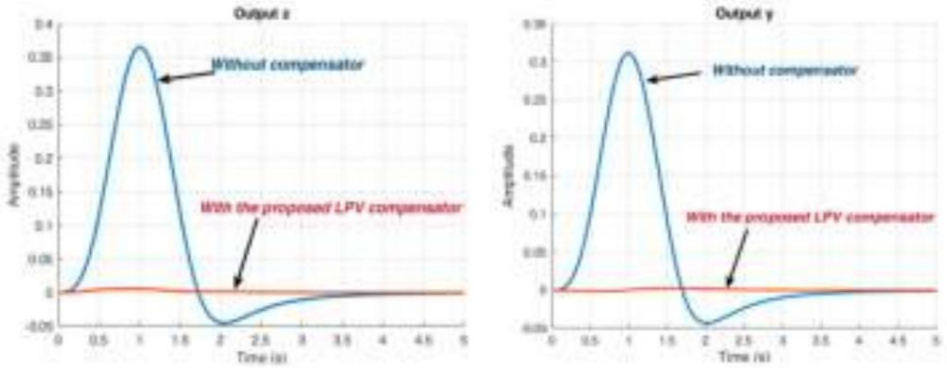


Abbildung 10. Die Ausgänge z und y für ein Stoßstörungssignal.

In ähnlicher Weise zeigen die Abbildungen 11 und 12 die Ausgänge z **und** y **für die Störungstestsignale in den Abbildungen 9b bzw. 9c. Die vom parameterabhängigen Kompensator gegen drei Störungen erzeugten Signale sind in Abbildung 13 dargestellt.**

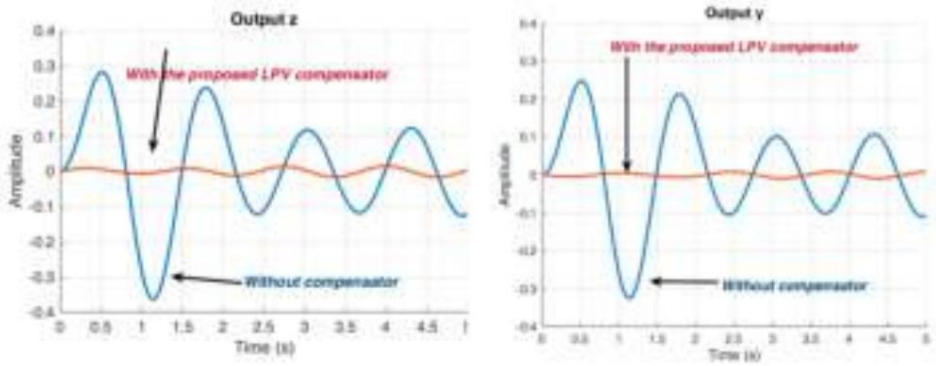


Abbildung 11. Die Ausgaben z und y für eine Sinusstörung von 5 rad/s.

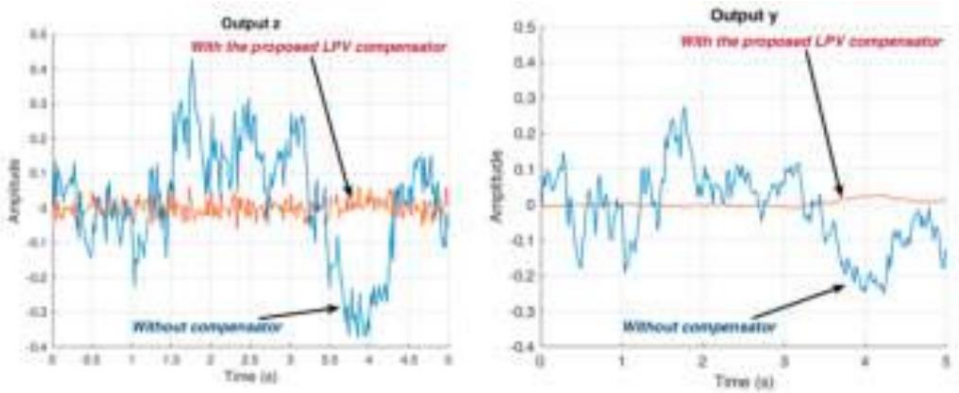


Abbildung 12. Die Ausgaben z und y bei weißer Rauschstörung.

Das System mit einem LPV-Feedforward-Kompensator wird mit dem System ohne Kompensator verglichen. Es ist zu beachten, dass das System stabil sein oder durch einen Rückkopplungsregler stabilisiert werden muss. Wie aus den Ausgaben in den Abbildungen 11 und 12 ersichtlich, ist der parameterabhängige Kompensator recht erfolgreich, ähnlich den Ergebnissen in Abbildung 10. Somit werden die Auswirkungen der externen Störungen um fast 100 % reduziert. Daher weist der entworfene LPV-Feedforward-Kompensator eine signifikante Leistung gegenüber den Störungen auf. Schließlich ist der Kompensatorentwurf für das LPV-System recht erfolgreich. Infolgedessen ist der vorgeschlagene Kompensator recht erfolgr

Machine Translated by Google

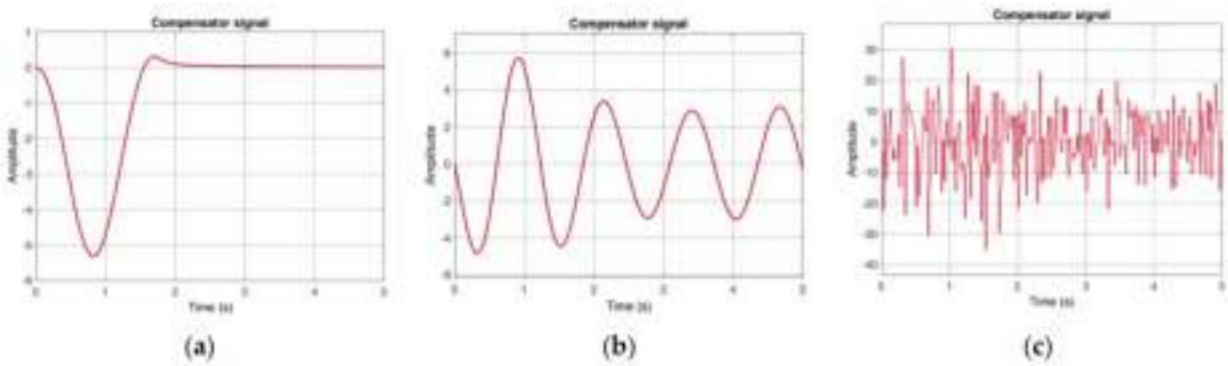


Abbildung 13. Die generierten Kraftsignale für: (a) ein Stoßsignal; (b) ein 5 rad/s Sinussignal; (c) ein weißes Rauschsignal.

Beim Entwurf des vorgeschlagenen LPV-Kompensators auf Basis des erweiterten LMI kann die Lya-Punov-Matrix in der LPV-Form bezeichnet werden, was den Konservativismus reduziert, ohne zusätzliche Ansätze wie Gitter, S-Prozedur, Multikonvexität in [11] und die Bedingungen in [18,19]. Beispielsweise existieren viele LMI-Bedingungen für den LPV-Reglerentwurf in [45], da Multikonvexitätsbedingungen aufgrund der Produkte $A(\dot{y}) \times P(\dot{y})$ in (6) verwendet werden. Daher ist die vorgeschlagene Lösung einfacher und weniger konservativ. Andernfalls werden entweder ineffiziente bilineare Matrixungleichung (BMI)-Löser oder iterative LMI-Algorithmen benötigt.
Die Lösungen erfordern keine iterativen LMI- oder BMI-Algorithmen.

6. Schlussfolgerungen und zukünftige Arbeiten

In diesem Artikel wird neben einem Zustandsrückkopplungsregler für das LPV-System auch ein Feedforward-Kompensator gegen Störungen behandelt. Die Lyapunov-Matrix wird von den Systemmatrizen durch Einfügen einer lockeren reellen Variable in die LMI-basierten Designs getrennt; daher wird die PDLM bezeichnet. Daher sind die vorgeschlagenen LMI-Formulierungen sowohl für den Zustandsrückkopplungsregler als auch für den Feedforward-Kompensator endlichdimensional und benötigen keine zusätzlichen Bedingungen wie das S-Verfahren, Gitter und Mehrfachkonvexität, selbst wenn die PDLM verwendet wird. Daher haben sie eine einfachere Struktur und sind weniger konservativ als ähnliche Zustandsrückkopplungs- und Feedforward-Designs. Die numerischen Ergebnisse des Zustandsrückkopplungsreglers zeigen, dass die Lösung selbst bei sehr großen Parameterbereichen weniger konservativ und einfacher ist. Insbesondere für den Kompensator zeigen die Simulationsergebnisse, dass sowohl die parameterabhängigen als auch die parameterunabhängigen Feedforward-Kompensatoren eine signifikante Leistung gegen Störungen aufweisen, da die Auswirkungen der externen Störungen fast vollständig eliminiert werden . Daher werden die Gültigkeit und Vorteile der vorgeschlagenen LMIs effektiv mit numerischen und Simulationsbeispielen ermittelt. Infolgedessen sind die vorgeschlagen

Für weitere Studien kann die erste davon als das Erhalten des optimalen realen Skalars μ über metaheuristische Optimierungsmethoden betrachtet werden. Die zweite kann sowohl eine Rückkopplungssteuerung als auch einen Feedforward-Kompensator gleichzeitig für instabile Systeme entwerfen. Die dritte ist das Design des Umschaltens des LPV-Feedforward-Kompensators über den erweiterten LMI. Darüber hinaus kann **dieser vorgeschlagene Feedforward-Kompensator mit einem geschalteten LPV-Feedforward-Kompensator hinsichtlich Kriterien wie Leistung u**

Beiträge der Autoren: Konzeptualisierung, M.ÿI. und YA; Methodik, M.ÿI. und YA; Software, M.ÿI. und YA; Validierung, YA; formale Analyse, YA; Untersuchung, M.ÿI. und YA; Ressourcen, M.ÿI.; Datenkuratierung, M.ÿI.; Schreiben – Vorbereitung des Originalentwurfs, M.ÿI.; Schreiben – Überprüfung und Bearbeitung, YA; Visualisierung, M.ÿI. und YA; Aufsicht, YA. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und ihr zugestimmt.

Finanzierung: Diese Forschung erhielt keine externe Finanzierung.

Datenverfügbarkeitserklärung: Die in der Studie vorgestellten Originalbeiträge sind im Artikel enthalten, weitere Anfragen können an den entsprechenden Autor gerichtet werden.

Machine Translated by Google

Interessenkonflikte: Der Autor Mustafa ÿInci war bei der Firma Volta Motor Industry and Trade Inc. angestellt. Die übrigen Autoren erklären, dass die Forschung ohne kommerzielle oder finanzielle Beziehungen durchgeführt wurde, die als potenzieller Interessenkonflikt ausgelegt we

Verweise

1. **Kim, C. Robuste Hÿ- statische Ausgangsrückkopplungssteuerung für TCP/AQM-Router basierend auf LMI-Optimierung.** Electronics **2024,** 13, **2165.** [Querverweis]
2. **Kim, Y.; Tran, TV; Kim, K.-H. Robuste EMPC-basierte frequenzadaptive sensorlose Netzspannungsregelung für einen LCL-gefilterten netzgekoppelten Wechselrichter.** Electronics **2024,** 13, **998.** [CrossRef]
3. **Németh, Z.; Kuczmann, M. Linear-Matrix-Inequality-Based Controller and Observer Design for Induction Machine.** Elektronik **2022,** 11, **3894.** [CrossRef]
4. **Ndabarushimana, E.; Qin, N.; Ma, L. Störungsentkopplung für einen einphasigen Pulsweitenmodulationsgleichrichter basierend auf einem Erweiterter H-Infinity-Filter.** Electronics **2023,** 12, **2765.** [CrossRef]
5. **Boyd, S.; El Ghaoui, L.; Feron, E.; Balakrishnan, V.** Lineare Matrixungleichungen in der System- und Kontrolltheorie; **Gesellschaft für industrielle und Applied Mathematics (SIAM) Publishing: Philadelphia, PA, USA, 1994; Band 15, ISBN 0898714850.**
6. **De Oliveira, MC; Geromel, JC; Bernussou, J. Erweiterte H2- und H-Norm-Charakterisierungen und Controller-Parametrisierungen für Diskrete Zeitsysteme.** Int. J. Control **2002,** 75, **666–679.** [CrossRef]
7. **Bara, GI Robuste Analyse und Kontrolle von Parameter-abhängigen unsicheren Deskriptorsystemen.** Syst. Control Lett. **2011,** 60, **356–364.** [CrossRef]
8. **Cao, Y.-Y.; Lin, Z. Ein Deskriptorsystemansatz zur robusten Stabilitätsanalyse und Controllersynthese.** IEEE Trans. Autom. Control **2004,** 49, **2081–2084.** [CrossRef]
9. **De Caigny, J.; Camino, JF; Oliveira, RCLF; Peres, PLD; Swevers, J. Gain-Scheduled H2 und Hÿ Steuerung der diskreten Zeit Polytopische zeitvariable Systeme.** IET Control Theory Appl. **2010, 4, 362–380.** [CrossRef]
10. **Oliveira, MCD; Geromel, JC; Bernussou, J. Ein LMI-Optimierungsansatz für den Entwurf multiobjektiver Regler für zeitdiskrete Systeme. In Proceedings der 38. IEEE-Konferenz über Entscheidung und Steuerung (Kat.-Nr. 99CH36304), Phoenix, AZ, USA, 7.–10. Dezember 1999; IEEE: Piscataway, NJ, USA, 1999; Band 4, S. 3611–3616.**
11. **Gahinet, P.; Apkarian, P.; Chilali, M. Affine Parameter-abhängige Lyapunov-Funktionen und reale parametrische Unsicherheit.** IEEE Trans. Automat. Control **1996,** 41, **436–442.** [CrossRef]
12. **Pipeleers, G.; Demeulenaere, B.; Swevers, J.; Vandenberghe, L. Systems & Control Letters Erweiterte LMI-Charakterisierungen für Stabilität und Leistung linearer Systeme.** Syst. Control Lett. **2009,** 58, **510–518.** [CrossRef]
13. **Ku, C.-C.; Chen, G.-W. Gain-Scheduled Controller Design für zeitdiskrete lineare Parametervariationssysteme mit multiplikativer Geräusche.** Int. J. Control Autom. Syst. **2015,** 13, **1382–1390.** [CrossRef]
14. **Becker, G.; Packard, A. Robuste Leistung linearer parametrisch variierender Systeme unter Verwendung parametrisch abhängiger linearer Rückkopplung.** Syst. Control Lett. **1994,** 23, **205–215.** [CrossRef]
15. **Feron, E.; Apkarian, P.; Gahinet, P. S-Verfahren zur Analyse von Kontrollsystemen mit parametrischen Unsicherheiten über parameterabhängige Lyapunov-Funktionen. In Proceedings der 1995 American Control Conference, Seattle, WA, USA, 21.–23. Juni 1995; IEEE: Piscataway, NJ, USA, 1995; Band 1, S. 968–972.**
16. **Wei, X.; Verhaegen, M. LMI-Lösungen für das gemischte H-/Hÿ- Fehlererkennungs-Beobachterdesign für lineare Parametervariation Systeme.** Int. J. Adapt. Kontrollsignalprozess. **2011,** 25, **114–136.** [CrossRef]
17. **Lim, S.** Analyse und Steuerung linearer Parametervariationssysteme; **Stanford University: Stanford, CA, USA, 1999; ISBN 0599241535.**
18. **Oliveira, RCLF; Peres, PLD Parameterabhängige LMIs in der robusten Analyse: Charakterisierung homogener polynomialer Parameterabhängiger Lösungen mittels LMI-Relaxationen.** IEEE Trans. Autom. Control **2007,** 52, **1334–1340.** [CrossRef]
19. **Ebihara, Y.; Peaucelle, D.; Arzelier, D.** S-Variable Approach to LMI-Based Robust Control; **Springer: Berlin/Heidelberg, Deutschland, 2015; ISBN 978-1-4471-6605-4.**
20. **Ebihara, Y.; Hagiwara, T. Neue erweiterte LMI-Charakterisierungen für die kontinuierliche multiobjektive Reglersynthese.** Automatica **2004,** 40, **2003–2009.** [CrossRef]
21. **Song, L.; Yang, J. Ein verbesserter Ansatz zur robusten Stabilitätsanalyse und Reglersynthese für LPV-Systeme.** Int. J. Robust Nonlinear Control **2010,** 21, **1574–1586.** [CrossRef]
22. **Apkarian, P.; Tuan, HD; Bernussou, J. Kontinuierliche Analyse, Eigenstrukturzuordnung und H2-Synthese mit verbesserten linearen Matrixungleichungen (LMI)-Charakterisierungen.** IEEE Trans. Autom. Control **2001,** 46, **1941–1946.** [CrossRef]
23. **Ebihara, Y.; Hagiwara, T. Ein erweiterter LMI-Ansatz zur robusten Leistungsanalyse linearer zeitinvarianter unsicherer Systeme.** Automatica **2005,** 41, **1933–1941.** [CrossRef]
24. **Trofino, A.; Coutinho, DF; Barbosa, KA Verbesserte H2- und Hÿ -Bedingungen für robuste Analyse und Kontrollsynthese linearer Systeme.** Controle Autom. **2005,** 16, **427–434.** [CrossRef]
25. **Shaked, U. Verbesserte LMI-Darstellungen für die Analyse und den Entwurf kontinuierlicher Systeme mit polytopischem Typ Unsicherheit.** IEEE Trans. Autom. Control **2001,** 46, **652–656.** [CrossRef]
26. **Wei, J.; Lee, L. Weitere Verbesserung von LMI-Darstellungen für die Analyse und das Design kontinuierlicher Systeme mit polytopischer Typunsicherheit 1. In Proceedings der 5. Asian Control Conference 2004 (IEEE Cat. No. 04EX904), Melbourne, VIC, Australien, 20.–23. Juli 2004; Band 2, S. 1–7.**

Machine Translated by Google

27. Sajjadi-Kia, S.; Jabbari, F. Über beschränkte reelle Matrixungleichungsdilatation. Int. J. Control **2012,** 85, **1593–1601.** [CrossRef]

28. Xie, W. Eine äquivalente LMI-Darstellung des beschränkten reellen Lemmas für zeitkontinuierliche Systeme. J. Inequalities Appl. **2008,** 2008, **672905.** [CrossRef]

29. Dabboussi, K.; Zrida, J. Ein erweiterter LMI-Ansatz zur robusten Synthese von **Hÿ- Statik-Feedback-Controllern. In Proceedings der 16. IEEE International Conference on Electronics, Circuits and Systems (ICECS 2009), Yasmine Hammamet, Tunesien, 13.–16. Dezember 2009; S. 491–494.**

30. Rodrigues, LA; Oliveira, RCLF; Camino, JF Parametrisierte LMIs für robuste und zustandsgekoppelte Steuerung von kontinuierlichen polytopischen Systemen. Int. J. Robust Nonlinear Control **2018,** 28, **940–952.** [CrossRef]

31. Hu, C.; Jaimoukha, IM Neues iteratives Verfahren auf Basis linearer Matrixungleichungen für die H2- und Hÿ- Zustandsrückkopplungssteuerung von kontinuierlichen polytopischen Systemen. Int. J. Robust Nonlinear Control **2021,** 31, **51–68.** [CrossRef]

32. Guzmán, JL; Hägglund, T. Abstimmungsregeln für die Vorwärtsregelung aus messbaren Störungen in Kombination mit PID-Regelung Eine Rezension. Int. J. Control **2021,** 97, **2–15.** [CrossRef]

33. Zhu, W.; Zhang, Z.; Chen, J.; Zhao, S.; Huang, S. Dynamischer Datenabgleich zur Verbesserung der Leistung von Feedforward- /Feedback-Regelsystemen mit Messrauschen. J. Process Control **2021,** 108, **12–24.** [CrossRef]

34. García-Mañas, F.; Guzmán, JL; Rodríguez, F.; Berenguel, M.; Hägglund, T. Experimentelle Bewertung von Feedforward-Tuning -Regeln. Kontroll-Ing. Üben. **2021,** 114, **104877.** [CrossRef]

35. Du, Y.; Cao, W.; She, J.; Wu, M.; Fang, M. Störungsunterdrückung durch Feedforward-Kompensation unter Verwendung eines verbesserten äquivalenten Eingangsstörungsansatzes. J. Frankl. Inst. **2020,** 357, **10977–10996.** [CrossRef]

36. Huang, M.; Deng, Y.; Li, H.; Wang, J. Dämpfung der Drehmomentwelligkeit von PMSM durch verbesserte robuste Zwei-Freiheitsgrade Regler über erweiterten Sliding Mode Parameter Observer. ISA Trans. **2022,** 129, **558–571.** [CrossRef]

37. Adam, EJ; Marchetti, JL Entwurf und Feinabstimmung robuster Feedforward-Regler. Comput. Chem. Eng. **2004,** 28, **1899–1911.** [Querverweis]

38. Peng, C.; Xu, H.; Zhang, J. Zwei-Freiheitsgrad-Steuerungsdesign mit verbesserter Hÿ LMI-Darstellung. J. Inequalities Appl. **2012,** 2012, **147.** [CrossRef]

39. Köse, IE; Scherer, CW Robuste L2-Gain-Feedforward-Regelung unsicherer Systeme mit dynamischen IQCs. Int. J. Robust Nonlinear Control **2009,** 19, **1224–1247.** [CrossRef]

40. Alma, M.; Martinez, JJ; Landau, ID; Buche, G. Design und Abstimmung von H-Infinity Feedforward-Kompensatoren reduzierter Ordnung für Aktive Vibrationskontrolle. IEEE Trans. Control Syst. Technol. **2012,** 20, **554–561.** [CrossRef]

41. Peng, C.; Zhang, Z.; Zou, J.; Li, K.; Zhang, J. Internes Modell basierend auf robuster Inversion Feedforward und Feedback 2DOF-Steuerung für LPV-System mit Störung. J. Process Control **2013,** 23, **1415–1425.** [CrossRef]

42. Yuan, C.; Wu, F. Robuste H2- und Hÿ -geschaltete Feedforward-Steuerung unsicherer LFT-Systeme. Int. J. Robuste nichtlineare Steuerung **2016,** 26, **1841–1856.** [CrossRef]

43. Fleps-Dezasse, M.; Bünte, T.; Svaricek, F.; Brembeck, J. LPV Feedforward Control of Semi-Active Suspensions for Improved Roll Stability. Control Eng. Pract. **2018,** 78, **1–11.** [CrossRef]

44. Altun, Y.; Gulez, K.; Mumcu, TV Statische LPV-Feedforward-Controller-Synthese für lineare Parametervariationssysteme. In Proceedings der 9. Asian Control Conference (ASCC) 2013, Istanbul, Türkei, 23.–26. Juni 2013; IEEE: Piscataway, NJ, USA, 2013; S. 1–4.

45. Altun, Y.; Gulez, K. Lineare Parameter variierende Feedforward-Steuerungssynthese unter Verwendung der parameterabhängigen Lyapunov-Funktion. Nichtlineare Dyn. **2014,** 78, **2293–2307.** [CrossRef]

46. Venkataraman, R.; Seiler, P. Konvexe LPV-Synthese von Schätzern und Feedforwards unter Verwendung von Dualität und integraler quadratischer Einschränkungen. Int. J. Robust Nonlinear Control **2018,** 28, **953–975.** [CrossRef]

47. Löfberg, J. Automatische robuste konvexe Programmierung. Optim. Methods Softw. **2012,** 27, **115–129.** [CrossRef]

48. Labit, Y.; Peaucelle, D.; Henrion, D. SEDUMI INTERFACE 1.02: Ein Tool zum Lösen von LMI-Problemen mit SEDUMI. In Proceedings des 2002 IEEE International Symposium on Computer Aided Control System Design, CACSD 2002—Proceedings, Glasgow, Großbritannien, 20. September 2002; Seiten 272–277.

49. Gahinet, P.; Nemirovskii, A.; Laub, AJ; Chilali, M. LMI Control Toolbox. In Proceedings der 33. IEEE-Konferenz über Entscheidung und Kontrolle 1994, Lake Buena Vista, FL, USA, 14.–16. Dezember 1994; The Mathworks, Inc.: Natick, MA, USA, 1995; Seiten 4–13.

50. Apkarian, P.; Adams, RJ Fortgeschrittene Gain-Scheduling-Techniken für unsichere Systeme. IEEE Trans. Control Syst. Technol. **1998, 6, 21–32.** [CrossRef]

51. Chen, H.; Guo, KH Eingeschränkte Hÿ- Steuerung aktiver Aufhängungen: Ein LMI-Ansatz. IEEE Trans. Control. Syst. Technol. **2005,** 13, **412–421.** [CrossRef]

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Ein adaptiver, hocheffizienter LED-Hintergrundbeleuchtungstreiber

Xinyu Xu, Qiyue Zhuo, Chunhui Jiang, Zichao Zhou, Junlong Tang und Wanghui Zou *

Fakultät für Physikalische und Elektronische Wissenschaften, Changsha University of Science and Technology, Changsha
410114, China * Korrespondenz: zouwh@csust.edu.cn

Zusammenfassung: **Um das Problem des zusätzlichen Leistungsverlusts durch Schwankungen der LED-Durchlassspannung zu lösen, wurde ein adaptives Treiberschema für die Hintergrundbeleuchtung von Leuchtdioden (LED) mit hohem Wirkungsgrad** vorgeschlagen. Bei **diesem Schema werden der Spitzenstrom und der Arbeitszyklus jedes LED-Kanals separat über einen adaptiven Steueralgorithmus eingestellt, um den Spannungsabfall am linearen Stromregler (LCR) jedes LED-Kanals zu minimieren und den übermäßigen Leistungsverlust in jedem LED-Kanal zu verringern und die Gesamtleistungseffizienz zu verbessern . Ein für die adaptive Steuerung geeigneter linearer Stromregler wurde auf Grundlage eines 0,18 µm 5 V komplementären Metalloxid-Halbleiter-Prozesses (CMOS) entwickelt. Simulationsergebnisse zeigen, dass der lineare Stromregler einen linear einstellbaren Kanalstrom im Bereich von 0 bis 48 mA mit einer Stromauflösung von 0,2 mA erreichen kann. Über verschiedene Prozessecken und Temperaturen hinweg beträgt der maximale Fehler für den gesamten Strombereich weniger als 0,1 %. Die Kernfläche des Chip-Layouts beträgt etwa 0,1 mm2. Der vollständige Treiberprototyp besteht aus den LCR-Chips, externen Power-MOS-Transistoren, dem Digitalmodul und den LED-Ketten. Die Testergebnisse zeigen, dass der Leistungsverlust des linearen Stromreglers** erheblich reduziert wurde und die Energieeffizienz jedes LED-Kanals bei **etwa 98,1 % liegt .**

Schlüsselwörter: LED-Hintergrundbeleuchtungstreiber; adaptive Steuerung; lineare Stromregelung; hohe Effizienz

check for updates

Zitat: Xu, X.; Zhuo, Q.; Jiang, C.; Zhou, Z.; Tang, J.; Zou, W. Ein adaptiver hocheffizienter LED-Hintergrundbeleuchtungstreiber. Electronics **2024,** 13, **3057. https://doi.org/10.3390/Elektronik13153057**

Wissenschaftlicher Redakteur: Francesco Driussi

Empfangen: 22. Juni 2024
Überarbeitet: 25. Juli 2024
Akzeptiert: 30. Juli 2024
Veröffentlicht: 2. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.
Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Einleitung

Leuchtdioden (LEDs) werden häufig als Hintergrundbeleuchtungsquellen für Flüssigkristallanzeigen (LCD) und andere Anwendungen eingesetzt. Im Vergleich zu herkömmlichen Lichtquellen bieten LEDs erhebliche Vorteile wie hohe Effizienz, ausgezeichnete Farbstabilität, hohe Zuverlässigkeit, lange Lebensdauer, kompakte Größe und Umweltverträglichkeit [1–4]. Da die Helligkeit einer LED proportional zu ihrem Strom ist, ist bei LED-Hintergrundbeleuchtungsanwendungen eine effiziente und genaue Treiberschaltung unerlässlich, um eine präzise Helligkeitsregelung zu erreichen. Darüber hinaus werden immer parallele LED-Kanäle (LED-Ketten) eingesetzt, um eine ausreichende Abdeckung zu gewährleisten [5,6]. Ein herkömmlicher LED-Hintergrundbeleuchtungstreiber besteht hauptsächlich aus einem Aufwärtswandler, parallelen LED- Kanälen und entsprechenden linearen Stromreglern (LCRs). Der Aufwärtswandler liefert ausreichend Spannung, um alle LEDs in jedem Kanal einzuschalten, und der LCR regelt den Strom für jeden Kanal, um die Helligkeit zu regeln [7–9]. Als wichtiger Bereich der Energieverwaltungsanwendung war das Design von LED-Hintergrundbeleuchtungstreibern schon immer ein beliebter Forschungsschwerpunkt zahlreicher Universitäten und Forschungseinrichtungen weltweit. Ziel ist es, die Systemleistung durch innovative Schaltungsdesigns und Systemarchitekturen zu optimieren, wobei der Schwerpunkt auf Kennzahlen wie System-Effizienz [10–15], Dimmverhältnis [16–20]

Aufgrund von Prozessschwankungen, Alterungseffekten und Temperaturschwankungen können LEDs desselben Typs unterschiedliche Durchlassspannungen aufweisen, was zu erheblichen Spannungsunterschieden zwischen verschiedenen LED-Kanälen führt. Daher ist es notwendig, einen gewissen **Spannungsspielraum einzubauen,** um die korrekte Funktionalität unter verschiedenen Arbeitsbedingunge **Dies führt jedoch auch zu erhöhten Leistungsverlusten und einer verschlechterten Energieeffizienz [26,27]. Um das Problem des erhöhten Leistungsverlustes durch zu hohe Spannung zu lösen,** wurde in **[28] ein digitaler Regelmechanismus auf Basis eines R-DAC vorgeschlagen.** Ein Vergleich zwischen der Span

Machine Translated by Google

am LCR wird eine voreingestellte Referenzspannung angelegt, die über einen Digital-Analog-Umsetzer (DAC) die Einstellung der Rückkopplungsspannung des DC–DC-Aufwärtswandlers steuert. Der Ausgang des Aufwärtswandlers wird dann angepasst, um die LCR-Spannung auf einem festen Niveau zu halten. In [29] wurde eine adaptive Referenzverfolgungstechnik vorgeschlagen, bei der eine „Minimalspannungs “-Schaltung und ein einstellbarer 3-Bit-Widerstand als Rückkopplungswiderstand verwendet werden, um den Ausgang des Aufwärtswandlers einzustellen. Die Genauigkeit der Ausgangsspannung wird von der Präzision und dem Einstellbereich des Widerstands beeinflusst. In [30–34] wird durch einen „Minimalspannungswähler“ oder „Niedrigster Kanalwähler“ die minimale LCR-Spannung zwischen den Kanälen erkannt und verwendet, um den Arbeitszyklus des Schaltsignals des Aufwärtswandlers dynamisch anzupassen und den Ausgang des Aufwärtswandlers zu ändern. In [35,36] werden Dioden verwendet, um die minimale LCR-Spannung auszuwählen, eine Methode, die in Standard-CMOS-Prozessen nicht einfach umzusetzen ist. In [2] wurden die Vor- und Nachteile verschiedener LED-Treibertopologien analysiert und zukünftige Forschungsrichtungen identifiziert.

Im Allgemeinen besteht die Hauptstrategie bei den meisten bestehenden Lösungen darin, die Ausgangsspannung des Aufwärtswandlers anzupassen, um die übermäßige Spannung am LCR zu reduzieren und dadurch Verluste zu verringern und die Effizienz zu verbessern. Bei einer Mehrkanalkonfiguration ist es jedoch aufgrund der Unterschiede zwischen den Kanälen effektiver, jeden Kanal separat zu optimieren, um Verluste weiter zu reduzieren und die Effizienz zu verbessern. In diesem Artikel schlagen wir ein adaptives, hocheffizientes LED-Hintergrundbeleuchtungs-Ansteuerungsschema vor, das das Problem der Unterschiede zwischen den Kanälen angeht. Für jeden Kanal werden sowohl eine lineare Durch lineare Stromregelung wird der Kanalstrom angepasst und die Spannung am LCR in jedem Kanal auf einen vordefinierten Festwert minimiert. Anschließend wird durch Pulsweitenmodulation der Arbeitszyklus des Kanalstroms angepasst, wodurch der Durchschnittsstrom jedes Kanals auf das gleiche Niveau kalibriert wird. Diese Strategie stellt sicher, dass jeder LED-Kanal mit minimalem Verlust arbeitet und so die Effizienz maximiert wird.

Die restlichen Teile dieses Artikels sind wie folgt strukturiert: Abschnitt 2 erläutert die in diesem Artikel entwickelte adaptive Steuerungstechnologie. Abschnitt 3 diskutiert und analysiert die Simulations- und Testergebnisse des vorgeschlagenen adaptiven hocheffizienten LED-Hintergrundbeleuchtungstreibers . Abschnitt 4 schließt den Artikel ab.

2. Adaptive Regelungstechnik

2.1. Prinzip der adaptiven Regelung

Abbildung 1 zeigt die Struktur des vorgeschlagenen adaptiven hocheffizienten LED-Hintergrundbeleuchtungstreibers , der aus einem Helligkeitsregler, einem Aufwärtswandler und mehreren LED- Kanälen besteht. Jeder LED-Kanal besteht aus einer LED-Kette, einem LCR und einem Hochspannungs-MOS- Transistor, der als Schalter zur Steuerung des Arbeitszyklus fungiert. Der Helligkeitsregler enthält einen 9-Bit-R-DAC zur Erzeugung von VC, der als Spannungsreferenz dient und den Ausgang des Aufwärtswandlers Vboost bestimmt. Das SENSE-Signal z Der Helligkeitsregler überwacht SENSE jedes Kanals und generiert VC-, PWM- und CM_SW-Signale nach einem bestimmten Algorithmus und in bestimmten Schritten. CM_SW und PWM sind digitale Mehrbitsignale, die jeweils den Spitzenstrom und den Arbeitszyklu Dieser Prozess stellt sicher, dass jeder LED-Kanal mit hoher Effizienz arbeitet.

Aus der Abbildung ergibt sich für Kanal n eine Spannungsbeziehung als

$$Vboost = VLED,n + VMP,n + VLCR,n \quad (1)$$

wobei VLED die Spannung über der LED-Kette oder die Summe der Durchlassspannungen jeder LED ist. VMP ist die Source-Drain-Spannung des Hochspannungs-MOS-Transistors MP und VLCR bezeichnet die LCR-Spannung. Das „n“ im Index steht für Kanal n, was für alle Symbole in diesem Dokument gilt. Wenn die adaptive Steuerung deaktiviert ist und der LED-Kanal nicht im optimalen Zustand ist, gibt es einen übermäßigen Anteil in der LCR-Spannung, der wie folgt ausgedrückt werden kann:

$$VLCR,n = VLCR_min + VLCR_exc,n \quad (2)$$

Machine Translated by Google

$$= + \quad (2)$$

wobei VLCR,min die minimal erforderliche LCR-Spannung ist, um einen ordnungsgemäßen Betrieb sicherzustellen, d. h. wobei VLCR,min die minimal erforderliche LCR-Spannung ist, um einen ordnungsgemäßen Betrieb si
erhalten ist zu sagen, der Ausgangstransistor im LCR funktioniert einwandfrei und hält die genaue Ausgabe aufrecht
zu sagen, der Ausgangstransistor im LCR funktioniert einwandfrei und hält die genaue Ausgabe aufrecht
Strom. VLCR_exc stellt den übermäßigen Teil der LCR-Spannung dar. Für einen Kanal ist der Strom. VLCR_exc
stellt den übermäßigen Teil der LCR-Spannung dar. Für einen Kanal ist der
Der Leistungsverlust kann berechnet werden durch

$$= \quad = \quad (3)$$

$$P_{loss,n} = I_{ave,n} \times V_{loss,n} = I_{ave,n} \times (V_{LCR,n} + V_{MP,n})$$

wobei Iave,n der durchschnittliche Strom ist, der durch den Spitzenstrom Ipeak,n und den bestimmt wird, wobei
Iave,n der durchschnittliche Strom ist, der durch den Spitzenstrom Ipeak,n und den bestimmt wird
aktuellen Tastverhältnis yn. Vloss,n ist die Summe aus VLCR,n und VMP,n. Wenn die Einschaltwiderstände des
aktuellen Tastverhältnisses yn. Vloss,n ist die Summe aus VLCR,n und VMP,n. Wenn die Einschaltwiderstände des
Die Hochspannungs-MOS-Transistoren sind niedrig genug, VMP hat nur begrenzte Auswirkungen auf Ploss. Weil
Die Hochspannungs-MOS-Transistoren sind niedrig genug, VMP hat nur begrenzte Auswirkungen auf Ploss. Weil
VMP und VLCR_min sind Voraussetzung für die ordnungsgemäße Funktion von MOS und LCR. VMP und
VLCR_min sind Voraussetzung für die ordnungsgemäße Funktion von MOS und LCR.
Übermäßiger Leistungsverlust ist definiert als
Der übermäßige Leistungsverlust wird definiert als

$$= \quad (4)$$

$$P_{loss_exc,n} = I_{ave,n} \times V_{LCR_exc,n}$$

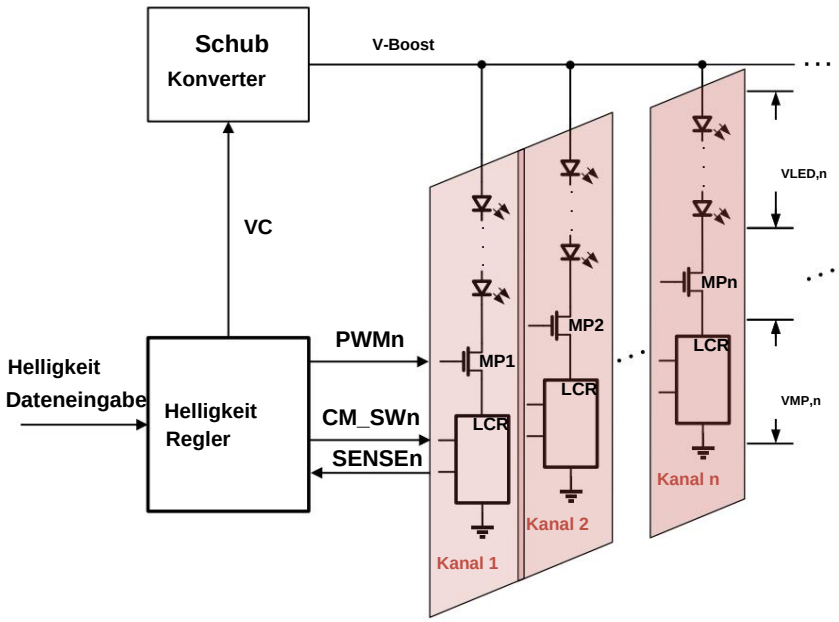


Abbildung 1. Struktur des vorgeschlagenen adaptiven hocheffizienten LED-Hintergrundbeleuchtungstreibers.
Abbildung 1. Struktur des vorgeschlagenen adaptiven hocheffizienten LED-Hintergrundbeleuchtungstreibers.

2.2. Adaptiver Steuerungsprozess
2.2. Adaptiver Steuerungsprozess

Ein adaptiver Steuerungsprozess, der für mehrkanalige LED-Treiber geeignet ist, soll übermäßige
Ein adaptiver Steuerungsprozess, der für mehrkanalige LED-Treiber geeignet ist, soll
Leistungsverluste durch zu hohe Spannung reduzieren. Beim Systemstart wird der
Reduzierung des übermäßigen Leistungsverlusts durch die Überspannung. Beim Systemstart leitet der
Helligkeitsregler einen Initialisierungsprozess ein, bei dem der Spitzenstrom und die
Der Helligkeitsregler leitet einen Initialisierungsprozess ein, bei dem der Spitzenstrom und der Arbeitszyklus
jedes LED-Kanals individuell eingestellt werden, um eine gleichmäßige Helligkeit zu erreichen und
Zyklus jedes LED-Kanals werden individuell eingestellt, um eine gleichmäßige Helligkeit zu erreichen und
gleichzeitig übermäßigen Leistungsverlust zu minimieren. Nach der Initialisierungsphase wird die Helligkeit
gleichzeitig zu geringe Leistungsverluste. Nach der Initialisierungsphase geht der Helligkeitsregler in die
Nachführphase über, in der er kontinuierlich den Status überwacht
Der Controller geht in die Tracking-Phase über, in der er den Status des LCR kontinuierlich überwacht. Die
Einstellungen für Spitzenstrom und Arbeitszyklus werden dann entsprechend angepasst, um
der LCR. Die Spitzenstrom- und Arbeitszykluseinstellungen werden dann entsprechend angepasst, um
Schwankungen bei Stromversorgung, Temperatur und anderen Faktoren auszugleichen.
Mäßige Schwankungen bei Stromversorgung, Temperatur und anderen Faktoren.

Abbildung 2 zeigt das Flussdiagramm der Initialisierungsphase. Zunächst wird der Helligkeitseingang
Abbildung 2 zeigt das Flussdiagramm der Initialisierungsphase. Zunächst werden die
Helligkeitseingangsdaten , einschließlich des nominalen Spitzenstroms Iset und des Arbeitszyklus yset,
Daten, einschließlich des nominalen Spitzenstroms Iset und des Arbeitszyklus yset, werden in den Helligkeitsregler
geladen . Anschließend wird der LCR-Strom jedes Kanals über die
Helligkeitsregler. Anschließend wird der LCR-Strom jedes Kanals über das CM_SW-Signal auf
Iset eingestellt . Der Helligkeitsregler beginnt dann, VC zu erhöhen , wodurch Vboost
CM_SW-Signal. Der Helligkeitsregler beginnt dann, VC zu erhöhen, wodurch Vboost entsprechend abnimmt .
Gleichzeitig überwacht der Helligkeitsregler die
entsprechend erhöht. Gleichzeitig überwacht der Helligkeitsregler das Statussignal SENSE von jedem Kanal.
Sobald alle SENSE-Signale von 0 auf 1 wechseln, wird der
tus-Signal SENSE von jedem Kanal. Sobald alle SENSE-Signale von 0 auf 1 wechseln, stoppt der Anstieg
von VC . An diesem Punkt haben die Spitzenströme aller Kanäle den voreingestellten Wert erreicht.
Anstieg in VC stoppt. An diesem Punkt haben die Spitzenströme aller Kanäle den voreingestellten Wert von Iset
erreicht, und es gibt immer einen Kanal, angenommen Kanal k, dessen SENSE

Das Signal ändert sich als letztes.

Machine Translated by Google

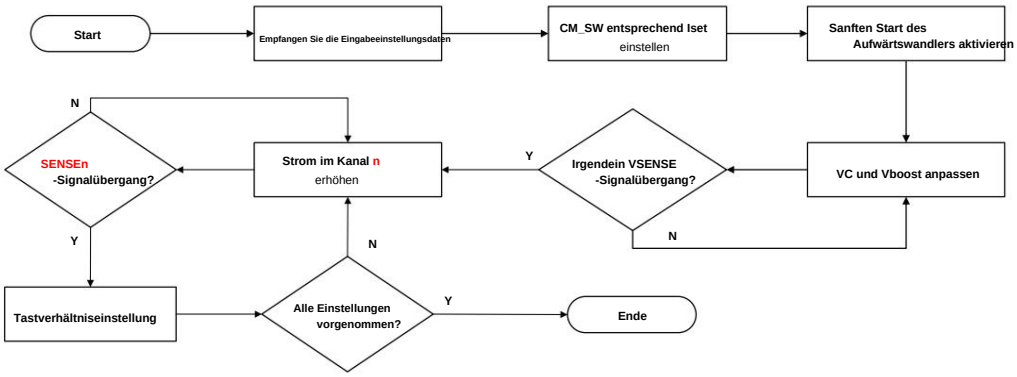


Abbildung 2. Flussdiagramm des adaptiven Steuerungsprozesses in der Initialisierungsphase.

Als nächstes wird nacheinander auf allen Kanälen außer Kanal k ein Spitzenstromanpassungsschritt ausgeführt . Der Helligkeitsregler versucht, den Kanalstrom tt durch Bereitstellung einer neuen CM_SW-Einstellung zu erhöhen. Wenn SENSE tt auf 1 bleibt, setzt der Helligkeitsregler diesen Vorgang fort. Wenn SENSE jedoch von 1 auf 0 wechselt, kehrt der Helligkeitsregler zur vorherigen CM_SW-Einstellung zurück und schließt den Anpassungsprozess tt für diesen Kanal ab. Der entsprechende Kanal hat dann einen neuen Spitzenstrom. Der Helligkeitsregler führt den Anpassungsprozess für den nachfolgenden Kanal weiter durch. Sobald alle Stromanpassungen abgeschlossen sind, nähert sich die Spannung an jedem LCR VLCR_min und die überschüssige Komponente wurde minimiert.

Um die gleiche Helligkeit zu gewährleisten, muss jeder LED-Kanal den gleichen Durchschnittsstrom haben. Daher wird der Arbeitszyklus von Kanal n (ÿn) vom Helligkeitsregler wie folgt angepasst:

$$\ddot{y}n = \ddot{y}set \times \frac{Iset^{N_{Satz}}}{ISpitze,n^{I_{CH}}} \quad (5)$$

wobei Ipeak,n der Spitzenstrom des n-ten LED-Kanals nach dem oben beschriebenen Anpassungsprozess ist. ÿset und Iset stellen den anfänglichen Arbeitszyklus bzw. den Spitzenstrom dar. Es ist anzumerken, dass gemäß Gleichung (4) der Durchschnittsstrom unverändert bleibt.

Durch den oben genannten Prozess werden der Spitzenstrom Ipeak,n und der Arbeitszyklus ÿn für jeden LED-Kanal angepasst, wodurch zusätzlicher Stromverbrauch durch übermäßige Spannung minimiert wird. Der angepasste Durchschnittsstrom entspricht weitgehend dem anfänglich eingestellten Durchschnittsstrom. Da die Helligkeitsabgabe von LEDs direkt mit ihrem Strom korreliert, stellt die adaptive Steuerungsmethode sicher, dass die Helligkeitsstufen über alle LED-Kanäle hinweg konsistent bleiben. Beim Eintritt in die Tracking-Phase ähnelt das Hauptprinzip dem der Initialisierungsphase. Beide kombinieren die Anpassung von Spitzenstrom und Arbeitszyklus ur

2.3. Linearer Stromregler

Da die LED-Helligkeit direkt von ihrem Strom beeinflusst wird, erfordert eine präzise Helligkeitssteuerung eine genaue Stromregelung. In vielen LED-Hintergrundbeleuchtungstreiberlösungen wird typischerweise ein LCR zur Erzeugung und Steuerung des LED-Stroms eingesetzt. Abbildung 3 zeigt zwei typische LCR-Schaltungskonfigurationen [37]. Ein wichtiges Designziel für LCR-Schaltungen ist das Erreichen einer hohen Stromgenauigkeit. In Abbildung 3a bilden der Operationsverstärker und der Transistor M1 eine negative Rückkopplung, die sicherstellt, dass VS gleich VREF ist. Folglich **ist der Ausgangsstrom ILED VREF/RS. Die Genauigkeit von RS wirkt sich offensichtlich direkt auf** die Präzision **von ILED** aus . **Es ist wichtig zu beachten, dass die Antriebsströme für Hintergrundbeleuchtungs-LEDs im Allgemeinen beträchtlich sind. In diesem Dokument wird der maximale Antriebsstrom auf 48 mA ausgelegt. Um den Leistungsverlust zu minimieren ,** sollten **VS oder VREF** außerdem **so niedrig wie möglich gehalten werden, was einen kleinen RS** erfordert , **um hohe Ströme zu erreichen .** Ein zu kleiner **RS** kann **jedoch die negativen Auswirkungen von Schaltungsparasiten auf die Genauigkeit des Ausgangsstroms erheblich verstärken.**

Machine Translated by Google

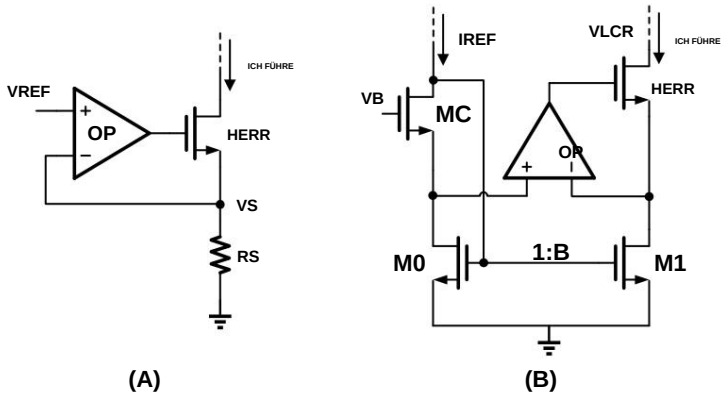


Abbildung 3. Typische lineare Stromreglerschaltungen. (a) Quellregelungsstruktur; (b) Spiegel aktuelle Struktur.

Abbildung 3b zeigt einen anderen LCR-Typ. In dieser Konfiguration wird der Ausgangsstrom ILED nicht direkt wie in Abbildung 3a erzeugt, sondern durch Stromspiegelung. In dieser Struktur wird **ILED durch B × IREF bestimmt, wobei IREF ein Referenzstrom und B das Stromspiegelverhältnis ist, das durch das Größenverhältnis der Transistoren M1 zu M0 definiert wird. Der Operationsverstärker (OP) spielt eine entscheidende Rolle, indem er sicherstellt, dass die Transistoren M1 und M0 die gleiche VDS haben, wodurch eine genaue Stromspiegelung gewährleistet wird.**

Tatsächlich können M0 und M1 im linearen Bereich statt im Sättigungsbereich arbeiten, indem VB gesenkt wird, wodurch **VDS reduziert wird , was dazu beiträgt, VLCR** zu senken **und übermäßigen Stromverbrauch** zu reduzieren **[38]. Der klassische Drainstromausdruck für MOS-Transistoren, die im linearen Bereich arbeiten, lautet**

$$C(\quad) = [2(V_{GS} - V_{TH}) - V_{DS}]V_{DS} \quad \text{ID} = \frac{1}{2}\mu_{COX}\frac{W}{L}[2(V_{GS} - V_{TH})V_{DS} - V_{DS}^2] \quad (6)$$

Der Transistor kann nun ungefähr einem linearen Widerstand entsprechen. Wie oben erwähnt, werden aufgrund des Effekts von OP die VDS von M0 und M1 gleich gehalten. Dann hat der Drainstrom von M0 und M1 eine Beziehung wie ff

$$\frac{\text{Nr. M0}}{\text{Nr. M1}} = \frac{(W/N)M0}{(W/L)M1} = \frac{1}{B} \quad (7)$$

(WL) B Eine exakte Stromspiegelung ist selbstverständlich gewährleistet.

Basierend auf der in Abbildung 3b gezeigten Schaltungsstruktur wurde ein verbesserter LCR für die adaptive Steuerung entwickelt, wie in Abbildung 4 dargestellt. Der Hauptunterschied zur Schaltung in Abbildung 3b ist die Einführung des Operationsverstärkers OP2. OP2 bildet in Verbindung mit dem Transistor M0 eine negative Rückkopplungsschleife, die sicherstellt, dass die Source-Drain-Spannung VDS von M0, in der Abbildung als VA bezeichnet , **gleich VREF bleibt. Durch Anpassen** von **VREF** kann **die Source** - **Drain-Spannung VDS von M0 präzise gesteuert werden. Der Vorteil des Designs besteht darin, dass es die Auswirkungen von Prozessschwankungen abmildert und eine g**

Wenn VREF niedrig ist, arbeitet der Transistor M0 im linearen Bereich. Die Transistoren M1, M2, ... , MN dienen als Stromspiegeltransistoren, die M0 entsprechen und jeweils die gleiche Größe haben. Der Decoder bestimmt anhand des binären Steuerworts CM_SW, ob jeder Spiegeltransistor aktiviert oder deaktiviert ist , und passt so den gesamten Ausgangsstrom an. Wenn das Gate eines Spiegeltransistors mit **VCOM verbunden ist , wird es eingeschaltet; wenn es mit Masse verbunden ist, wird es ausgeschaltet.**

OP1 funktioniert ähnlich wie OP in Abbildung 3b und stellt sicher, dass die Spannung am Knoten D (bezeichnet als VD) mit VA übereinstimmt , wodurch eine genaue Stromspiegelung aufrechterhalten wird. Der Komparator COMP arbeitet als einfacher Schaltkreiszustandsdetektor, wobei VCOM in diesem Design auf 2,5 V eingestellt ist. OP1 ist ein Operationsverstärker mit hoher Verstärkung und einer verstärkungssteigernden Ausgangsstruktur [39], während OP2 ein **Operationsverstärker mit einer Rail-to-Tail-Klasse-AB-Ausgangsstufe ist [40].**

Machine Translated by Google

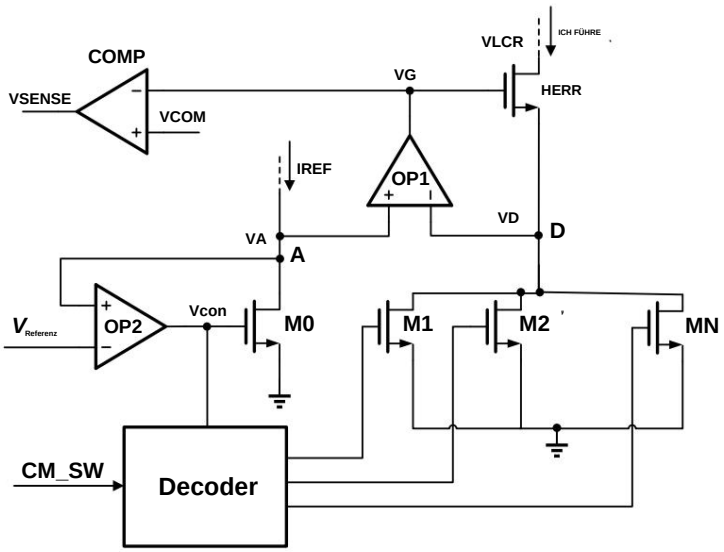


Abbildung 4. LCR-Schaltung zur adaptiven Steuerung.

Wenn Vboost ausreichend hoch ist, **bleibt selbst nach Berücksichtigung des Spannungsabfalls über die** LED-Kette genügend **VLCR** übrig , **um den ordnungsgemäßen Betrieb des LCR sicherzustellen, sodass ILED seinen eingestellten Wert erreichen oder beibehalten kann. Unter dieser Bedingung beträgt die Gate-Spannung des Transistors MR gemäß den Simulationsergebnissen an der typischen Prozessecke ungefähr 1,86 V (VGS, MR + VREF) , und SENSE bleibt bei 1. Wenn Vboost abnimmt oder ILED zunimmt und die Durchlassspannung der LED-Kette zunimmt, nimmt VLCR** entsprechend ab. **Solange der Transistor MR jedoch wirksam bleibt, kann der LCR weiterhin normal funktionieren. Wenn VLCR weiter abnimmt und den Transistor MR in den tiefen linearen Bereich drückt,** muss **VG erhöht werden, um den Auswirkungen des abnehmenden VLCR entgegenzuwirken. Wenn VG VCOM überschreitet , ändert der Ausgang des Komparators seinen Zustand und SENSE wechselt auf 0. Dies zeigt an ,** dass der LCR möglicherweise nicht mehr in der Lage ist, den eingestellten **ILED aufrechtzuerhalten oder zu**

Alle Spiegeltransistoren haben die gleiche Größe und jeder Transistor hat ein Größenverhältnis von B:1 mit M0. Der LED-Strom kann mit folgender Gleichung bestimmt werden:

$$ILED = N \times B \times IREF \quad (8)$$

wobei N die Anzahl der aktivierten Spiegeltransistoren ist und durch CM_SW bestimmt wird. IREF ist der Referenzstrom, der von der externen Vorspannungsschaltung erzeugt wird. Es ist zu beachten, dass der ff

Der Ausgangsstrom ILED des LCR bildet tatsächlich den oben genannten Spitzenstrom Ipeak.

Sogar auf demselben Chip gibt es ein gewisses Maß an Prozessabweichungen, die sich in erster Linie auf die Gerätegröße und die charakteristischen Parameter auswirken und zu Ungenauigkeiten oder Fehlanpassungen bei der Stromspiegelung führen. Unter Berücksichtigung des Einflusses des Breite-Länge-Verhältnisses (W/L) und der Schwellenspannung (Vth) kann die durch Prozessabweichungen verursachte Stromänderung (WL) wie folgt ausgedrückt werden:

$$\ddot{y}Vth\ \ddot{y}(B/L)\ \ddot{y}Vth - \ddot{y}(B/L) + \ddot{y}I = \frac{\ddot{y}Ich}{} \quad (9)$$

Die Stromfehlanpassung wird üblicherweise als die Änderungsrate des Stroms definiert. Unter der Annahme kein Zusammenhang zwischen den verschiedenen Einflussfaktoren besteht, lässt es sich wie folgt ausdrücken:

$$\frac{\ddot{y}I}{} = \frac{1}{} \quad \frac{\ddot{y}Ich}{} \ddot{y}(B/L) + \ddot{y}Vth\ \frac{\ddot{y}Ich}{} \ddot{y}(B/L)\ \ddot{y}Vth \quad (10)$$

Wenn der Transistor im Sättigungsbereich arbeitet, entsprechend dem entsprechenden Drain aktuelle Formel, die aktuelle Nichtübereinstimmung ist

$$\frac{\ddot{y}I}{} = \frac{\ddot{y}(B/N)}{} + \frac{2\ddot{y}Vth}{(VGS\ \ddot{y}\ Vth)} \quad (11)$$

Machine Translated by Google

Wenn der Transistor jedoch im linearen Bereich arbeitet, wird die Stromfehlanpassung gemäß der entsprechenden Drainstromformel

$$\frac{\ddot{y}I}{_{ICH}} = \frac{\ddot{y}(B/N)}{_{(Sieg/Niederlage)}} + \frac{2\ddot{y}Vth}{2(VGS\ \ddot{y}\ Vth)\ \ddot{y}\ VDS} \quad (12)$$

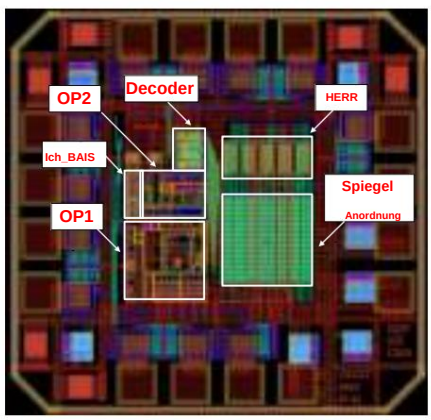
Vergleicht man die beiden obigen Formeln, sieht man, dass die Auswirkung der Größenvariation auf die Stromfehlanpassung sowohl im linearen als auch im Sättigungsbereich gleich ist. Die Auswirkung der Schwellenspannungsvariation ist im linearen Bereich jedoch geringer, was darauf hindeutet, dass es möglich ist, im linearen Bereich eine geringere Stromfehlanpassung

3. Simulation und Messergebnisse

3.1. System- und Chipdesign

Basierend auf Abbildung 2 konstruiert und validiert dieses Dokument einen adaptiven dreikanaligen LED- Treiber, wobei jeder Kanal aus 6 LEDs besteht. Es wird ein Aufwärtswandlerchip verwendet, insbesondere der TPS61046 von Texas Instruments. Der LCR wird mithilfe eines 0,18 µm komplementären Metalloxidhalbleiters (CMOS) mit 5 V entwickelt. Da dieser Prozess keine Hochspannungsgeräte unterstützt, werden externe MP-Transistoren verwendet. Im Idealfall könnten diese MP-Transistoren, wenn ein BCD-Prozess verfügbar ist, direkt auf dem Chip integriert werden. Darüber hinaus ist eine weitere Integration des Helligkeitsreglers vorgesehen, um einen monolithisch integrierten Treiberchip zu erreichen, was eine geplante zukünftige Arbeit ist. Bei der Schaltungssimulation werden PSPICE-Modelle für den Aufwärtswandler, die LEDs und die MP-Transistoren verwendet. Der Einschaltspannungsparameter im PSPICE-Modell für LED wird absichtlich angepasst, um Unterschiede zwischen den Kanälen einzuführen. Insbesondere für Kanal 3 ist die LED-Einschaltspannung auf 3,20 V eingestellt, was einen typischen Wert darstellt. Für die Kanäle 1 und 2 sind die Parameter auf 2,73 V bzw. 2,76 V eingestellt. Wie bereits beschrieben, wird der Mindestwert von Vboost durch die VLED von Kanal 3 bestimmt. Der Helligkeitsregler ist mit Verilog HDL implementiert und läuft mit einer Frequenz von 100 MHz. Die Frequenz des PWM-Signals beträgt 10 KHz. Die gesamte Schaltung läuft in einem Mixed-Signal-Simulationsmodus, um den adaptiven Steuerungsprozess zu überprüfen und die Leistung des LCR zu bewerten.

Unter Berücksichtigung des Stromsteuerbereichs und der Auflösung wird die Größe von M0 im LCR auf 1 µm/1 µm gewählt und sein Verhältnis zum Spiegeltransistor auf 1:4 eingestellt. Insgesamt gibt es 240 Stromspiegeltransistoren . Der Referenzstrom IREF wird auf 50 µA eingestellt, was zu einem Einheitsausgangsstrom von 200 µA und einem maximalen Ausgangsstrom von 48 mA führt. Die Größe von MR wird auf 40 µm/2 µm × 50 gewählt. Das Gesamtlayout der LCR-Schaltung ist in Abbildung 5 dargestellt. Alle Spiegeltransistoren und ihre entsprechenden Steuerschalter sind in Form eines 15 × 16-Arrays angeordnet. Der Kernbereich beträgt nur 325 µm × 310 µm, während der gesamte Bereich einschließlich der Pads 700 µm × 720 µm beträgt. Die wichtigsten Transistorgrößen sind auch in Abbildung 5 zusammengefasst.



Transistor	W/L
MR	40µm/2µm×50
M0	1µm/1µm
M1-MN	4µm/1µm

Abbildung 5. Layout des LCR, entwickelt in einem 0,18-µm-CMOS-Prozess, die Kernfläche beträgt 325 × 310 µm2 .

Abbildung 5. Layout des LCR, entworfen in einem 0,18-µm-CMOS-Prozess, die Kernfläche beträgt 325 × 310 µm2 .
Außerdem sind die wichtigsten Transistorgrößen zusammengefasst.
Außerdem sind die wichtigsten Transistorgrößen zusammengefasst.

3.2. Simulationsergebnisse

Abbildung 6 zeigt die simulierten Ausgangseigenschaften des LCR unter verschiedenen Prozessbedingungen und Temperaturen. Es ist ersichtlich, dass der Ausgangsstrom des LCR im Bereich von 0 mA bis 48 mA unter verschiedenen Betriebsbedingungen eine gute Linearität aufweist. Konkret beträgt bei einer Ausgangsstromeinstellung von 20 mA die maximale Stromabweichung 13

Machine Translated by Google

3.2. Simulationsergebnisse

Abbildung 6 zeigt die simulierten Ausgangseigenschaften des LCR unter verschiedenen Prozessbedingungen und Temperaturen. Es ist ersichtlich, dass der LCR bei einem Ausgangsstrombereich von 0 mA bis 48 mA unter verschiedenen Betriebsbedingungen eine gute Li Konkret beträgt bei einer Ausgangsstromeinstellung von 20 mA die maximale Stromabweichung 13 µA, was zu einem Stromfehler von weniger als 0,1 % führt. Insgesamt weist der LCR eine gute Stromlinearität und Genauigkeit innerhalb des angegebenen Strombereichs auf.

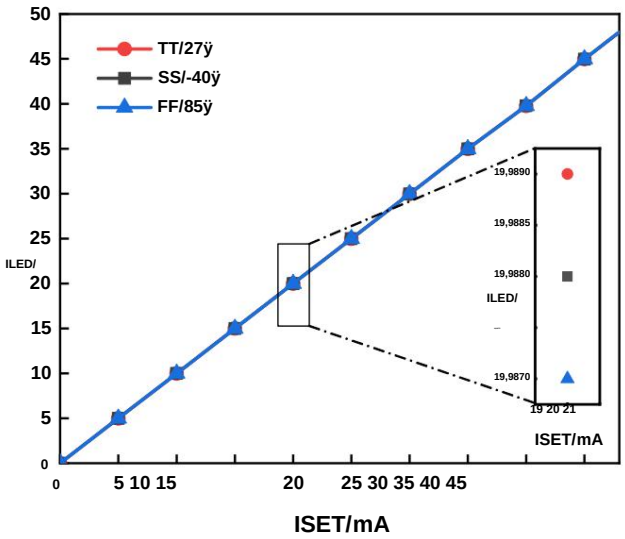


Abbildung 6. Ausgangsstrom des LCR im Vergleich zur Stromeinstellung unter verschiedenen Arbeitsbedingungen.

Abbildung 7 zeigt die Spannungswellenform an mehreren wichtigen Knoten während des Anpassungsprozesses . SENSE1, SENSE2 und SENSE3 stellen die SENSE-Signale der entsprechenden Kanäle dar. T0 bezeichnet die Softstartphase des Aufwärtswandlers. Beim Einschalten erreicht der Ausgang des Aufwärtswandlers Vboost nach einer gewissen Zeit allmählich einen voreingestellten Wert. Dieser voreingestellte Wert ist die Mindestspannung für den ordnungsgemäßen Betrieb der Schaltung unter verschiedenen Arbeitsbedingungen und ist in diesem Design auf 25,19 V eingestellt. Wenn Vboost zunimmt, wechseln die SENSE-Signale jedes Kanals nacheinander auf 0. T1 markiert die Vboost- Anpassungsphase . Während dieser Zeit erhöht der Helligkeitsregler VC allmählich, was zu einer entsprechenden Verringerung von Vboost führt. Es ist ersichtlich, dass, wenn Vboost 21,52 V erreicht, SENSE3 auf 1 wechselt, was eine Änderung des Status von Kanal 3 anzeigt. Schließlich bleibt **VC bei 0,33 V und Vboost stabilisiert sich bei 21,62 V. T2 kennzeichnet die Stromregelungsphase, während der der Strom von Kanal 1 (Kanal 2) zunimmt. Wenn SENSE1 (SENSE2) auf 1 wechselt, zeigt dies an, dass sich der Status von Kanal 1 (Kanal 2) geändert hat. Die Stromregelung von Kanal 1 (Kanal 2) wird beendet und der Arbeitszyklus von Kanal 1 (Kanal 2) wird gemäß Gleichung (4) eingestellt. T3 stellt die normale Arbeitsphase dar, die auch als kontinuierliche Tracking-Phase bezeichnet wird, und SENSE-Signale werden dynamisch auf 0 g**

Abbildung 8 zeigt die Stromwellenform in jedem Kanal. Zunächst ist der Spitzenstromwert jedes Kanals auf 20 mA eingestellt, mit einem Standard-Arbeitszyklus von 100 %. Während der frühen Phase von T0 ist **Vboost noch nicht vollständig aufgebaut und alle LEDs sind ausgeschaltet, was zu einem Nullstrom in jedem Kanal führt. Da die LEDs in Kanal 3 die höchste Durchlassspannung haben , steigt der Strom in Kanal 3 als letzter auf den eingestellten Wert, was mit dem Timing der SENSE-Signale in Abbildung 7 übereinstimmt. Nach Abschluss der Vboost- Anpassungsphase beginnt die Stromregelungsphase. Es ist ersichtlich, dass die Spitzenströme in Kanal 1 und Kanal 2 auf 24,2 mA bzw. 23,2 mA erhöht werden, wobei die entsprechenden Arbeitszyklen auf 82,6 % bzw. 86,2 % angepasst werden, wodurch gleichmäßige Durchschnittsströme über die drei Kanäle hinweg sichergestellt werden. Wenn die normale Arbeitsphase beginnt, überwacht der Helligkeitsregler kontinuierlich den Zustand des LCR. Wenn sich der Zustand ändert,**

Machine Translated by Google

auf 0. Basierend auf der simulierten Stromwellenform betragen die maximalen Anstiegs- und Abfallzeiten 173 ns bzw. 74 ns, ohne dass ein merkliches Überschwingen des Stroms auftritt.

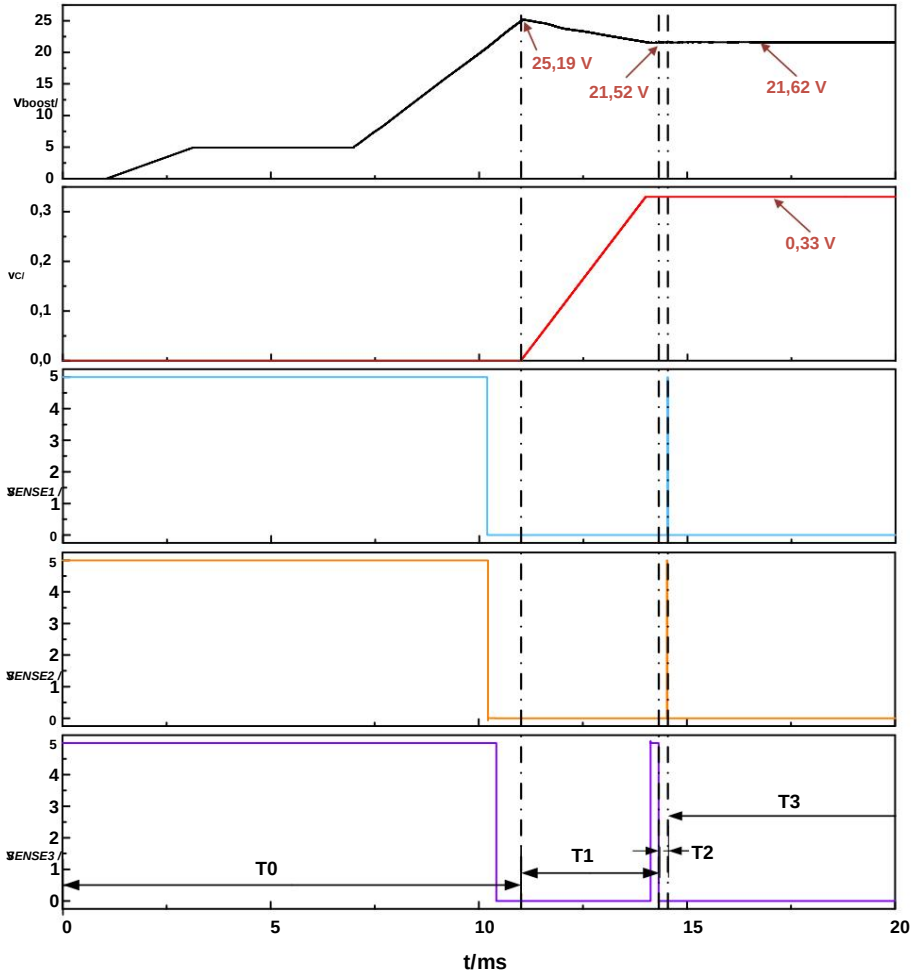


Abbildung 7. Spannungswellenformen im adaptiven Steuerungsprozess.

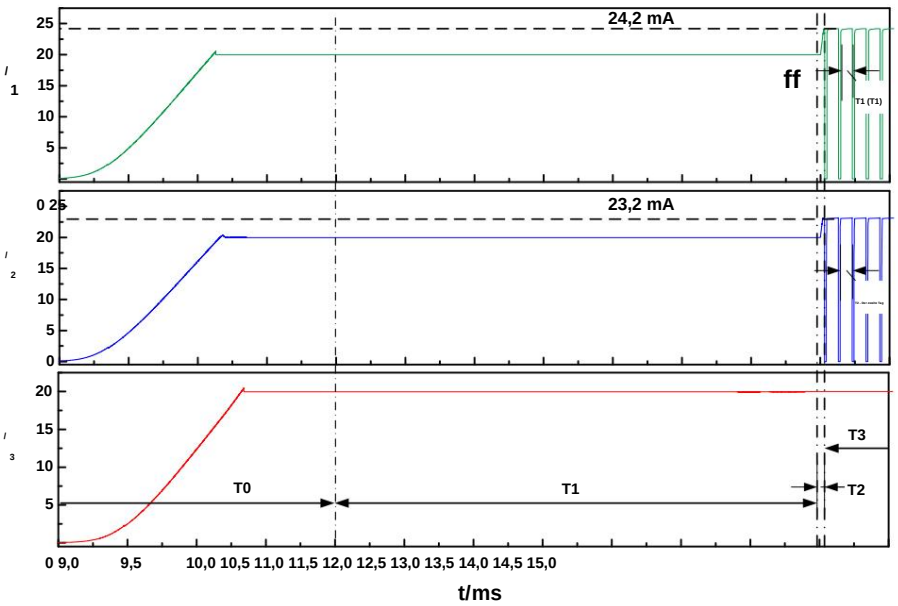


Abbildung 8. Stromwellenformen jedes Kanals im adaptiven Steuerungsprozess.

ff

Machine Translated by Google

ff

Abbildung 9 vergleicht die Leistungsverlustunterschiede zwischen der Steuermethode mit festem Spitzenstrom und der in diesem Dokument vorgeschlagenen adaptiven Steuermethode. Bei der Methode mit festem Spitzenstrom wird nur der Vboost- Anpassungsschritt ausgeführt, während der Spitzenstrom und der Arbeitszyklus unverändert bleiben. Wie bereits zuvor ff **beschrieben, erfährt Kanal 3 bei der Methode mit festem Spitzenstrom den geringsten Leistungsverlust, während Kanal 1 und Kanal 2 aufgrund übermäßiger Spannung größere Leistungsverluste erleiden. Die Abbildung veranschaulicht den Leistungsverlust der drei Kanäle bei der Methode mit festem Spitzenstrom. Im Gegensatz** ff **dazu werden bei der adaptiven Steuermethode sowohl der Spitzenstrom als auch der Arbeitszyklus angepasst, wodurch übermäßige Spar**
Wie in der Abbildung dargestellt, **reduziert die adaptive Steuerungsmethode den Leistungsverlust um 74 %, wenn der nominale Kanalstrom auf 20 mA eingestellt ist .**

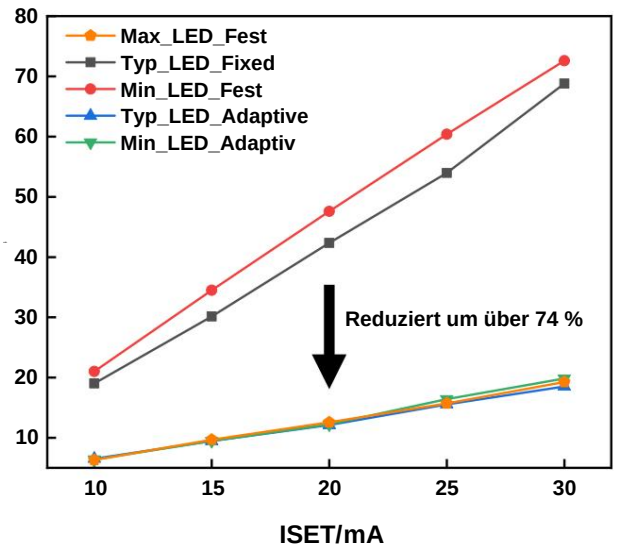


Abbildung 9. Leistungsverluste von LCR mit der Methode des festen Spitzenstroms und der vorgeschlagenen adaptiven Steuerungsmethode.

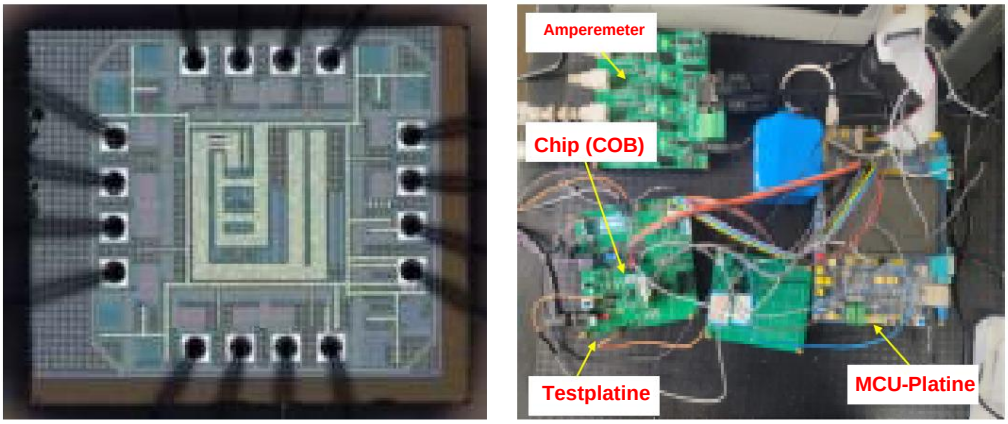
3.3. Messergebnisse

Abbildung 10 zeigt eine Mikrofotografie des hergestellten Chips mit Verbindungsdrähten sowie den Chip-Testaufbau. Der Chip ist in einem Chip-on-Board-Gehäuse (COB) untergebracht und auf der Testplatine montiert. Der Referenzstrom IREF wird von der Testplatine geliefert. Eine Entwicklungsplatine STM32F407 MCU (von ST Microelectronics) dient als Helligkeitsregler, und der integrierte DAC in der MCU wird zur Erzeugung von VC verwendet. Ein mehrkanaliges Amperemeter ist in Reihe zwischen jeder LED-Kette und dem Drain des Hochspannungs-MOS- Transistors geschaltet, um den Durchschnittsstrom der LED-Kanäle zu messen. Dieses Amperemeter unterstützt die Echtzeit-Datenaufzeichnung über einen seriellen Anschluss. Spannungswellenformen an Schaltungsknoten werden mit einem digitalen Oszilloskop erfasst, insbesondere dem RTB2004 von Rohde & Schwarz. Der Einfachheit halber werden beim Testen nur zwei LED-Kanäle implementiert, die jeweils aus 6 LEDs bestehen. Es ist bemerkenswert, dass die LEDs absichtlich so ausgewählt werden, dass ein Unterschied von ungefähr 500 mV in der Durchlassspannung über die LED-Kette zwischen den beiden Kanälen entsteht. Kanal 2 hat die kleinere Durchlassspannung. Dieser gewollte Unterschied verstärkt die Demonstration der Wirksamkeit der Schaltung.

Die Innentemperatur wurde bei etwa 27 °C gehalten. Abbildung 11 zeigt die gemessenen Spannungswellenformen von VPWM und Vloss für Kanal 2 nach dem anfänglichen adaptiven Steuerungsprozess. Die Kanalstromeinstellung Iset beträgt 20 mA. Die vier Abbildungen entsprechen dem Fall, wenn VREF auf 1 V, 0,6 V, 0,4 V bzw. 0,3 V eingestellt ist. Der Arbeitszyklus von VPWM wird in jedem Fall mit 68,57 %, 68,86 %, 68,43 % bzw. 68,53 % aufgezeichnet. Wenn VPWM hoch ist, wird der Kanal eingeschaltet und Vloss wird in jedem Fall mit 1,06 V, 0,69 V, 0,49 V bzw. 0,36 V gemessen. Wenn VPWM auf niedrig wechselt, wird der Kanal geschlossen und der Anstieg von Vloss ist hauptsächlich auf

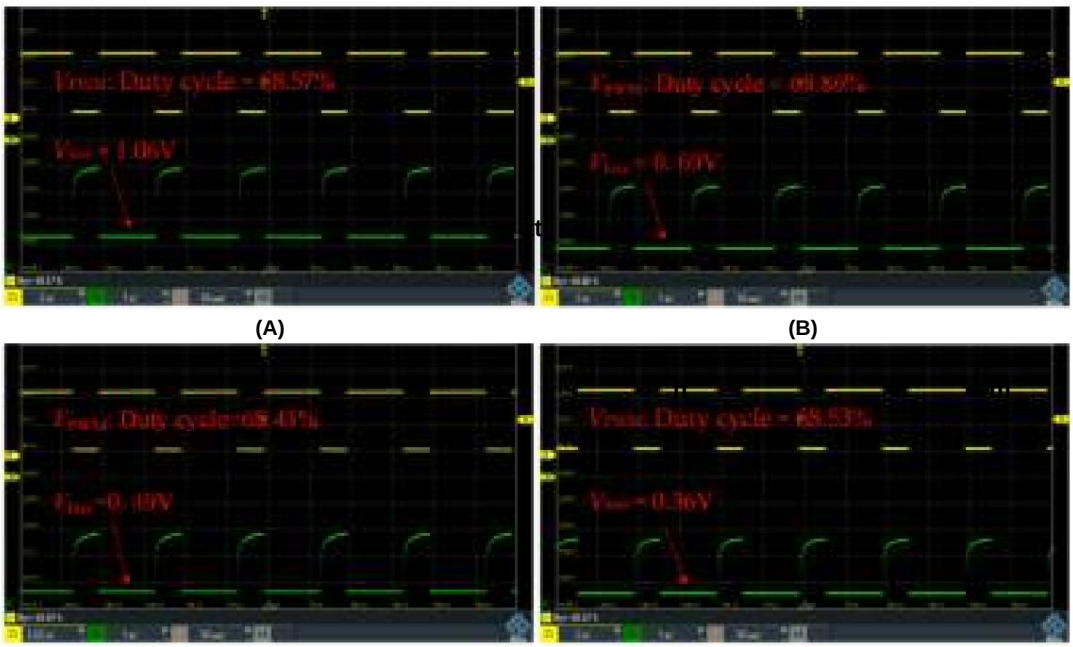
Machine Translated by Google

Aus den Abbildungen lässt sich schließen, dass trotz des Rückgangs **von VREF von 1 V auf 0,3 V, der dazu führen kann, dass die Stromspiegeltransistoren im LCR in den linearen Bereich eintreten, die Genauigkeit des Ausgangsstroms weiterhin aufrechterhalten werden kann und der Arbeitszyklus von VPWM bei diesen Einstellungen konsistent bleibt .**



(A) **(B)**

Abbildung 10. (a) Mikrofotografie des LCR-Chips; (b) Testsystem.



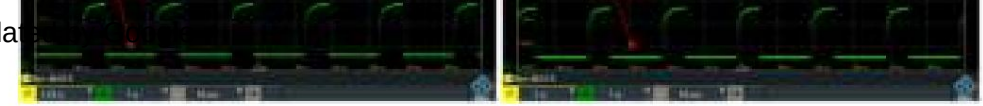
(A) **(B)**

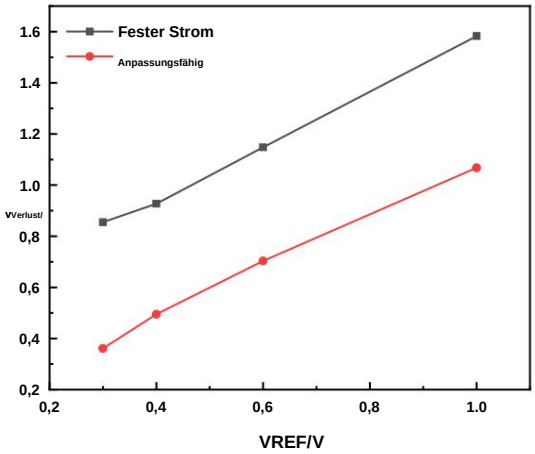
(C) **(D)**

Abbildung 11. Spannungswellenformen von VPWM und Vloss unter (a) VREF = 0,3 V, (b) VREF = 0,4 V, (c) VREF = 0,6 V und (d) VREF = 1 V. Iset = 20 mA.

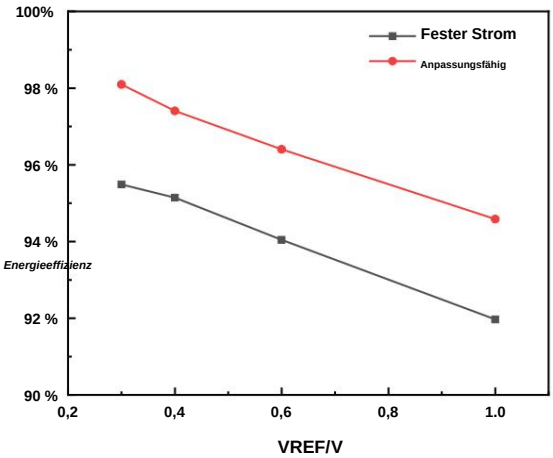
Abbildung 12 **zeigt den Unterschied zwischen Vloss und Energieeffizienz von Kanal 2 sowohl bei der Steuerungsmethode mit festem Spitzenstrom als auch bei der adaptiven Steuerungsmethode. In Abbildung 12a (Adaptiv) ist sehr deutlich zu erkennen, dass in Kanal 2 eine übermäßige Spannung von etwa 500 mV vorliegt. Mit der adaptiven Steuerungsmethode wird diese übermäßige Spannung durch Anpassungen in ILED effektiv gemildert, was zu einer erheblichen Verbesserung der Energieeffizienz führt. Insbesondere wenn VREF auf 300 mV eingestellt ist, kann die Energieeffizienz bei einem ILED von 20 mA 98,1 % erreichen. Im Prinzip ist es möglich, die Energieeffizienz weiter zu verbessern, indem VREF noch weiter reduziert wird, wobei die Obergrenze durch VCOM begrenzt wird, das wiederum von der Stromversorgu**

Machine Translat





(A)



(B)

Abbildung 12. Vergleich zwischen der Methode des festen Spitzenstroms und der vorgeschlagenen adaptiven Steuerung Methode im Hinblick auf (a) Spannungsverlust und (b) Leistungseffizienz.

Tabelle 1 vergleicht verschiedene Parameter und Leistungskennzahlen zwischen den vorgeschlagenen adaptiver Treiber und andere vorhandene Referenzen. Das vorgeschlagene Design weist Vorteile wie einen breiten Ausgangsstrombereich und eine hohe Energieeffizienz auf. Tatsächlich zeigt T zeigt nur die Effizienz eines einzigen Kanals. Wichtiger noch ist jedoch, dass die vorgeschlagene Die adaptive Methode zielt darauf ab, die Effizienz jedes Kanals in einer Mehrkanal-LED zu verbessern Treiber. Es kompensiert effektiv übermäßige Spannung am LCR durch adaptive Steuerung mit mäßiger Implementierungskomplexität.

Tabelle 1. Vergleich der Leistungsparameter.

Ref.	[17]	[41] [42]	[43]	Diese Arbeit
Verfahren	0,35 µm BCD 0,5 µm CMOS	40 nm CMOS	40 nm CMOS	0,18 µm CMOS
Kernfläche (mm2)	N/A +	N/A	N/A +	0,1
Stromversorgung (V)	6–27	3–4,2	3,2–4,2	5
Maximaler ILED (mA)	30	2 6	25	48
LED-Kettenkonfiguration.		5,36 22,5		6
V-Boost (V)	8	~300	6	18,96 *
Vverlust (mV)	26,41	~675	20	360 *
Energieeffizienz	~400	94,4	~480	98,1 % *
Vboost- Anpassung	98,4	%	97,6	Ja
Kanalweise Anpassung	% Ja Nein	97 % Ja Ja Nein Nein	% Ja Nein	Ja

*** Gemessen bei Iset = 20 mA, ÿset = 100 %.** **+ Nicht verfügbar.**

Machine Translated by Google

4. Zusammenfassung

In diesem Dokument wird ein adaptiver hocheffizienter LED-Hintergrundbeleuchtungstreiber entwickelt, um das Problem des übermäßigen Leistungsverlusts zu lösen, der durch Schwankungen der Durchlassspannung von LEDs verurs
Durch die Anpassung des Spitzenstroms wird die Konsistenz der Durchlassspannung der LED-Kette innerhalb jedes Kanals erreicht. Dies minimiert die Spannung am LCR in jedem Kanal und optimiert so den übermäßigen Leistungsverlust in jedem Kanal und die Gesamtl
Darüber hinaus wird durch die Anpassung des Arbeitszyklus eine gleichbleibende Helligkeit über alle Kanäle hinweg gewährleistet.
Basierend auf dem 0,18 µm 1P4M CMOS-Prozess wurde eine für die adaptive Steuerung geeignete LCR-Schaltung entwickelt und hergestellt. Die Schaltung verwendet eine Stromspiegelstruktur und im linearen Bereich arbeitende MOS-Transistoren, um Stromgenauigkeit und Energieeffizienz sicherzustellen. Die entwickelte LCR-Schaltung unterstützt einen Ausgangsstrombereich von 0–48 mA. Um die Wirksamkeit der adaptiven Steuerschaltung zu überprüfen, wurde ein experimenteller Zweikanal-LED-Hintergrundbeleuchtungstreiber unter Verwendung der entwickelten LCR-Chips konstruiert. Jeder Kanal besteht aus 6 LEDs, und die Durchlassspannungsdifferenz zwischen den Kanälen wird absichtlich auf ungefähr 500 mV erweitert. Die Testergebnisse zeigen, dass die vorgestellte adaptive Steuerschaltung diese Durchlassspannungsdifferenz eliminiert und den übermäßigen Leistungsverlust reduziert. Wenn die Referenzspannung auf 300 mV eingestellt wird, wird bei 20 mA eine Energieeffizienz von 98,1 % erreicht. Tatsächlich kann der experimentelle Zweikanaltreiber auf eine Mehrkanalimplementierung erweitert werden. Während in dieser Studie nur ein LCR-Chip entworfen und hergestellt wird, können sowohl die MP-Transistoren als auch der Helligkei

Beiträge der Autoren: Konzeptualisierung, WZ und QZ; Methodik, XX, QZ, CJ und ZZ; Software, XX; Validierung, XX, QZ, CJ und ZZ; formale Analyse, XX; Untersuchung, XX, QZ, CJ und ZZ; Ressourcen, XX und QZ; Datenkuratierung, XX und QZ; Schreiben – Vorbereitung des Originalentwurfs, XX und QZ; Schreiben – Überprüfung und Bearbeitung, WZ und XX; Aufsicht, WZ und JT; Projektverwaltung, WZ und JT; Mittelbeschaffung, WZ. Alle Autoren haben die **veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu .**

Finanzierung: Diese Forschung wurde vom wissenschaftlichen **Forschungsprojekt (23A0260) des Bildungsministeriums der Provinz Hunan** und dem Graduierten-Praxis-Innovationsprojekt **(CLSJCX24079) der Universität Changsha für Wissenschaft und Technologie finanziert.**

Datenverfügbarkeitserklärung: Die in dieser Studie präsentierten Daten sind auf Anfrage beim entsprechenden Autor erhältlich.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Verweise

1. **Tsao, JY Festkörperbeleuchtung: Lampen, Chips und Materialien für morgen.** IEEE Circuits Devices Mag. **2004,** 20, **28–37.** [CrossRef]
Yadlapalli, RT; Narasipuram, RP; Kotapati, A. Ein Überblick über energieeffiziente Halbleiter-LED-Treibertopologien. Int. J. Energy **2.** Res. **2020, 44 , 612–630.**
3. **Morrow, RC LED-Beleuchtung im Gartenbau.** HortScience **2008,** 43, **1947–1950.** [CrossRef]
4. **Cho, J.; Park, JH; Kim, JK; Schubert, EF Weiße Leuchtdioden: Geschichte, Fortschritt und Zukunft.** Laser Photonics Rev. **2017,** 11, **1600147.** [CrossRef]
5. **Surkanti, PR; Mehrotra, D.; Furth, PM Mehrkanal-LED-Treiber mit genauer Stromanpassung für tragbare Anwendungen. In Proceedings of the 2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS), Dallas, TX, USA, 4.–7. August 2019; S. 941–944.**
6. **Pal, N.; Fish, A.; McIntyre, W.; Griesert, N.; Winter, G.; Eichhorn, T.; Pilawa-Podgurski, R.; Hanumolu, PK Ein 91 % effizienter 30-V- Hybrid-Boost-SC-Konverter-basierter LED-Hintergrundbeleuchtungstreiber in 180-nm-CMOS. In Proceedings der 2020 IEEE Custom Integrated Circuits Conference (CICC), Boston, MA, USA, 22.–25. März 2020; S. 1–4.**
7. **Pressman, A.** Design von Schaltnetzteilen; **McGraw-Hill, Inc.: New York, NY, USA, 1997.**
8. **Wang, Y.; Alonso, JM; Ruan, X. Ein Überblick über LED-Treiber und verwandte Technologien.** IEEE Trans. Ind. Electron. **2017,** 64, **5754–5765.** [Querverweis]
9. **Gupta, N.; Verma, D. Eine Überprüfung und Analyse der Topologien von Leuchtdioden-Treibern (LED). In Proceedings der 2021 IEEE 2nd International Conference on Electrical Power and Energy Systems (ICEPES), Bhopal, Indien, 10.–11. Dezember 2021; S. 1–6.**
10. **Oh, T.-J.; Cho, A.; Ki, S.-L.; Hwang, I.-C. Ein stromsparender und kostengünstiger digital gesteuerter Boost-LED-Treiber-IC für Hintergrundbeleuchtungen. In Proceedings der 2012 IEEE Asian Solid State Circuits Conference (A-SSCC), Kobe, Japan, 12.–14. November 2012; Seiten 237–240.**

Machine Translated by Google

11. Lee, S.-W.; Choe, H.-J.; Yun, J.-J. Leistungsverbesserung eines Boost-LED-Treibers mit hoher Spannungsverstärkung für randbeleuchtete LED-Hintergrundbeleuchtungen. IEEE Trans. Circuits Syst. II Express Briefs **2017,** 65, **481–485. [CrossRef]**

12. Chen, H.; Zhang, Y.; Ma, D. Ein SIMO-Parallelstring-Treiber-IC für dimmbare LED-Hintergrundbeleuchtung mit lokaler Busspannungsoptimierung und einzelne zeitgeteilte Regelschleife. IEEE Trans. Power Electron. **2011,** 27, **452–462. [CrossRef]**

13. Kim, H.-C.; Yoon, CS; Jeong, D.-K.; Kim, J. Ein eininduktiver, mehrkanaliger stromausgleichender LED-Treiber für Displays Anwendungen für Hintergrundbeleuchtung. IEEE Trans. Ind. Appl. **2014,** 50, **4077–4081. [CrossRef]**

14. Chen, J.-H.; Wang, S.-C.; Liu, Y.-H.; Cheng, Y.-S.; Yang, Z.-Z. RGB-LED-Hintergrundbeleuchtungs-Antriebssystem mit dynamischer Spannungsregelungsfunktion . In Proceedings der 2013 IEEE 10th International Conference on Power Electronics and Drive Systems (PEDS), Kitakyushu, Japan, 22.–25. April 2013; Seiten 349–353.

15. Liu, P.-J.; Hong, Z.-Y.; Hung, C.-C. Ein einstufiger AC–DC-RGB-LED-Treiber mit geringem Stromverbrauch und Schaltkondensator-Steuerungsschema. IEEE Trans. Ind. Electron. **2020,** 68, **3103–3112. [CrossRef]**

16. Kim, J.-S.; Lee, Y.-K.; Lee, J.-S.; Shin, Y.-K.; Tark, J.-H.; Ryu, K.-C.; Yang, B.-D. Ein schnell schaltender Stromimpulstreiber für LED- Hintergrundbeleuchtung. In Proceedings des 2009 IEEE International Symposium on Circuits and Systems, Taipei, Taiwan, 24.–27. Mai 2009; S. 1775–1778.

17. Hong, S.-I.; Han, J.-W.; Kim, D.-H.; Kwon, O.-K. Ein LED-Hintergrundbeleuchtungstreiber-IC mit Doppelschleifensteuerung für mittelgroße LCDs. In Proceedings der 2010 IEEE International Solid-State Circuits Conference-(ISSCC), San Francisco, CA, USA, 7.–11. Februar 2010; Seiten 116–117.

18. Huang, Y.-C.; Chen, H.-W. Ein neuartiger schnell schaltender Stromimpulstreiber für LED-Hintergrundbeleuchtungsanwendungen. In Proceedings der 2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), Nantou, Taiwan, 27.–29. Mai 2016; S. 1–2.

19. Ahn, H.-A.; Hong, S.-K.; Kwon, O.-K. Ein schnell schaltender Stromregler mit einer Methode zur Reduzierung der Anstiegszeit für hohe Dimmleistung Verhältnis von LED-Hintergrundbeleuchtungstreibern. IEEE Trans. Circuits Syst. II Express Briefs **2016,** 63, **1014–1018. [CrossRef]**

20. Han, H.; Zhang, F.; Liu, M. PWM-Dimmmethode für kondensatorgeklemmte Stromteilungsschaltungen in LED-Hintergrundbeleuchtungssystemen. IEEE J. Emerg. Sel. Top. Leistungselektron. **2018, 6, 1190–1197. [CrossRef]**

21. Hong, S.-I.; Nam, K.-S.; Jung, Y.-H.; Ahn, H.-A.; In, H.-J.; Kwon, O.-K. Hocheffizienter und zuverlässiger Treiber-IC für Leuchtdioden-Hintergrundbeleuchtung zur gleichmäßigen Stromansteuerung mittelgroßer Flüssigkristallanzeigen. J. Inf. Disp. **2012,** 13, **73–82. [CrossRef]**

22. Lim, B.-M.; Ko, Y.-H.; Jang, Y.-S.; Kwon, O.-H.; Han, S.-K.; Lee, S.-G. Ein 200-V-Buck-LED-Treiber mit 98,16 % Wirkungsgrad und integrierter Stromregelung zur Verbesserung der Stromgenauigkeit für großflächige Einzelstrang-LED-Hintergrundbeleuchtungsanwendungen. IEEE Trans. Power E **2015,** 31, **6416–6427. [CrossRef]**

23. Parthasarathy, MK; Ganesan, N. Ein neuartiger Einzelinduktor-Achtkanal-Leuchtdiodentreiber für Displays mit geringem Stromverbrauch Anwendungen für Hintergrundbeleuchtung. Turk. J. Electr. Eng. Comput. Sci. **2018,** 26, **2130–2139. [CrossRef]**

24. He, J.-P.; Zhang, B.; Gao, X.; Wu, C.-D.; Li, Z.-J.; Zhou, J.-Y. Eine hochpräzise analoge Dimmmethode für LEDs mit schwacher Spannung. IEEE Zugriff **2019, 7, 172362–172373. [CrossRef]**

25. Lin, C.-L.; Chen, S.-C.; Deng, M.-Y.; Ho, Y.-H.; Lin, C.-A.; Tsai, C.-L.; Liao, W.-S.; Liu, C.-I.; Wu, C.-E.; Peng, J.-T. AM PWM-Antriebsschaltung für Mini-LED-Hintergrundbeleuchtung in Flüssigkristallanzeigen. IEEE J. Electron Devices Soc. **2021, 9, 365–372. [CrossRef]**

26. Schubert, EF Leuchtdioden (2018); **Rensselaer Polytechnic Institute: Tory, NY, USA, 2018.**

27. Liu, C.-H.; Hsieh, C.-Y.; Hsieh, Y.-C.; Tai, T.-J.; Chen, K.-H. SAR-gesteuerte adaptive Ausschaltzeittechnik ohne Messwiderstand zur Erzielung eines hocheffizienten und genauen LED-Beleuchtungssystems. IEEE Trans. Circuits Syst. I Regul. Pap. **2009,** 57, **1384–1394.**

28. Lohaus, L.; Rossius, A.; Dietrich, S.; Wunderlich, R.; Heinen, S. Ein dimmbarer LED-Treiber mit resistiver DAC-Rückkopplungssteuerung zur adaptiven Spannungsregelung. In Proceedings of the 2013 IEEE Energy Conversion Congress and Exposition, Denver, CO, USA, 15.–19. September 2013; S. 3126–3133.

29. Hsieh, C.-Y.; Wei, Y.-S.; Chen, K.-H.; Lin, T.-C. Effizienter LED-Treiber mit einer adaptiven Referenzverfolgungstechnik. In Proceedings der 16. IEEE International Conference on Electronics, Circuits and Systems (ICECS 2009), Yasmine Hammamet, Tunesien, 13.–16. Dezember 2009; Seiten 291–294.

30. Liu, P.-J.; Hsu, S.-R.; Chang, C.-W.; Liao, C.-Y.; Chien, L.-H. Dimmbarer weißer LED-Treiber mit adaptiver Spannungsrückkopplungssteuerung. In Proceedings der 2015 IEEE 2nd International Future Energy Electronics Conference (IFEEC), Taipeh, Taiwan, 1.–4. November 2015; S. 1–4.

31. Jung, YH; Hong, SK; Kwon, OK Ein schnell transienter LED-Treiber mit adaptiver Frequenzregelung je nach Lastvariation für großformatige LCD-Hintergrundbeleuchtungen. J. Soc. Inf. Disp. **2017,** 25, **712–724. [CrossRef]**

32. Hsu, Y.-C.; Lin, J.-Y.; Chen, CC-P. Flächensparender und hocheffizienter RGB-LED-Treiber mit adaptiver Antriebsspannung und Energiespartechnik . Energies **2018,** 11, **1422. [CrossRef]**

33. Yong, W.; Baoxing, D.; Yi, L.; Ting, L.; Yintang, Y. Ein mehrkanaliger, hocheffizienter Boost-LED-Hintergrundbeleuchtungstreiber mit hohem Dimmverhältnis. In Proceedings der 2017 International Conference on Electronic Industry and Automation (EIA 2017), Suzhou, China, 23.–25. Juni 2017; S. 110–116.

34. Hsieh, Y.-T.; Liu, B.-D.; Wu, J.-F.; Fang, C.-L.; Tsai, H.-H.; Juang, Y.-Z. Ein LED-Treiber mit hohem Dimmverhältnis für LCD-Hintergrundbeleuchtungen. IEEE Trans. Power Electron. **2012,** 27, **4562–4570. [CrossRef]**

35. Chiu, H.-J.; Lo, Y.-K.; Chen, J.-T.; Cheng, S.-J.; Lin, C.-Y.; Mou, S.-C. Ein hocheffizienter dimmbarer LED-Treiber für Beleuchtungsanwendungen. IEEE Trans. Ind. Electron. **2009,** 57, **735–743. [CrossRef]**

36. Hu, Y.; Jiang, L.; Wen, W.; Yan, W.; Ding, R. Ein neuer adaptiver Antriebsspannungsansatz für LED-Treiber. In Proceedings der 2019 IEEE Applied Power Electronics Conference and Exposition (APEC), Anaheim, CA, USA, 17.–21. März 2019; S. 721–726.

Machine Translated by Google

37. Wu, C.-H.; Chen, C.-L. Hocheffiziente stromgeregelte Ladungspumpe für einen weißen LED-Treiber. IEEE Trans. Circuits Syst. II Express Briefs **2009,** 56, **763–767.**

38. Mu, H.; Geng, L.; Liu, J. Eine hochpräzise Konstantstromquelle für LED-Treiber. In Proceedings des Symposiums 2011 für Photonik und Optoelektronik (SOPO), Wuhan, China, 16.–18. Mai 2011; S. 1–4.

39. Sansen, WM Analog Design Essentials; **Springer Science & Business Media: Berlin, Deutschland, 2007; Band 859.**

40. Hogervorst, R.; Tero, JP; Eschauzier, RG; Huijsing, JH Ein kompakter, energieeffizienter 3 V CMOS Rail-to-Rail-Ein-/Ausgang Operationsverstärker für VLSI-Zellenbibliotheken. IEEE J. Solid-State Circuits **1994,** 29, **1505–1513.** **[CrossRef]**

41. Surkanti, PR; Mehrotra, D.; Gangineni, M.; Furth, PM Mehrsträngiger LED-Treiber mit genauer Stromanpassung und dynamischer Aufhebung von Durchlassspannungsfehlanpassungen. In Proceedings of the 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), Boston, MA, USA, 6.–9. August 2017; S. 229–232.

42. Shahroury, FR; Ahmad, HH; Abuishmais, I. Designaspekte eines Single-Output-Multi-String-WLED-Treibers mit 40 nm CMOS- Technologie. J. Low Power Electron. Appl. **2022,** 12, **5.** **[CrossRef]**

43. Abuishmais, I.; Shahroury, FR; Ahmad, H. Eine Designmethodik für hocheffiziente dimmbare Stromsenken für stromgeregelte Treiber. Electronics **2022,** 11, **2566.** **[CrossRef]**

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere

Machine Translated by Google

Elektronik

MDPI

Artikel

Objekterkennung und monokulare stabile Distanzschätzung für Straßenumgebungen: Eine Fusionsarchitektur mit YOLO-RedeCa und abnormalem Springen Filter ändern

Hejun Lv, Yu Du *, Yan Ma und Ying Yuan

Beijing Key Laboratory of Information Service Engineering, College of Robotics, Beijing Union University, Peking 100101, China; lyuhejun@163.com (HL); jqrmayan@buu.edu.cn (YM); yueeing@163.com (YY)
*** Korrespondenz: duyu@buu.edu.cn**

Zusammenfassung: Eine schnelle und genaue umfassende Umgebungswahrnehmung für Fahrzeuge zu ermöglichen, stellt eine große Herausforderung dar. Objekterkennung und monokulare Entfernungsschätzung sind die beiden Haupttechnologien, obwohl sie oft getrennt voneinander verwendet werden. Daher ist es notwendig, die Interaktion zwischen ihnen zu stärken und zu optimieren. Fahrzeugbewegungen oder Objektverdeckungen können zu plötzlichen Abweichungen in den Positionen oder Größen von Erkennungsboxen innerhalb der zeitlichen Daten führen, was zu Schwankung Daher schlagen wir eine Methode zur Integration eines Detektors basierend auf YOLOv5-RedeCa, eines Bot-Sort-Trackers und eines Anomaly-Jump-Change-Filters vor. Diese Kombination ermöglicht eine genauere Erkennung und Verfolgung von Objekten. Der Anomaly-Jump-Filter glättet Distanzschwankungen, die durch plötzliche Änderungen der Erkennungsboxgrößen verursacht werden. Unsere Methode erhöht die Genauigkeit bei gleichzeitiger Reduzierung des Rechenaufwands und zeigt bei mehreren Datensätzen eine hervorragende Leistung. Insbesondere beim KITTI-Datensatz bleibt die Standardabweichung der kontinuierlichen Entfernungsmessungsergebnisse konstant niedrig, insbesondere in Szenarien mit mehreren Objektverdeckungen oder -verschwinden. Diese Ergebnisse bestätigen die Wirksamkeit und Präzision unserer Methode bei der Bewältigung dual

Schlüsselwörter: automatische Fahrtechnik; Objekterkennung; monokulare Distanzmessung; Kalman-Filter

check for updates

Zitat: Lv, H.; Du, Y.; Mai.; Yuan, Y. Objekterkennung und Monokular Stabile Entfernungsschätzung für Straßen Umgebungen: Eine Fusionsarchitektur Verwendung von YOLO-RedeCa und Abnormal Jumping Change Filter. Electronics **2024,** 13, **3058. https://doi.org/10.3390/Elektronik13153058**

Wissenschaftlicher Herausgeber: Felipe Jiménez

Empfangen: 24. Juni 2024
Überarbeitet: 19. Juli 2024
Akzeptiert: 26. Juli 2024
Veröffentlicht: 2. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.
Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

Eine effiziente Wahrnehmung und ein Verständnis der Straßenumgebung sind für die autonome Fahrtechnologie von entscheidender Bedeutung [1]. Eine umfassende Wahrnehmung der Umgebung für Fahrzeuge zu erreichen, stellt jedoch eine große Herausforderung dar. Dieses umfassende Wahrnehmungssystem erfüllt mehrere Aufgaben. Es erkennt verschiedene Objekte rund um das Fahrzeug. Zusätzlich werden die Abstände zwischen diesen Objekten und dem Fahrzeug ermittelt. Daher konzentrieren wir uns auf Methoden, die eine doppelte Aufgabe erfüllen. Unsere Methoden erkennen gleichzeitig Objekte rund um das Fahrzeug und bestimmen die Abstände zu diesen Objekten. Weitere Details finden Sie in Abbildung 1.

Beim autonomen Fahren ist eine genaue Objekterkennung für die dynamische Abstandsschätzung von entscheidender Bedeutung und wirkt sich direkt auf die Sicherheit und Effizienz des Systems aus [2]. Die dynamische Abstandsschätzung kann als Teil der Wahrnehmung auf höherer Ebene die intuitive Wahrnehmung menschlicher Fahrer simulieren.

Bestehende Entfernungsmessverfahren auf Basis von Radar und Stereovision sind zwar effektiv, stehen aber vor Herausforderungen wie hohen Kosten, Systemkomplexität und schlechter Echtzeitleistung . Darüber hinaus lohnt es sich zu untersuchen, ob die aktuelle Vogelperspektiven-Wahrnehmungstechnologie (BEV) eine stärkere Unterstützung der menschlichen Seheigenschaften erfordert [3]. Bestehende monokulare Entfernungsmessverfahren sind bei plötzlichen Wertänderungen nicht robust genug , was sich direkt auf die Wahrnehmungs- und Entscheidungsqualität autonomer Fahrsysteme in dynamisch

Machine Translated by Google

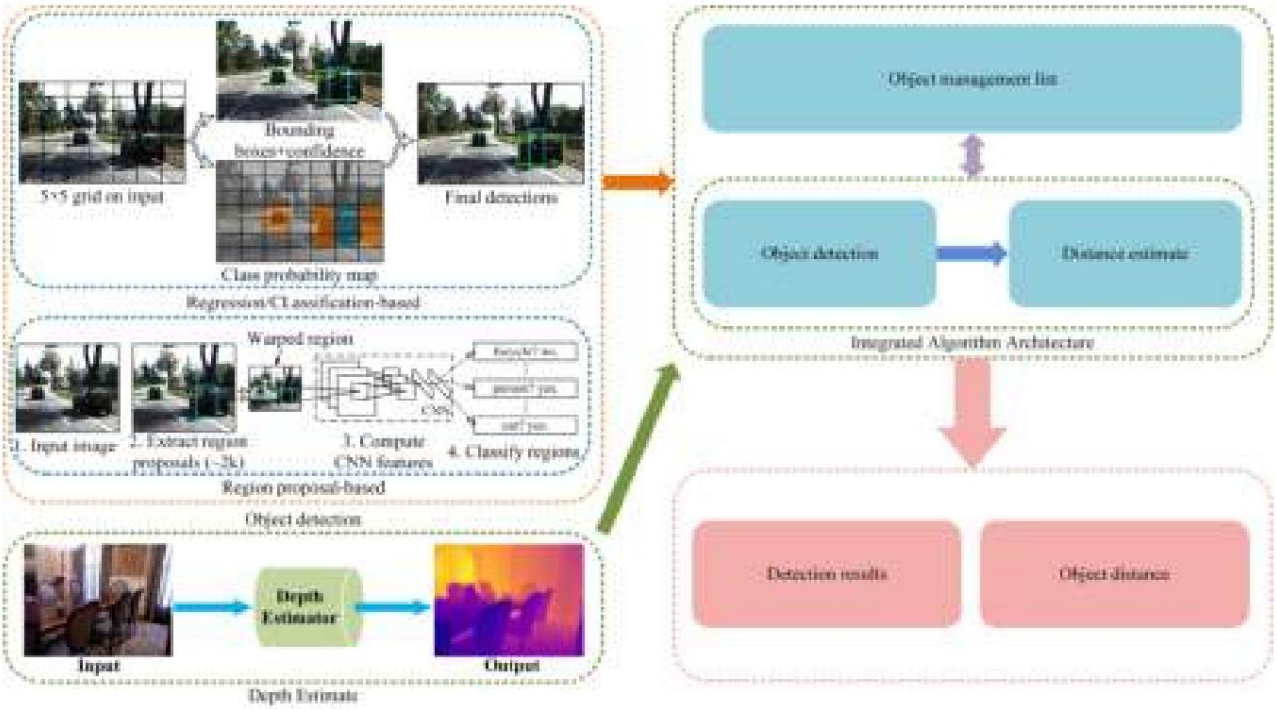


Abbildung 1. Die linke Seite zeigt das Hauptgerüst der Objekterkennung mit monokularer Tiefenschätzung. Auf der rechten Seite ist die Architektur des Algorithmus, den wir implementieren werden, und die Ausgabe.

Das Ziel dieser Studie ist es, eine monokulare Methode zur Distanzmessung vorzuschlagen. Diese Die Methode zielt darauf ab, sowohl Objekte zu erkennen als auch die Entfernung zwischen den Objekten zu bestimmen und des Fahrzeugs. Durch die Kombination der YOLOv5-RedeCa- und Anomaly-Jump-Filter-Methoden wird die Dieser Ansatz zielt darauf ab, die Genauigkeit und Robustheit der dynamischen Distanzschätzung zu verbessern, Dadurch werden Wahrnehmungs- und Entscheidungsherausforderungen in dynamischen Umgebungen bewältigt. Unsere Methode weist ein erhebliches Potenzial bei der Merkmalsextraktion und Distanzschätzung auf . Darüber hinaus wird sie zuverlässigere und umfassendere Wahrnehmungsinformationen liefern für autonome Fahrsysteme, wodurch die Fahrsicherheit und -effizienz deutlich verbessert werden . Dies macht unsere Methode zu einem vielversprechenden Kandidaten für bahnbrechende Fortschritte in Technologie für autonomes Fahren.

Überprüfung bisheriger Arbeiten im Bereich der 2D-Objekterkennung, Region-Proposal-basiert Methoden (z. B. R-CNN) [4] haben die Erkennungsgenauigkeit durch einen zweistufigen Erkennungsansatz deutlich verbessert und funktionieren auch bei komplexen Hintergründen außergew Sie sind relativ langsam in der Berechnung. Im Gegensatz dazu, obwohl einstufige Detektoren (z. B. YOLO) [5] gehen bei der Genauigkeit leichte Kompromisse ein, da sie die Positionen und Kategorien von Objekten in einem Bild direkt durch Regression/Klassifizierung vorhersagen und so das Modell vereinfachen Struktur und Berechnungsprozess, wodurch sie für die Echtzeitverarbeitung geeignet sind [6]. Daher verwenden wir sie als Basis für unseren Detektor.

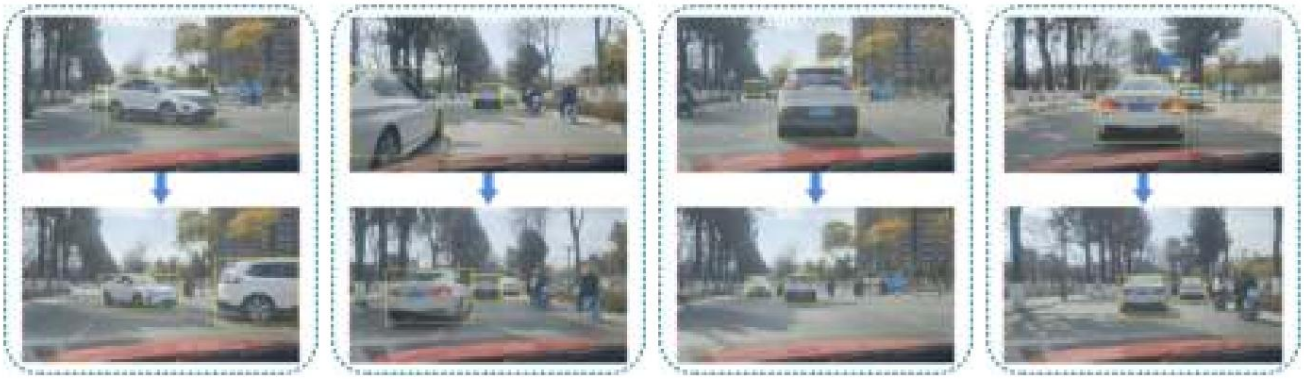
Die Objekterkennung allein kann die Umgebung ohne Entfernungsinformationen nicht vollständig erfassen [7]. Traditionelle Methoden zur Tiefenschätzung verwenden monokulare Hinweise oder manuell entwor Funktionen [8,9], aber ihre Leistung ist in Szenarien mit unterschiedlicher Tiefe oft unzureichend.

Deep Learning, insbesondere mit Convolutional Neural Networks (CNNs), bietet direkte Pixel-zu-Tiefen-Mapping [10]. Multiskalen-Netzwerke und skaleninvariante Verlustfunktionen haben verbesserte Genauigkeit bei der Tiefenschätzung. Im Gegensatz zu herkömmlichen Methoden generiert Deep Learning automatisch Disparitätskarten und Tiefeninformationen ohne umfangreiche Annotation, normalerweise mithilfe von Videosequenzen oder Stereobildern. Diese Methoden erfordern jedo Rechenressourcen.

Beim autonomen Fahren kontinuierliche Objekterkennungsinformationen aus dem Zeitbereich wird zur Objektverfolgung und Entfernungsschätzung verwendet. Allerdings können Vibrationen durch Fahrzeug Bewegungen und Okklusionen können die Stabilität von Objekten in der Videosequenz stören. Dies kann

Machine Translated by Google

zu plötzlichen Änderungen der Position oder Größe der Erkennungsboxen führen, wie in Abbildung 2 dargestellt. Daher ist es notwendig, die Erkennungsboxen zu glätten.



(A) **(B)** **(C)** **(D)**

Abbildung 2. Die Begrenzungsrahmen für nicht ideale und ideale Aufnahmen von Straßenszenenobjekten werden dargestellt. (a–d) stellen aufeinanderfolgende Frames aus vier unterschiedlichen Videosequenzen dar. In (a,b) zeigen die Begrenzungsrahmenabmessungen abrupte Änderungen, wenn das Zielobjekt plötzlich erscheint, verschwindet oder verdeckt wird. (c,d) hingegen zeigen Szenarien, in denen das Ziel klar sichtbar ist und der Begrenzungsrahmen das Zielobjekt unter normalen Bedingungen genau umschließt.

Obwohl Algorithmen zur 2D-Objekterkennung deutliche Fortschritte gemacht haben, sind sie immer noch spezialisiert. Auch Algorithmen zur monokularen Tiefenschätzung haben sich erheblich weiteren Die Forschung in diesen beiden Bereichen ist jedoch relativ getrennt. Die Integration zwischen Objekterkennung und Tiefenschätzung fehlt noch. In der bestehenden Forschung fehlt der Objekterkennung häufig eine integrierte Tiefenschätzung, und die Tiefenschätzung umfasst keine Objekterkennungsfunktionen. Diese Situation unterstreicht den dringenden Bedarf. Die Entwicklung eines integrierten Systems, das Objekterkennung und Tiefenschätzung kombiniert, ist von entscheidender Bedeutung. Ein solches System würde die Interaktion zwischen den beiden Techniken optimieren. Folglich würde es ihre Effizienz in praktischen Anwendungen verbessern.

Daher schlagen wir eine ganzheitliche Algorithmenarchitektur vor und entwerfen diese, die Objekt Erkennung und Entfernungsschätzung. Unsere wichtigsten Beiträge sind wie folgt:

1. **Wir entwickeln das RedeCa-Modul, um die Merkmalsextraktionsfunktion zu verbessern und eine präzisere Objekterkennung in Straßenverkehrsszenarien zu ermöglichen. Um eine Echtzeitleistung zu gewährleisten, verwenden wir eine tief trennbare Faltungsnetzwerkarchitektur.**
2. **Wir integrieren einen Multi-Frame-Assoziationsmechanismus, um die Robustheit des Algorithmus bei Objektverdeckung oder temporärem Verlust zu verbessern und so die gleichzeitige Verfolgung und Entfernungsbestimmung mehrerer verschiedener Objekte zu ermöglichen.**
3. **Wir entwickeln einen Anomalie-Sprungwechselfilter, um die Genauigkeit und Robustheit der Entfernungsschätzung zu verbessern. Dieser Filter eliminiert Anomalien, die durch plötzliche Änderungen der Erkennungsfeldgröße innerhalb eines Rahmens verursacht werden. Er glättet die Variation der Objektentfernung. Folglich wird die Zuverlässigkeit der Entfernungsmessung verbessert.**

Unser vorgeschlagener Algorithmus gewann die Goldmedaille bei der 3. Konferenz für intelligent vernetzte Fahrzeuge in China (Shenyang). Außerdem erhielt er die Bronzemedaille in der Gruppe Wahrnehmung und Entscheidungskontrolle bei der Metaverse Virtual Simulation Competition der World Intelligent Driving Challenge 2023.

2. Verwandte Arbeiten

Im Bereich des autonomen Fahrens ist es eine große Herausforderung, dem Fahrzeug die vollständige Wahrnehmung seiner Umgebung zu ermöglichen. Diese umfassende Wahrnehmung umfasst nicht nur die Erkennung verschiedener Objekte in der Umgebung, sondern auch die Erfassung der Entfernung zwischen diesen Objekten und dem Fahrzeug. Derzeit ist die Verwendung von Kameras mit sichtbarem Licht ein gängiger und kostengünstiger Ansatz zur Lösung dieses Problems. Ausgestattet mit lichtempfindlichen Elementen können Kameras mit sichtbarem

Machine Translated by Google

Führen Sie eine eingehende Analyse dieser Farbbilder durch, um wichtige Wahrnehmungsaufgaben wie Objekterkennung und Entfernungsschätzung zu erfüllen.

Verkehrsobjekterkennung. Die Objekterkennung ist eines der Kernprobleme im Bereich der Computer Vision [11]. Ziel ist die Klassifizierung und Bestimmung der Position von Objekten in Bil Tiefe Convolutional Neural Networks (DCNNs) haben mit ihren herausragenden Fähigkeiten zur Merkmalsdarstellung die Entwicklung von Objekterkennungstechniken stark unterstützt [12,13]. Im Allgemeinen können die wichtigsten Architekturen zur Objekterkennung in zwei Kategorien unterteilt werden: auf Region Proposals basierende Objekterkennungsmethoden und auf Regression/ Klassifizierung basierende Objekterkennungsmethoden [6]. Jede von ihnen hat ihre eigenen Merkmale und befasst sich mit unterschiedlichen technischen Herausforderungen.

Repräsentative Modelle von auf Region Proposals basierenden Objekterkennungsarchitekturen sind RCNN [4] und seine nachfolgenden Verbesserungen wie SPPNet [14], Fast RCNN [15], Faster R-CNN [16], Mask RCNN [17] und R-FCN [18]. Die erste Stufe dieser Architekturen generiert Kandidatenregionsvorschläge. Die zweite Stufe führt eine detaillierte Klassifizierung und Verfeinerung der Begrenzungsrahmen für diese Regionen durch. Diese stufenweise Verarbeitung verbessert die Erkennungsgenauigkeit erheblich und zeigt insbesondere bei komplexen Hintergründen eine hervorragende Leistung. Um beispielsweise das Problem der langsamen Vorschlagsgenerierungsgeschwindigkeit zu lösen, führt Faster R-CNN das Region Proposal Network (RPN) ei Mask RCNN prognostiziert Objektmasken noch genauer, indem es einen parallelen Zweig hinzufügt, sodass das Modell sowohl Objekterkennung als auch Instanzsegmentierung gleichzeitig durchführen kann. Dies ist besonders wertvoll bei Anwendungen, die feinkörnige visuelle Informationen erfordern.

Um die Modellstruktur und die Rechenprozesse zu vereinfachen, werden auf Regression/ Klassifizierung basierende Objekterkennungsarchitekturen entwickelt, um Objektpositionen und -kategorien direkt auf Bildern vorherzusagen, wie etwa OverFeat [19], die YOLO-Reihe [20–25], SSD [26], DSSD [27] und FSSD [28]. Der Hauptvorteil dieser Architekturen ist ihre hohe Erkennungsgeschwindigkeit, wodurch sie sich für Echtzeit-Verarbeitungsszenarien eignen. Unter ihnen sticht die YOLO-Reihe hervor . Indem sie das Erkennungsproblem als einzelne Regressionsaufgabe neu formuliert, kann sie die Kategorien und Positionen mehrerer Objekte gleichzeitig in einem einzigen Netzwerk-Vorwärtsdurchlauf vorhersagen. Um die Erkennungsprobleme für Objekte verschiedener Größen effektiv zu bewältigen und die Fähigkeit des Modells zur Erkennung kleiner

Monokulare Tiefenschätzung. Die Bestimmung der Entfernung zwischen Objekten und der Kamera ist seit ihrer Einführung ein zentraler Forschungsschwerpunkt im Bereich der Computervision . In den ersten Ansätzen stützte sich die monokulare Tiefenschätzung hauptsächlich auf Shape from Shading (SFS) und Depth from Focus/Defocus (DFF/DFD). SFS leitet die Form des Objekts anhand von Änderungen der Beleuchtung auf seiner Oberfläche ab [29]. DFF/DFD -Techniken schätzen die Tiefe anhand fokussierter und defokussierter Informationen in Bildern sowie Kamerakalibrierungsparametern [30]. Die grundlegende Annahme ist, dass die fokussierten Teile des Bildes klarer sind, während die defokussierten Teile allmählich unscharf werden. Obwohl diese Methoden in bestimmten Szenarien effektiv sind, basieren sie oft auf zusätzlichen Szenenannahmen und weisen bei Szenen mit erheblicher Tiefenschärfe eine unterdurchschnittliche Leistung auf .

Mit der Weiterentwicklung der maschinellen Lerntechnologie haben Forscher begonnen, das Problem der Tiefenschätzung einzelner Pixel als probabilistisches Lernproblem auf der Grundlage von Zufallsfeldern zu betrachten. Beispielsweise werden Tiefenkarten mithilfe von Markov Random Field (MRF) [8,31] optimiert und Posterior-Wahrscheinlichkeiten mithilfe von Conditional Random Field (CRF) zur Tiefenschätzung modelliert [9]. Die verbesserte Rechenleistung von Grafikprozessoren (GPUs) und die Verfügbarkeit großer Bilddatensätze haben das Deep Learning vorangetrieben und den Weg für neue Forschungsrichtungen in der Tiefenschätzung geebnet.

Deep-Learning-Techniken erfordern typischerweise reale Tiefendaten als Trainingslabels. Diese Techniken ordnen Pixel direkt Tiefenwerten zu, indem sie Convolutional Neural Networks (CNNs) konstruieren. Die Autoren [32,33] waren die ersten, die CNNs im Bereich der Tiefenschätzung an Sie verwendeten eine Multiskalen-Netzwerkstrategie. Ein globales Grobschätzungsnetzwerk liefert erste Entfernungsschätzungen. Zusätzlich optimiert ein lokales Verfeinerungsnetzwerk die Details. Sie führten eine skaleninvariante Verlustfunktion zur Behandlung von Fehlern ein. Nachfol

Machine Translated by Google

haben die Schätzgenauigkeit weiter verbessert, indem sie Netzwerkstrukturen verfeinert [29,34–36] und Zusatzinformationen integriert haben [37,38]. Einige Forschungsarbeiten stützen sich nicht auf eine große Menge an markierten Tiefendaten. Stattdessen verwenden sie Videosequenzen oder Stereobilder, um mithilfe von Deep-Learning-Algorithmen automatisch Disparitätskarten und Tiefeninformationen abzuleiten. Beispielsweise verarbeitet Ref. [39] die linke Ansicht mithilfe eines Convolutional Neural Network. Dadurch wird eine Disparitätskarte erzeugt. Anschließend rekonstruieren sie die linke Ansicht, indem sie diese durch inverses Mapping mit der rechten Ansicht kombinieren. Der Rekonstruktionsfehler wird als Überwachungssignal für das Training verwendet. Ref. [40] verarbeitet gleichzeitig die linke und die rechte Ansicht, um die Generierung der Disparitätskarte zu optimieren, wodurch eine genaue Tiefenschätzung ermöglicht wird, selbst wenn während des Tests nur monokulare Bilder verwendet werden. Darüber hinaus leitet Ref. [41] Pose und Tiefe aus Videosequenzen aufeinanderfolgender Frames ab und demonstriert damit die Wirksamkeit der Kombination mehrerer Netzwerkmodelle zur Ansichtsrekonstruktion. Diese Methoden erzielen zwar eine gute Leistung, sind jedoch mit einem erheblichen Rechenaufwand verbunden.

Bei der 2D-Objekterkennung und der monokularen Tiefenschätzung wurden erhebliche Fortschritte erzielt. In der Forschung werden diese beiden Verfahren jedoch weitgehend getrennt voneinander betrieben und es ist keine effektive Integration zu b

Gegenwärtig verfügen Objekterkennungssysteme häufig nicht über die Fähigkeit zur Entfernungsschätzung. Und das Tiefenschätzungssystem integriert nicht die Fähigkeit zur Objekterkennung. Diese Trennung deutet auf einen Entwicklungsbedarf hin. Eine ganzheitliche algorithmische Architektur ist erforderlich. Eine solche Architektur sollte Objekterkennung und Entfernungsschätzung integrieren. Diese Integration würde die Synergie zwischen diesen beiden Techniken verbessern. Folglich würde sie ihre Leistung in praktischen Anwendungen optimieren.

3. Methode

Um die Integration von Objekterkennung und Entfernungsschätzung zu ermöglichen, entwerfen und entwickeln wir eine algorithmische Architektur. Sie besteht aus drei Hauptkomponenten: einem Detektor basierend auf YOLOv5-RedeCa, einem Multi-Objekt-Tracker und einem Anomaly Jumping Change Filter. In autonomen Fahrszenarien treffen Fahrzeuge oft auf komplexe und dynamische Umgebungen mit mehreren autonom bewegten Objekten, die plötzlich auftauchen, verdeckt werden oder sich überlappen können [21]. Diese Situationen stellen hohe Anforderungen an die Genauigkeit und Robustheit des Algorithmus. Darüber hinaus muss der Algorithmus aufgrund der hohen Geschwindigkeit der Fahrzeugbewegung Sensordaten verarbeiten und innerhalb von Millisekunden reagieren, um die Fahrzeugsicherheit zu gewährleisten . Daher benötigen autonome Fahrsysteme einen Wahrnehmungsalgorithmus mit hoher Genauigkeit und Robustheit.

Zu diesem Zweck haben wir das Modul RedeCa entwickelt, das die Architektur des Objekterkennungsalgorithmus verbessert und die Genauigkeit der Objektlokalisierung in Straßenszenarien verbessert. Darüber hinaus integrieren wir den Bot-Sort-Tracker. Dieser Tracker verwendet einen Multi- Frame-Assoziationsmechanismus. Er verbessert die Robustheit des Algorithmus bei Objektverdeckungen oder vorübergehendem Verlust. So kann die Methode mehrere Objekte gleichzeitig verfolgen und ihre Entfernungen schätzen. Der Ansatz bewertet auch regelmäßig die Kontinuität der Verfolgung, überwacht signifikante Änderungen des Objektzustands und überprüft die Stabilität des Verfolgungsprozesses. Um die Genauigkeit und Robustheit der Entfernungsschätzung zu verbessern, integrieren wir auch einen Anomaliesprung-Änderungsfilter. Der Filter glättet effektiv die Variation der Objektentfernung und eliminiert potenzielle Anomalien, die durch plötzliche Änderungen der Erkennungsboxgröße verursacht werden . Dadurch kann er die Zuverlässigkeit der Entfernungsmessung verbessern. Die Ges

Durch die Implementierung von YOLOv5-RedeCa verbessert unsere Methode die Genauigkeit der Objekterkennung und verfeinert die Präzision der Begrenzungsrahmen über Videobilder hinweg. Darüber hinaus berechnen wir die Abstände zwischen Objekten und der Kamera mithilfe eines ausgeklügelten physikalischen E

Um mehrere Objekte effektiv zu verfolgen und zu messen, integrieren wir den Bot-Sort-Tracker. Dieser Multi-Frame-Assoziationsmechanismus verbessert die Robustheit in Szenarien mit Verdeckungen oder vorübergehendem Verlust von Objekten erheblich. Selbst wenn das Objekt in bestimmten Frames vorübergehend aus dem Blickfeld verschwindet, verfolgt und misst es weiterhin konsistent die Entfernung. Wir weisen neu erkannten Objekten mithilfe einer inkrementellen Sequenz eindeutige Kennungen zu. Anschließend verwendet es den Bot-Sort, um diese mit vorhandenen Objekten abzugleichen und zwischen denen zu unterscheiden, die

Machine Translated by Google

neu aufgetauchte und bereits verfolgte Objekte. Bei erfolgreich zugeordneten Objekten werden deren Statusinformationen aktualisiert und die ursprünglichen Kennungen bleiben erhalten. Umgekehrt werden neuen Objekten neue Tracking-Datensätze und Kennungen zugewiesen. Wenn ein Objekt im aktuellen Frame nicht erkannt wird, bleibt sein Eintrag vorübergehend erhalten. Wenn das Objekt innerhalb einer vorgegebenen Zeitspanne nicht erneut auftaucht , wird es aus der Tracking-Liste

Heutzutage konzentrieren sich die meisten monokularen Entfernungsmessungsmethoden in erster Linie auf die Genauigkeit einzelner Bilder und übersehen dabei die Korrelation von Entfernungsänderungen in aufeinanderfolgenden Bildern im zeitlichen Bereich. Daher schlagen wir eine Strategie zur Datenanalyse mit mehreren Bildern vor, die Kalman-Filter integriert. Sie verbessert nicht nur die Präzision und Robustheit der Entfernungsschätzung , sondern unterstützt auch die detaillierte Analyse der Bewegungsmuster des Objekts und der Trends bei Entfernungsänderungen. Der Kalman-Filter glättet Entfernungsschwankungen, indem er Änderungen desselben Objekts in aufeinanderfolgenden Bildern analysiert. Außerdem werden potenzielle Fehler eliminiert, die durch abrupte Änderungen der Größe von

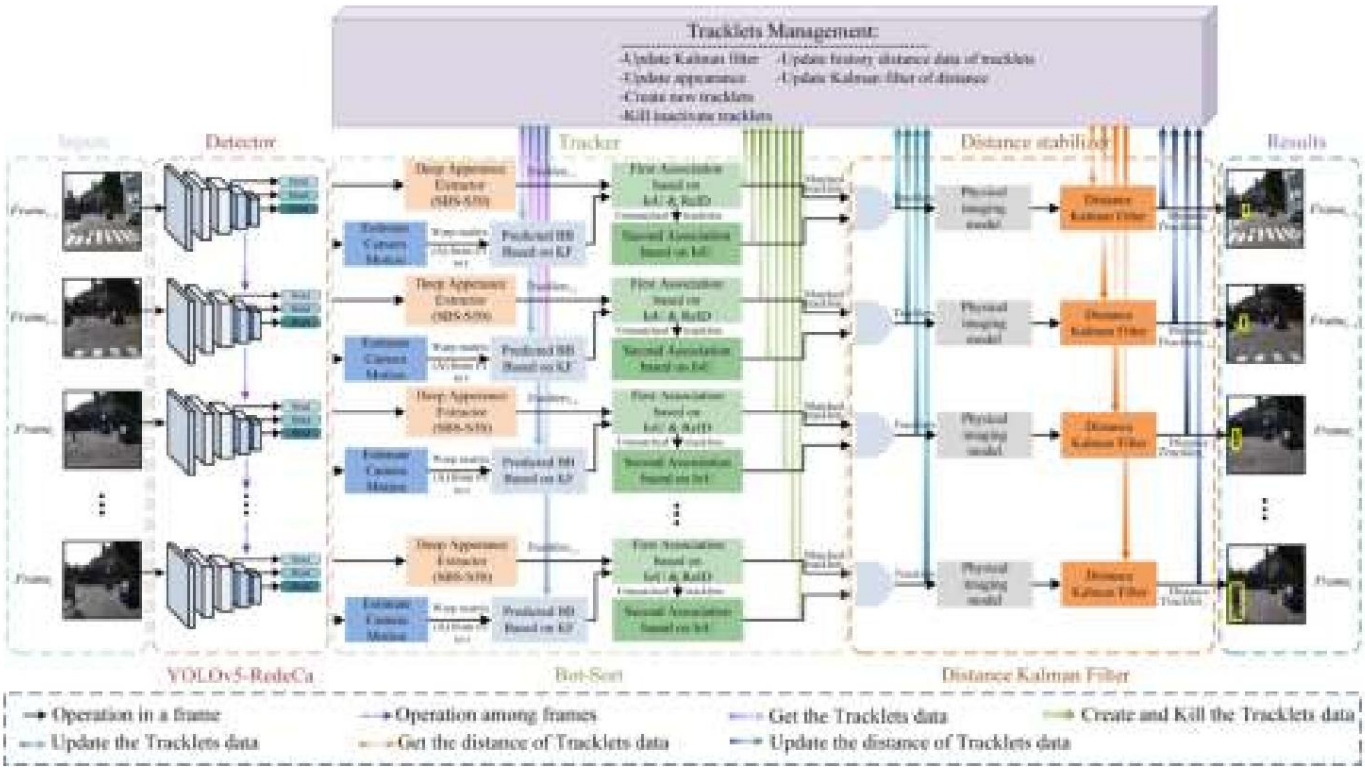


Abbildung 3. Die Grafik der Gesamtarchitektur.

3.1. Objekterkennung

Um eine genaue Objekterkennung und präzise Extraktion von Begrenzungsrahmen in jedem Videobild sicherzustellen, haben wir den Objekterkennungsalgorithmus YOLOv5 verwendet. Er dient als Kerngerüst für unser Erkennungsmodul, das auf tiefen Faltungsneuronalen Netzwerken basiert . YOLOv5 sagt die Klasse und Position von Objekten, d. h. die Begrenzungsrahmen, in einem einzigen Vorwärtsdurchgang direkt aus dem Eingabebild voraus. Im Vergleich zu den zweistufigen Erkennungsmethoden , die zuerst Kandidatenbereiche extrahieren und sie dann klassifizieren, verbessert YOLOv5 die Erkennungsgeschwindigkeit und -effizienz erheblich. Es verwendet ein tiefes Faltungsnetzwerk, um Merkmale aus Bildern zu extrahieren und schrittweise zunehmend komplexere Merkmale he Es verbessert seine Fähigkeit, verschiedene Objekte zu erkennen. Durch die Integration eines Feature Pyramid Network (FPN) kombiniert YOLOv5 Feature Maps aus verschiedenen Ebenen und verbessert so seine Erkennungsfähigkeiten für eine Reihe von Objektgrößen.

Angesichts der Vielfalt der Objektgrößen in Straßenszenen, der Genauigkeit der Objekterkennung und der spezifischen Anforderungen des Straßenverkehrs haben wir die YOLOv5-Netzwerkarchitektur verbessert . Zunächst haben wir das RedeCa-Modul in die Restblöcke integriert. Dies ermöglicht eine genauere Erkennung kleinerer Objekte unter unterschiedlichen Lichtbedingungen und

Machine Translated by Google

Wetteränderungen. Dann haben wir die traditionellen Faltungen in den Restblöcken durch tiefenweise trennbare Faltungen ersetzt. Diese Änderung reduziert die Anzahl der Parameter im Netzwerk und verbessert die Echtzeitleistung. Die Architektur des YOLOv5-RedeCa- Netzwerks ist in Abbildung 4 dargestellt.

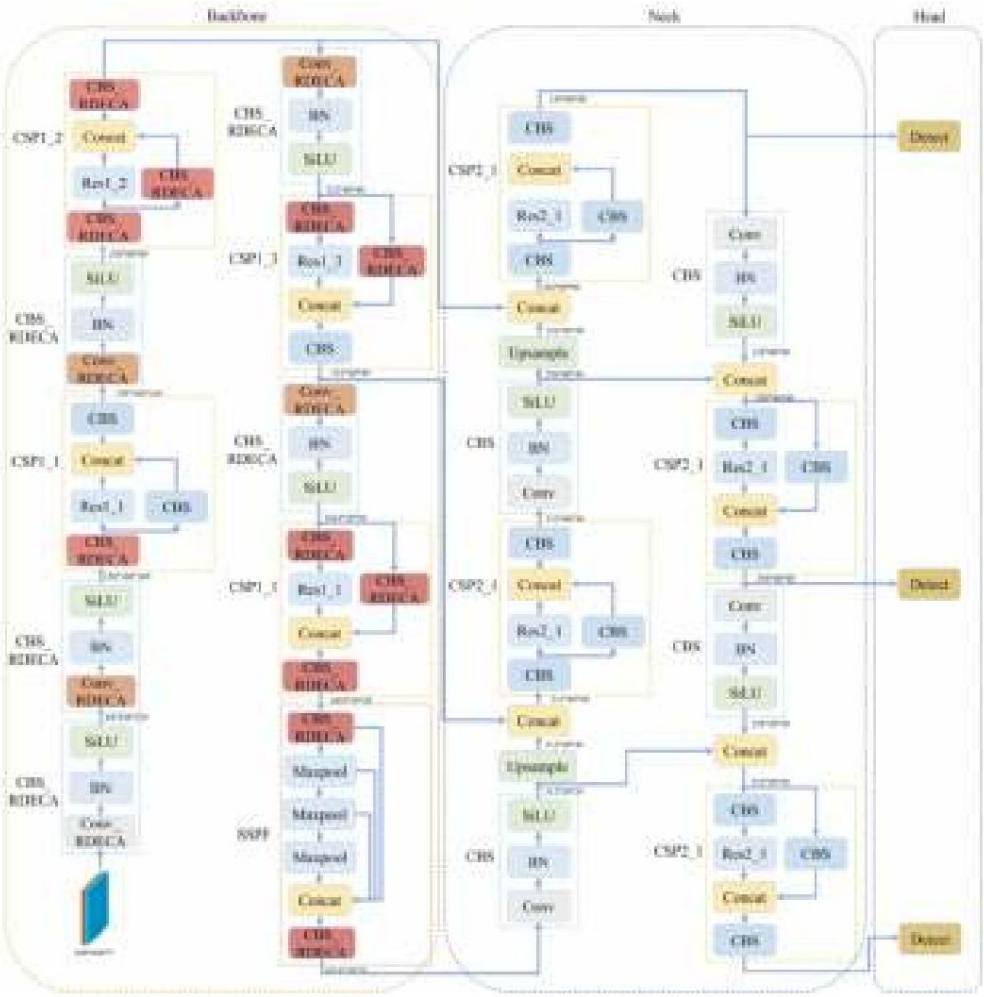


Abbildung 4. Das YOLOv5-RedeCa-Netzwerkarchitekturdiagramm.

RedeCa-Modul

YOLOv5 hat in einer Vielzahl von Erkennungsszenarien eine herausragende Leistung bewiesen . Die komplexe Straßenumgebung erfordert jedoch einen vielseitigen Ansatz, um den unterschiedlichen Arten und Größen von Verkehrsobjekten gerecht zu werden. Um diese Herausforderung zu meistern, haben wir die Netzwerkstruktur von YOLOv5 optimiert, insbesondere durch die Entwicklung des RedeCa-Moduls zur Verbesserung der Erkennungsgenauigkeit für Verkehrsobjekte verschiedener Größen. Wie in Abbildung 5 dargestellt, basiert das RedeCa-Modul auf einem Restblock. Es enthält eine Folge von residualen, tiefenweise trennbaren Faltungsblöcken und ECA

Tiefenweise trennbare Faltungen reduzieren die Anzahl der Parameter und den Rechenaufwand in Modellen deutlich. Dies wird erreicht, indem Standardfaltungen in zwei einfachere Operationen unterteilt werden: tiefenweise Faltungen und punktweise Faltungen. Der Prozess ist in Abbildung 6 detailliert dargestellt. Dabei wird zunächst die tiefenweise Faltung unabhängig auf jeden Eingangskanal angewendet. Anschließend aggregiert die punktweise Faltung diese Ausgaben mithilfe von 1 × 1 Faltungskernen. Dieser Ansatz reduziert die Parameteranzahl und den Rechenaufwand erheblich und senkt den Speicherbedarf und den Energieverbrauch. Folglich ist das Modell dadurch besser für den Einsatz a

Machine Translated by Google

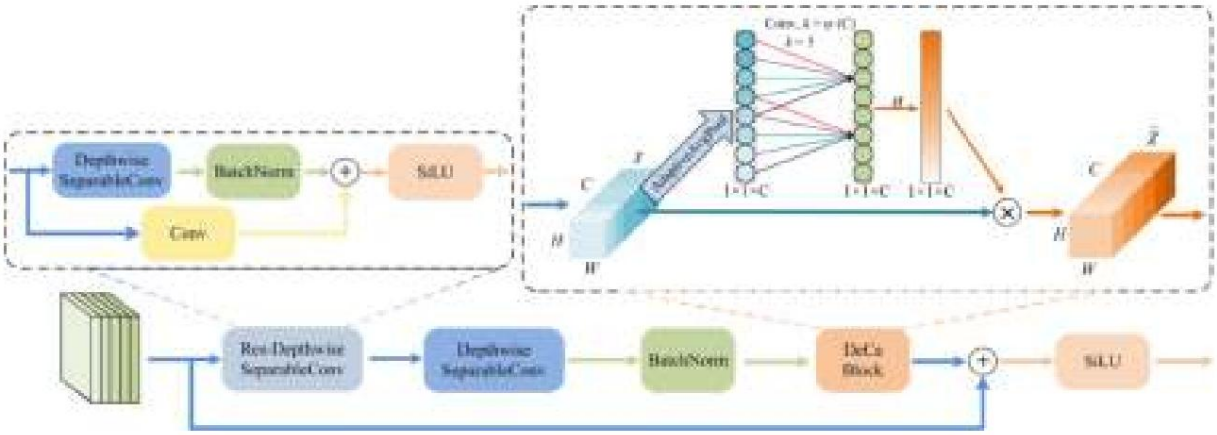


Abbildung 5. Der Graph der RedeCa-Struktur.

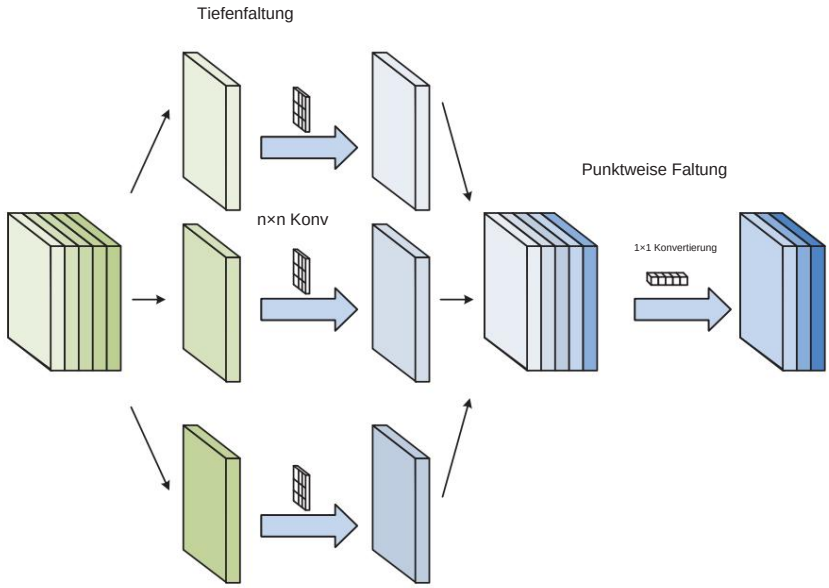


Abbildung 6. Der Graph der tiefenweise trennbaren Faltungsstruktur.

Tiefenmäßig trennbare Faltungen minimieren das Risiko einer Überanpassung, was insbesondere dann von entscheidender Bedeutung ist, wenn Modelle mit begrenzten Daten trainiert werden.

Durch die Reduzierung der Parameter wird das Training einfacher und die Wahrscheinlichkeit von Überanpassung wird reduziert. Es verbessert die Generalisierungsfähigkeit des Modells in der realen Welt Anwendungen. Die mit tiefenweise trennbaren Faltungen verbundene Rechenkomplexität wird weiter unten ausführlich beschrieben.

Tiefenfaltung:

$$ParamDC = K \times K \times Cin \quad (1)$$

$$FLOPSDC = Hout \times Wout \times K \times K \times Cin \quad (2)$$

Punktweise Faltung:

$$ParamPC = Cin \times Cout \quad (3)$$

$$FLOPSPC = Hout \times Wout \times Cin \times Cout \quad (4)$$

Machine Translated by Google

Tiefenmäßig trennbare Faltungen:

$$\text{ParamDW} = \text{ParamDC} + \text{ParamPC} \quad \textbf{(5)}$$

$$\text{FLOPSDW} = \text{FLOPSDC} + \text{FLOPSPC} \quad \textbf{(6)}$$

wobei ParamDC **die Parameteranzahl quantifiziert, die mit der tiefen Faltung verbunden ist, und** FLOPSDC **den Rechenaufwand beschreibt, der für die tiefen Faltung erforderlich ist . Entsprechend** gibt ParamPC **die Parameteranzahl für die punktweise Faltung an, während** FLOPSPC **deren Rechenaufwand charakterisiert. Darüber hinaus** gibt ParamDW **die Parameteranzahl bei der tiefenweise trennbaren Faltung an, während** FLOPSDW **den Rechenaufwand für diese Art der Faltung erläutert.**

Das RedeCa-Modul wird mit einem tiefenweise trennbaren Faltungsblock initiiert, der die Prinzipien eines Residualnetzwerks integriert. Es verwendet Identitätsmapping, um die Lerneffizienz tiefer Netzwerkschichten zu verbessern. Es kombiniert die hohe Parametereffizienz tiefenweise trennbarer Faltungen mit der schnellen Lernfähigkeit von Residualverbindungen und verbessert so die Merkmalsextraktionsfähigkeit des Moduls. Anschließend wird ein standardmäßiger tiefenweise trennbarer Faltungsblock integriert, um die Komplexität und den Rechenaufwand des Modells effektiv zu reduzieren . Schließlich passt das Deca-Modul die Abhängigkeiten zwischen den Kanälen an und optimiert die Merkmalsreaktionen dynamisch, um die Genauigkeit der Objekterkennung zu verbessern.
Die mathematische Formulierung des RedeCa-Moduls lautet wie folgt:

$$F_{DWConv}^{Res} = \text{SiLU}(\text{BN}(\text{DWConv}(F\ \ddot{y}\ )) + \text{Conv}(F\ \ddot{y}\ )) \quad \textbf{(7)}$$

$$\text{FDWConv} = \text{BN}(\text{Punktweise}(\text{Tiefenweise}(F_{DWConv}^{Res}))) \quad \textbf{(8)}$$

$$F_{\text{RedeCa}} = \text{SiLU}(\text{FDWConv} \times \ddot{y}(\text{Conv}(\text{AdaptAvgPool}(\text{FDWConv}))) + F\ \ddot{y}\ ) \quad \textbf{(9)}$$

wobei $F^{\ddot{y}}$ **stellt die Feature-Map-Eingabe aus der vorherigen Ebene dar.**

3.2. Verbesserung der Robustheit der Multi-Frame-Assoziation bei der Entfernungsschätzung

Bei der Objekterkennung und Entfernungsschätzung in Straßenverkehrsszenen reicht es nicht aus, nur ein einzelnes Objekt zu betrachten. In praktischen Anwendungen ist es notwendig, die Entfernung mehrerer Objekte gleichzeitig genau zu schätzen. Und diese Objekte können plötzlich auftauchen, verdeckt werden, sich überlappen oder verschwinden. Unsere vorgeschlagene Methode integriert den Bot-Sort-Tracker, der jedem Objekt in jedem Frame eine eindeutige Kennung zuweist und so eine effektive Objektidentifizierung und -verfolgung ermöglicht. Obwohl vorhandene monokulare Entfernungsmessungsalgorithmen die Genauigkeit bei einzelnen Bildern verbessert haben, übersehen sie häufig die Kohärenz von Entfernungsänderungen für dasselbe Objekt in aufeinanderfolgenden Frames. Das Ignorieren dieser Kohärenz kann zu Schwankungen bei den Entfernungsmessungen führen, insbesondere wenn sich die Begrenzungsrahmengröße der Objekterkennung innerhalb kurzer Zeit drastisch ändert. Unsere Methode verbessert die Genauigkeit und Robustheit der Entfernungsschätzung durch die Integration von Kalman-Filtertechniken über Zeitreihendaten . Der Ansatz glättet Variationen bei Objektentfernungen und eliminiert Anomalien, die durch plötzliche Änderungen der Einzelbild-Erkennungsrahmengrößen verursacht werden, was zu einer

3.2.1. Bot-Sort

Im Bereich der visuellen Analyse des Straßenverkehrs sind die bloße Erkennung und Entfernungsschätzung einzelner Objekte nicht ausreichend. Von größter Bedeutung für dieses Feld ist die Fähigkeit, die Entfernungen mehrerer Objekte innerhalb einer Szene genau zu messen und jedem Objekt eindeutige Kennungen zuzuweisen, um eine effektive Verfolgung und Unterscheidung zu ermöglichen. Unsere Methode integriert den Bot- Sort-Tracker, einen fortschrittlichen Algorithmus, der jedes Objekt in Videobildern zuverlässig identifiziert und analysiert und neue Objekte zur Verfolgung ermittelt. Die Methodik des Algorithmus 1 wird in Form von Pseudocode dargestellt [43].

Machine Translated by Google

Algorithmus 1 Frame-basierter Objekt-Tracking-System-Workflow

1: Eingabe: Videobilder
2: Ausgabe: Verfolgte Objekttrajektorien 3: für jedes Bild im Video
4: Potenzielle Objekte mit YOLOX erkennen
5: Schätzen Sie die Transformationsmatrix der Kamerabewegung aus dem vorherigen Frame
6: Extrahieren Sie Merkmale für erkannte Objekte mit dem SBS-S50-Modell
7: Vorhersage von Objektpositionen im aktuellen Frame mithilfe des Kalman-Filters
8: Erste Assoziation basierend auf fusionierten IoU- und ReID-Scores
9: Aktuelle Erkennungen mit vorhandenen Tracklets abgleichen
10: Zweite Assoziation für nicht übereinstimmende Tracklets basierend auf IoU
11: Statusinformationen der Tracklets aktualisieren
12: Neue Tracklets erstellen und neue eindeutige Kennungen zuweisen
13: Inaktive Tracklets entfernen, wenn sie innerhalb einer bestimmten Zeitspanne nicht erkannt werden
14: (Optional) **Fehlende Teile in Tracklets interpolieren**
15: Ende für
16: Rückkehr Abgeschlossene Flugbahnen für alle Objekte

BoT-SORT ist ein überlegenes Multi-Objekt-Tracking-System, das die Tracking-Genauigkeit durch die synergetische Integration von Bewegungs- und visuellen Erscheinungsdaten deutlich verbessert. Das System nutzt fortschrittliche Kalman-Filter zusammen mit Techniken zur Kamerabewegungskompensation, um die Dynamik der realen Welt gekonnt zu handhaben. Es verbindet auf einzigartige Weise Intersection over Union (IoU) mit Kosinus-Distanzmaßen aus Re-Identification (ReID) -Algorithmen, um konsistente und zuverlässige Objektzuordnungen herzustellen, insbesondere in Umgebungen mit hoher Dynamik.

3.2.2. Objektidentifikation und -pflege

Unsere Methode beinhaltet die folgende Logik zur Objektidentifizierung und -wartung. Zunächst weist die Methode jedem neu erkannten Objekt durch eine inkrementelle Sequenz eine eindeutige Kennung zu, um zwischen verschiedenen Tracking-Objekten unterscheiden zu können. Anschließend wendet die Methode den Bot-Sort-Algorithmus an. Dieser vergleicht neu entdeckte Objekte mit bereits verfolgten, um deren Status zu bestimmen.

Bei erfolgreich zugeordneten Objekten aktualisiert das System deren Statusinformationen in der Tracking-Liste, wobei ihre ursprünglichen eindeutigen Kennungen unverändert bleiben. Neu erscheinende Objekte werden als neue Tracking-Einträge erstellt und mit neuen eindeutigen Kennungen versehen. Wenn ein verfolgtes Objekt im aktuellen Frame nicht erkannt wird, möglicherweise aufgrund einer Okklusion oder weil es das Sichtfeld verlässt, behält die Methode seinen Eintrag und seine ID vorübergehend bei. Wenn das Objekt nicht innerhalb der eingestellten Zeit wieder auftaucht, wird es aus der Tracking-Liste entfernt.

3.2.3. Anomalie-Sprungfilter ändern

Zur Modellierung der Objektbewegung auf der Bildebene verwenden wir normalerweise ein diskretes Kalman- Filter, das mit einem Modell mit konstanter Geschwindigkeit ausgestattet ist. Der Zustandsvektor dieses Filters wird als Septupel definiert, x = [xc, **yc,s, a,** xÿc, **yÿc,sÿ], wobei** (xc, yc) **die 2D-Koordinaten des Objektmittelpunkts innerhalb der Bildebene angibt. Hier** bezeichnet s **den Maßstab (die Fläche) des Begrenzungsrahmens und** a **dessen Seitenverhältnis. Um die Präzision und Belastbarkeit der Entfernungsschätzung zu verbessern, die Schwankungen der Objektentfernung zu stabilisieren und Anomalien zu beheben, die möglicherweise durch abrupte Änderungen der Größe des Erkennungsrahmens innerhalb eines Frames ausgelöst werden, wird der Zustandsvektor auf ein Achtupel erweitert,** x = [xc, yc, **a, h,** xÿc, yÿc, **aÿ,** ÿh] **T . Dementsprechend ist der Zustandsvektor des Kalman-Filters** (KF) wie in Gleichung **(10) dargestellt definiert und der Messvektor ist in Gleichung (11) dargestellt.**

$$xk = [xc(k), yc(k), a(k), h(k), xÿc(k), yÿc(k), aÿ(k), ÿh(k)]T \quad \textbf{(10)}$$

$$zk = [zxc\ (k), zyc\ (k), za(k), zh\ (k)]T \quad \textbf{(11)}$$

Machine Translated by Google

Dieser Ansatz, der Kalman-Filterung in die Zeitreihenanalyse integriert, soll die Genauigkeit und Robustheit der Entfernungsschätzung deutlich verbessern. Der Kalman -Filter-Algorithmus glättet Schwankungen in Objektentfernungen methodisch, indem er eine Analyse historischer Daten durchführt. Er behebt effektiv Anomalien, die durch plötzliche Änderungen der Größe von Erkennungsboxen innerhalb einzelner Frames entstehen. Die Zustandsvorhersage des Algorithmus wird in den Gleichungen **(12) und (13) formuliert , während der Beobachtungsaktualisierungsmechanismus in den Gleichungen (14)–(16) detailliert beschrieben wird.**

$$\mathbf{X} = Fkx\hat{}\,k\ddot{y}1|k\ddot{y}1 + Bkuk\,k|k\ddot{y}1 \quad (12)$$

$$Pk|k\ddot{y}1 = \mathbf{FkPk\ddot{y}1|k\ddot{y}1F}^{T}_{k} + Qk \quad (13)$$

wobei $x\hat{}\,k|k\ddot{y}1$ **stellt den vorhergesagten Zustand dar,** Fk **ist die Zustandsübergangsmatrix,** Bk **bezeichnet das Steuereingabemodell,** uk **ist der Steuervektor,** Pk|kÿ1 **ist die vorhergesagte Kovarianz und** Qk **bezeichnet die Kovarianz des Prozessrauschens.**

$$Kk = Pk|k\ddot{y}1H^{T}_{k}\,(HkPk|k\ddot{y}1H^{T}_{k} + Rk)^{\ddot{y}1} \quad (14)$$

$$x\hat{}\,k|k = x\hat{}\,k|k\ddot{y}1 + Kk\,(zk\ \ddot{y}\ Hkx\hat{}\,k|k\ddot{y}1) \quad (15)$$

$$Pk|k = (I\ \ddot{y}\ KkHk)Pk|k\ddot{y}1 \quad (16)$$

wobei Kk **den Kalman-Gewinn bezeichnet,** Hk **die Beobachtungsmatrix darstellt,** zk **die tatsächliche Beobachtung ist,** Rk **die Beobachtungsrauschkovarianz angibt und** I **die Identitätsmatrix ist.**

4. Experimente

4.1. Datensätze

COCO-Datensätze. Der vom Microsoft-Team entwickelte Datensatz Common Objects in Context (COCO) ist ein wichtiges Gut in der Computer Vision-Forschung. Er enthält über 330.000 Bilder , die in 80 Objektkategorien ausführlich kommentiert sind. Indem der COCO-Datensatz einen vielfältigen, groß angelegten und anspruchsvollen Benchmark bietet, hat er umfangreiche Forschungsaktivitäten ermöglicht [44]. Folglich wurde die von uns vorgeschlagene Methode anhand des COCO-Datensatzes streng evaluiert, um ihre Leistung mit führenden Algorithmen auf diesem Gebiet zu vergleichen. Tabelle 1 fasst den Versuchsaufbau und die Datensatzpartitionierung zusammen und veranschaulicht unsere Methode innerhalb des COCO-Datensatzes.

Tabelle 1. Aufteilung der COCO-Datensätze für Training, Validierung und Testsätze.

Datensatzaufteilung	Anzahl der Bilder	Abgedeckte Kategorien
Trainingsset	118.287	80 Unterrichtsstunden
Validierungssatz	5000	80 Unterrichtsstunden
Test-Set	40.670	80 Unterrichtsstunden

VOC-Datensätze. Der PASCAL Visual Object Classes (VOC)-Datensatz dient als entscheidender Maßstab im Bereich der Computervision und unterstützt Aufgaben wie die Objekterkennung. Es umfasst Annotationen für 20 verschiedene Objektkategorien, einschließlich Klassenbezeichnungen und Begrenzungsrahmen. Die PASCAL VOC-Challenge-Serie spielte eine entscheidende Rolle bei der Weiterentwicklung visionsbasierter Technologien [45]. Zusätzlich evaluieren wir das verbesserte YOLOv5 auf dem VOC-Datensatz und vergleichen es mit anderen gängigen Methoden. Tabelle 2 enthält eine detaillierte Beschreibung der Partitionierung für den VOC-Datensatz.

Machine Translated by Google

Tabelle 2. Aufteilung des VOC-Datensatzes in Training und Validierungssätze.

Datensatzaufteilung	Anzahl der Bilder	Abgedeckte Kategorien
Trainingsset	11.540	20 Unterrichtsstunden
Validierungssatz	10.991	20 Unterrichtsstunden

KITTI-Datensätze. Der KITTI-Datensatz ist ein entscheidender Maßstab im Bereich der autonomen Fahrforschung. Es umfasst eine Vielzahl von realen Fahrszenarien, darunter Stadt-, ländliche und Autobahnumgebungen. Es bietet umfassende Anmerkungen für Aufgaben wie wie 3D-Objekterkennung, Tracking und Tiefenschätzung [46]. Wir verwenden den KITTI-Datensatz, um die Genauigkeit unserer vorgeschlagenen Methode zur Entfernungsschätzung zu validieren und zu demonstrieren seine Anwendbarkeit und Wirksamkeit unter realistischen Fahrbedingungen.

4.2. Einzelheiten zur Implementierung

4.2.1. Trainingsaufbau

Um die Fairness und Konsistenz des Modelltrainingseffekts zu gewährleisten, werden die Details der Hardware, Software und der experimentellen Umgebung sind in Tabelle 3 aufgeführt. **Die Parametereinstellungen sind in Tabelle 4 aufgeführt.**

Tabelle 3. Experimentierumgebung.

Umfeld	Ausführung
Betriebssystem	Ubuntu 20.04
CPU	AMD Ryzen 9 5900HX
Grafikkarte	Nvidia GeForce RTX 3080 16 G
Kompilierungsumgebung	Python 3.9.0
CUDA	12.2
Rahmenwerk für Deep Learning	Pytorch 1.13.1
Compiler	Pycharm 2022.3.1

Tabelle 4. Trainingsparameter.

Parameter	Wert
Batchgröße	16
Initialisierungsrate	0,1
Endgültige Lernrate	0,0001
Bildgröße	640 × 640
Optimierer	SGD
Arbeitskräfte	2

4.2.2. Verlustfunktion

Die im YOLOv5-RedeCa-Modell verwendete Verlustfunktion wird wie folgt berechnet:

$$L(tp, tgt) = \sum_{k=0}^{K} \ddot{y}\, [\ddot{y}_k^{\text{Gleichgewicht}} \ddot{y}_{box} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \ddot{y}_{kij}^{obj} L_{CIoU} + \ddot{y}_{obj} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \ddot{y}_{kij}^{obj} L_{obj} + \ddot{y}_{cls} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \ddot{y}_{kij}^{obj} L_{cls}] \quad (17)$$

$$L_{Schuldschein}^{CIoU} = 1 \ddot{y} IoU + \frac{\ddot{y}^2(b, b^{gt})}{C^2} + \ddot{y}\ddot{y} \quad (18)$$

wobei K, S^2, und B stellen die Ausgabe-Feature-Map, die Zelle und die Anzahl der Anker dar jede Zelle. $\ddot{y}_x$ **stellt das Gewicht des entsprechenden Elements dar.** $\ddot{y}_{box}$, $\ddot{y}_{obj}$ **und** $\ddot{y}_{cls}$ **nehmen Werte von 0,05, 0,7 bzw. 0,3 an. Wir haben die Parameterwerte von** $\ddot{y}_{box}$, $\ddot{y}_{obj}$ **und** $\ddot{y}_{cls}$ **auf dem ursprünglichen YOLOv5-Algorithmus. Unter Verwendung einer stabilen und anerkannten Baseline Konfiguration ist für eine konsistente Leistungsbewertung unerlässlich.** $\ddot{y}_k^{\text{Gleichgewicht}}$ **dient zum Ausgleich die Gewichte der Ausgabe-Feature-Maps bei jedem Maßstab, mit einem Standardwert von [4,0, 1,0, 0,4].**

Machine Translated by Google

obj **gibt an, ob die k-te Ausgabe-Feature-Map, die i-te Zelle und die j-te Ankerbox ÿ** kij **positive Stichproben sind. Wenn es sich um eine positive Stichprobe handelt, ist es 1, andernfalls ist es 0.** tp **und** tgt **sind der vorhergesagte Vektor und der Ground-Truth-Vektor.**

Die Verlustfunktion von CIoU ist wie folgt:

$$_{\text{LC Schuldschein}} = \mathbf{1} \ddot{y} \text{ IoU} + \frac{\ddot{y}^{\mathbf{2}}(\text{b, bgt})}{\mathbf{2}\,\text{c}} + \dot{y}\ddot{y} \quad \textbf{(19)}$$

wobei b **die Mittelpunkte der vorhergesagten Box und** b gt **die Ground- Truth-Box darstellt.** *ẏ* **ist der Ausgleichsfaktor zwischen dem durch das Seitenverhältnis verursachten Verlust und dem durch IoU verursachten Verlust.** *ÿ* **ist die Normalisierung des Unterschieds im Seitenverhältnis zwischen der vorhergesagten Box und der Ground-Truth-Box. Sie werden mathematisch wie folgt definiert:**

$$\ddot{y}(\text{b, b gt}) = \text{b } \ddot{y} \text{ b gt} \quad \mathbf{2} \quad \textbf{(20)}$$

$$\dot{y} = \frac{\ddot{y}}{(\mathbf{1} \ddot{y} \text{ IoU}) + \ddot{y}} \quad \textbf{(21)}$$

$$\ddot{y} = \frac{\mathbf{4(arctan}\ \underline{\text{wG}}\ \text{wP} \ddot{y} \mathbf{arctan}\ \text{hG hP}\ \underline{\ddot{y}\mathbf{2}})^{\mathbf{2}}}{} \quad \textbf{(22)}$$

Die Verlustfunktionen für Objekt und Klasse sind wie folgt definiert:

$$\text{Lobj(po, pIoU)} = \text{BCEsig ojb(po, pIoU; wobj)} \quad \textbf{(23)}$$

$$\text{Lcls(cp, cgt)} = \text{BCEsig (cpcls gt; wcls)} \quad \textbf{(24)}$$

wobei (wG, hG) **die Breite und Höhe der Objektbox und** (wP, hP) **die Breite und Höhe der vorhergesagten Box darstellen .** po **ist der Konfidenzwert des Objekts in der vorhergesagten Box.** pIoU **ist der Konfidenzwert des Objekts in der vorhergesagten Box.** BCEsig **und** BCEsig cls **stellen den binären Kreuzentropieverlust dar.** wobj **und** BCEcls$_{\text{cls}}$ **stellen die Gewichte der positiven Proben dar.**

4.3. Bewertungsindex

In den folgenden Experimenten bewerten wir die Genauigkeit der Objekterkennung anhand der Genauigkeit und Stabilität der Entfernungsschätzung anhand der folgenden Messwerte. •

Präzision: Dieser Messwert ist von entscheidender Bedeutung für die Beurteilung der Genauigkeit von Objekterkennungsm
Es quantifiziert das Verhältnis von Instanzen, die korrekt als Objekte identifiziert wurden, zu allen Fällen, die vom Modell als Objekte gekennzeichnet wurden. Die Formel ist definiert als

$$\text{Präzision} = \frac{\text{TP}}{\text{TP + FP}} \quad \textbf{(25)}$$

wobei TP **(True Positives) die Anzahl der korrekt erkannten Objekte darstellt und** FP **(False Positives) die Instanzen bezeichnet, die fälschlicherweise als Objekte klassifiziert wurden. •**

Rückruf: Der Rückruf bewertet den Anteil der echten Objekte, die das Modell erfolgreich aus der gesamten Menge der tatsächlichen Objekte erkennt. Ein höherer Rückruf bedeutet, dass das Modell in der Lage ist, die meisten echten Objekte zu identifizieren, auch wenn dadurch möglicherweise die Rate der falschen Positivwerte steigt. Die Formel für den Rückruf lautet

$$\text{Rückruf} = \frac{\text{TP}}{\text{TP + FN}} \quad \textbf{(26)}$$

wobei FN **(False Negatives) die Anzahl der Objekte angibt, die nicht erkannt wurden. • Average Precision (AP): AP berechnet den Mittelwert der Präzision über verschiedene Rückrufstufen hinweg. Er kann für jede Kategorie bestimmt werden, und die Mean Average Precision (MAP) stellt den Durchschnitt der APs über alle Kategorien hinweg dar. MAP wird häufig verwendet**

Machine Translated by Google

zur Beurteilung der umfassenden Leistung bei Aufgaben mit der Erkennung mehrerer Objektklassen.

$$mAP = \frac{1}{Klassen} \sum_{ich=1}^{Klassen} \int_{0}^{1} P(R)d(R) \quad (27)$$

- **Standardabweichung (SD): SD dient als statistisches Maß zur Quantifizierung des Ausmaßes der Abweichung oder Streuung vom Mittelwert innerhalb eines Datensatzes. Im Bereich der Zeitbereichsanalyse ist es von entscheidender Bedeutung, die Nähe von Distanzschätzungen genau zu be aus verschiedenen monokularen Entfernungsmessalgorithmen zu den wahren Werten abgeleitet, sowie zur Messung ihrer Streuung. Die Berechnung der Standardabweichung wird definiert durch die folgende Formel:**

$$\ddot{y} = \sqrt{\frac{1}{N} \sum_{ich=1}^{N} (f(xi) \ddot{y} \mu i)^2} \quad (28)$$

wobei $\ddot{y}$ **SD bezeichnet,** f(xi) **die geschätzte Distanz zum Zeitpunkt i darstellt und** µi **dem tatsächlichen Wert zu diesem Zeitpunkt entspricht. Diese Metrik ist entscheidend für die Beurteilung der Zuverlässigkeit und Genauigkeit dieser Algorithmen in realen Anwendungen.**

4.4. Vergleichsexperiment

Wir haben eine verbesserte YOLOv5-Methode entwickelt, die speziell für die Erkennung optimiert ist kleine Objekte in Straßenszenen unter verschiedenen Licht- und Wetterbedingungen. Zur Bewertung die Leistung unserer Methode in Bezug auf Merkmalsdarstellung und Echtzeitverarbeitung Fähigkeiten haben wir eine Reihe von Experimenten mit den COCO- und VOC-Datensätzen durchgeführt. Unsere Verbesserte Methoden, die je nach Maßstab als Ours-s, Ours-m und Ours-l bezeichnet werden, wurden verglichen mit mehreren hochmodernen Objekterkennungsmethoden.

Wir haben Präzision, mAP und Recall als primäre Leistungsmetriken verwendet. Zusätzlich Wir haben die Komplexität durch die Auswertung der Anzahl der Parameter (Params) und der Gleitkommaoperationen pro Sekunde (FLOPS) berücksichtigt. Wir haben eine vergleichende Analyse verschied Methoden auf den COCO- und VOC-Datensätzen, um die Robustheit der experimentellen Design und Validität der Datenanalyse. Die Ergebnisse werden im beigefügten Tabelle 5.

Tabelle 5. Erkennungsergebnisse verschiedener Algorithmen auf COCO- und VOC-Datensätzen.

Methode	COCO VOC Präzision (%) mAP@.5 (%) Rückruf (%) Präzision (%) mAP@.5 (%) Rückruf (%)						Parameter (M)	FLOPS (G)
SSD	50,7	45,1	38,8	31,4	54,7	57,3	26,3	65,7
Schnelleres R-CNN	53,1	48,0	46,0	34,1	57,5	54,7	60,4	93,6
Maske-RCNN	51,9	48,0	45,7	33,9	57,2	69,1	63,4	147,0
YOLOv5s	63,5	54,3	52,1	69,4	60,3	52,9	7,1	16.1
YOLOv5m	70,4	61,3	55,4	82,5	74,3	67,2	21,0	48,5
YOLOv5l	72,9	64,0	59,7	82,5	75,4	69,3	46,2	108,6
YOLOv6s	64,7	52,8	46,2	65,3	58,3	51,6	18,5	45.3
YOLOv7s	65,8	47,6	49,5	66,8	58,3	52,5	6,1	13.2
YOLOv8s	68,1	56,3	53,1	68,8	56,5	53,0	11,2	28,8
YOLOv9s	67,1	55,7	50,7	66,7	61,7	54,1	9,6	38,8
YOLOv10s	66,3	57,2	51,8	68,4	60,7	56,1	8,1	24.9
Unsere-s	67,2	55,0	51,3	67,0	57,5	53,5	6,7	13.2
Unsere-m	71,6	61,5	57,9	81,9	77,5	73,4	17,2	34,6
Unsere-l	73,6	64,4	60,1	82,5	80,0	76,0	34,3	70,4

Experimentelle Ergebnisse des COCO-Datensatzes zeigen, dass unser optimiertes Ours-m-Modell erreichte 71,6 % Präzision, 61,5 % mAP und 57,9 % Recall. Diese Ergebnisse sind die besten unter alle Methoden verwenden ähnliche Rechenressourcen. Im VOC-Datensatz hat das Modell auch übertraf mit 77,5 % mAP@.5 und 73,4 % Recall. Obwohl sie weniger Parameter als das äquivalente YOLOv5-Modell, Ours-s, Ours-m und Ours-l übertrafen

Machine Translated by Google

andere Methoden in Precision und mAP. Es zeigt unseren Erfolg bei der Leistungssteigerung bei gleichzeitiger Optimierung der Rechenleistung.

Im Vergleich zu SSD und Faster R-CNN zeigte Ours-m sowohl bei den COCO- als auch den VOC-Datensätzen erhebliche Vorteile. Es war besonders effektiv beim Recall für die Erkennung kleiner Objekte. Beispielsweise war der Recall von Ours-l beim COCO-Datensatz um 21,3 % höher als der von SSD und um 14,1 % höher als der von Faster R-CNN, was die Effektivität des **RedeCa-Moduls bei der Verbesserung der Recall-Fähigkeiten unterstreicht .**

Obwohl Mask-RCNN für die Instanzsegmentierung entwickelt wurde, hat das Ours-l-Modell bei Objekterkennungsaufgaben hervorragende Ergebnisse erzielt. Im VOC-Datensatz erreichte es einen mAP@.5 von 80,0 %, verglichen mit 57,2 % bei Mask-RCNN. Dies unterstreicht, dass unser verbessertes Modell eine hervorragende Erkennungsleistung liefert, selbst in Domänen, auf die Mask-RCNN spezialisiert ist.

Wir haben auch die Modellkomplexität kontrolliert und gleichzeitig eine höhere Leistung angestrebt. Im Vergleich zum ursprünglichen YOLOv5 zeigte die Ours-Reihe eine bessere Erkennung kleiner Objekte in komplexen Umgebungen mit weniger Parametern. Im COCO-Datensatz verbesserte Ours-l mAP und Recall jeweils um 0,4 % gegenüber YOLOv5l. Unsere Methoden übertreffen bestehende Methoden in wichtigen Leistungsmetriken bei gleichbleibender Rechenleistung und legen damit eine solide Grundlage für praktische Anwendungen.

Wir haben unsere Methode mit den neuesten Objekterkennungsmethoden von YOLOv6 bis YOLOv10 verglichen **[47–51]. Unsere Methode entspricht YOLOv8 und YOLOv9 in Bezug auf die Genauigkeit und übertrifft YOLOv6, v8, v9 und v10 in Bezug auf Modellgewichtung und Rechenlast. Die** in Tabelle **5 dargestellten Daten zeigen, dass unser Modell zwar eine mit den neuesten YOLO-Versionen vergleichbare Genauigkeit erreicht, aber dennoch Vorteile in Bezug auf Modellgewichtung und Rechenlast bietet. Diese Vorteile tragen zur Bereitstellung und Betriebseffizienz unserer Methode in praktischen Anwendungen bei.**

Das RedeCa-Modul verbessert die Fähigkeit unserer Methode, Merkmale darzustellen, insbesondere bei der genauen Verarbeitung von Kanten- und Detailinformationen, die für die Erkennung kleiner Objekte von entscheidender Bedeutung sind . Es verbessert auch die Anpassungsfähigkeit an unterschiedliche Licht- und Wetterbedingungen und erhöht so die Robustheit in komplexen Umgebungen.

In Straßenverkehrsszenarien ist die Echtzeitleistung entscheidend. Wir haben unsere Methode im Vergleich zu gängigen Methoden in Bezug auf FPS und Parameteranzahl bewertet, wie in Tabelle 5 aufgeführt . **Die Parameteranzahl und die Rechenkomplexität unserer Methode sind deutlich niedriger als die von Faster R-CNN und Mask-RCNN. Im Vergleich zum Basismodell YOLOv5 erreicht unsere Methode eine vergleichbare Leistung bei gleichzeitiger Reduzierung sowohl der Parameteranzahl als a**

Die experimentellen Ergebnisse zeigen, dass wir durch Hinzufügen des RedeCa-Moduls zu YOLOv5 die Genauigkeit der Erkennung kleiner Objekte deutlich verbessern und gleichzeitig hohe Rückrufraten l Diese Erkenntnisse sprechen nachdrücklich für die weitere Erforschung von Deep-Learning-Methoden zur Erkennung kleiner Objekte in komplexen Umgebungen und bestätigen die Wirksamkeit und das Potenzial unserer Verbesserungen.

Um die überlegene Genauigkeit und Robustheit unserer Methode zu bestätigen, haben wir eine umfassende quantitative und qualitative Vergleichsanalyse durchgeführt. Wir haben unsere Methode mit gängigen monokularen Distanzmessverfahren im Straßenverkehr verglichen.

Die Ergebnisse sind in Abbildung 7 dargestellt.

Wie in Abbildung 7 dargestellt, reduziert unsere monokulare Distanzmessverfahren effektiv potenzielle Fehler, die durch plötzliche Änderungen der Größe des Erkennungsbegrenzungsrahmens verursacht Dies führt zu einer geringeren Streuung bei den Entfernungsmessungen und stellt sicher, dass diese im Laufe der Zeit besser mit den tatsächlichen Entfernungen übereinstimmen.

Wir haben auch eine qualitative Analyse durchgeführt, um unsere Methode mit herkömmlichen Methoden zu vergleichen. Im Gegensatz zu Methoden, die Tiefenkarten für das gesamte Bild erstellen, um Objektentfernungen zu ermitteln, erkennt unsere Methode Objekte und misst ihre Entfernungen gleichzeitig, was die Effizienz verbessert. Die Ergebnisse der Entfernungsmessung sind in Abbildung 8 dargestellt.

Machine Translated by Google

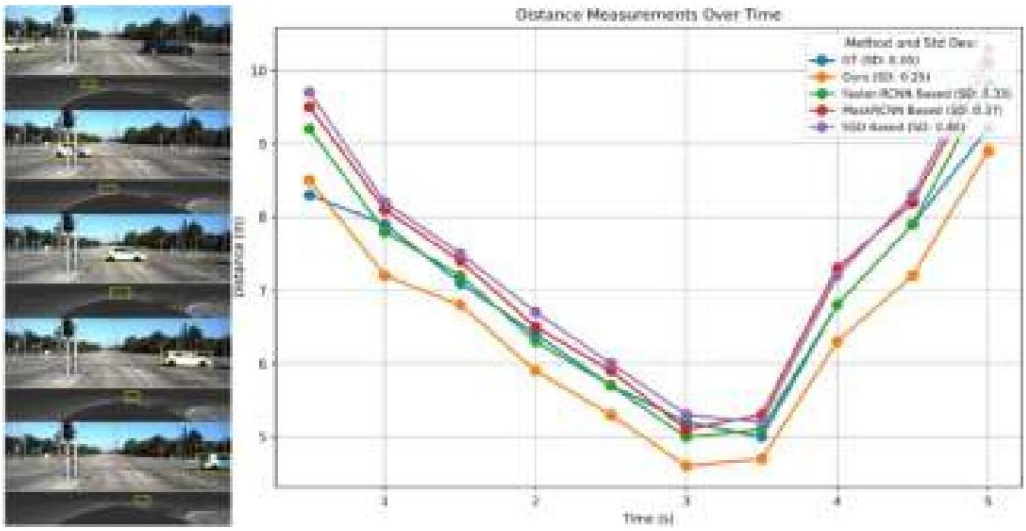


Abbildung 7. Statistische Ergebnisse und Berechnung der Standardabweichung für unterschiedliche monokulare Entfernungen Messmethoden. Die Standardabweichung unserer Methode beträgt 0,25. Die Standardabweichung von Faster-RCNN beträgt 0,33. Die Standardabweichung von Mask R-CNN beträgt 0,37. Die Standardabweichung von SSD beträgt 0,40.



(A) **(B)** **(C)** **(D)**

Abbildung 8. Die Visualisierungsergebnisse verschiedener monokularer Entfernungsmessverfahren. (a) ist die Grundwahrheit. (b) ist das Ergebnis basierend auf Faster R-CNN. (c) ist das Ergebnis basierend auf YOLOv5. (d) ist die Ergebnis basierend auf YOLO-RedeCa.

Machine Translated by Google

4.5. Ablationsexperiment

Wir haben die Methode des Ablationsexperiments angewendet, um die Auswirkungen verschiedener Module auf die Erkennungsleistung des YOLOv5m-Modells im VOC-Datensatz zu bewerten. Das Die Ergebnisse sind in Tabelle 6 dargestellt. Wir untersuchen den Effekt der tiefenweise trennbaren Faltung (DW) auf die Modellleistung. Darüber hinaus untersuchen wir die Auswirkungen effektiver Kanal Aufmerksamkeit (ECA). Wir untersuchen auch die direkte Kombination von DW und ECA. Schließlich bewerten wir die Leistung des RedeCa-Moduls.

Tabelle 6. Erkennungsergebnisse bei Verwendung verschiedener Module im VOC-Datensatz.

Methode	DW	ECA	RedeCa (DW + ECA)	Präzision (%)	mAP (%)	IOU (%)	Parameter (M)	FLOPS (G)
YOLOv5m				82,5	74,3	67,2	21,0	48,5
YOLOv5m	ÿ			70,5	62,9	38,4	9,1	18.3
YOLOv5m		ÿ		82,7	76,3	74,1	21,8	48,6
YOLOv5m	ÿ	ÿ		79,3	74,6	71,3	20,8	47,6
YOLOv5m + RedeCa			ÿ	81,9	77,5	73,4	17,2	34,6

Wie aus Tabelle 6 hervorgeht, erreichte das Basismodell YOLOv5m eine Präzision von 82,5 %, mAP von 74,3% und IOU von 67,2% auf dem VOC-Datensatz. Es hatte 21,0 Millionen Parameter und 48,5 GFLOPS.

Nach dem Hinzufügen von Depthwise Separable Convolutions (DW) zum Basismodell sank die Präzision auf 70,5 %, die mAP auf 62,9 % und die IOU deutlich auf 38,4%. Die Anzahl der Parameter und FLOPS wurde jedoch auf 9,1 Millionen reduziert und 18,3 GFLOPS. Dies zeigt, dass das DW-Modul zwar effektiv reduziert Aufgrund der Modellkomplexität wirkt sich dies negativ auf die Erkennungsleistung aus.

Die Einbindung des Efficient Channel Attention (ECA) Moduls führte zu einer leichten Verbesserung der Präzision auf 82,7 %, eine Steigerung der mAP auf 76,3 % und eine deutliche Steigerung der IOU auf 74,1%. Die Anzahl der Parameter und FLOPS stieg leicht an, was zeigt, dass das ECA-Modul die Erkennungsleistung verbessert, ohne die Rechenaufwand.

Verwendung von Depthwise Separatable Convolutions (DW) und Efficient Channel Attention (ECA) direkt, das Modell zeigte leichte Verbesserungen in der mittleren durchschnittlichen Präzision (mAP) und Intersection over Union (IoU). Der Anstieg war jedoch minimal und es gab keine signifikanten Reduzierung der Netzwerkparameter und der Rechenlast. Durch die effektive Durch die Integration von DW und ECA in das RedeCa-Modul nahm die Genauigkeit des Modells leicht ab auf 81,9 %, aber der mAP stieg auf 77,5 % und der IoU auf 73,4 %, während die Parameter und FLOPS wurden auf 17,2 M bzw. 34,6 G reduziert. Dies deutet darauf hin, dass die Verwendung des RedeCa-Moduls kann die Erkennungsleistung und die Modellkomplexität in gewissem Maße ausgleichen Umfangreich und bietet eine effektive Lösung für Anwendungen zur Echtzeit-Objekterkennung.

Um zu überprüfen, ob die Verwendung eines Anomaliesprungfilters dazu beiträgt, Distanzschwankungen auszugleichen und Beseitigung möglicher Fehler durch plötzliche Änderungen der Größe des Erkennungsbegrenzungsrahmens, haben wir eine vergleichende Analyse durchgeführt. Wir verglichen die Algorithmusarchitektur mit der Anomaliesprungfilter zu dem ohne.

Wie in Abbildung 9 dargestellt, ist die Verwendung eines Kalman-Filters, der in die Zeitreihe integriert ist, effektiv eliminiert potenzielle Fehler, die durch plötzliche Änderungen der Größe des Erkennungsrahmens verursacht werden. Dies reduziert die Streuung bei Distanzmessungen und macht sie konsistenter mit der tatsächliche Entfernungen im Laufe der Zeit.

Machine Translated by Google

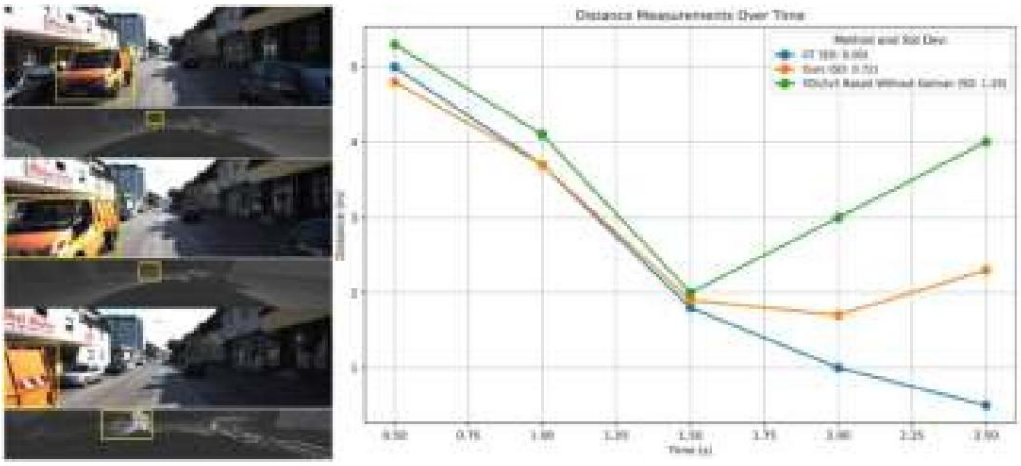


Abbildung 9. Der Einfluss der Verwendung des Kalman-Filters auf die Schwankungen bei der Entfernungsschätzung. Die Standardabweichung unserer Methode beträgt 0,72. Die Standardabweichung für die auf YOLOv5-RedeCa basierende Methode ohne Verwendung eines Kalman-Filters beträgt 1,29.

5. Schlussfolgerungen

Unser verbessertes Objekterkennungsnetzwerk liefert im COCO-Datensatz eine hervorragende Leistung. Es erreicht eine Präzision von 71,6 %, eine durchschnittliche Präzision von 61,5 % (mAP@.5) und einen Rückruf von 57,9 % . Diese Ergebnisse sind besser als die Vergleichsmodelle mit denselben Rechenressourcen . Darüber hinaus zeigt unser verbessertes Objekterkennungsnetzwerk im VOC-Datensatz auch bei mAP@.5 und dem Rückruf eine hervorragende Leistung. Die Ergebnisse liegen bei 77,5 % bzw. 73,4 % . Durch die Integration des Bot-Sort-Algorithmus wird die Robustheit der vorgeschlagenen Methode beim Umgang mit Objektverdeckungen oder temporärem Verlust deutlich verbessert, sodass das System mehrere Objekte gleichzeitig verfolgen und ihre Entfernung messen kann. Gleichzeitig kann der Filter für abnormale Sprünge plötzliche Änderungen der Objektreichweite wirksam glätten, potenzielle Anomalien durch die Veränderung der Erkennungsboxgröße beseitigen und so die Zuverlässigkeit der Entfernungsmessung verbessern.

Die von uns entwickelte und implementierte ganzheitliche Algorithmenarchitektur durchbricht erfolgreich die traditionelle Forschungslücke zwischen 2D-Objekterkennung und monokularer Tiefenschätzung. Wir kombinieren die beiden Techniken, indem wir eine neue Algorithmenarchitektur entwickeln Dadurch wird die Leistung der Objekterkennung und Entfernungsschätzung in dynamischen und komplexen Umgebungen deutlich verbessert. Die Integration dieser Technologien optimiert nicht nur die jeweiligen Funktionen, sondern verbessert auch die Effizienz des Systems in praktischen Anwendungen. Insbesondere im Bereich der Verkehrsüberwachung und des autonomen Fahrens hat sich großes Anwendungspotenzial gezeigt .

Beiträge der Autoren: Konzeptualisierung, HL, YM, YD und YY; Methodik, HL und YM; Software, HL; Validierung, HL und YY; formale Analyse, HL, YY und YM; Untersuchung, HL; Ressourcen, HL, YD und YM; Datenkuratierung, HL; Schreiben – Vorbereitung des Originalentwurfs, HL; Schreiben – Überprüfung und Bearbeitung, HL, YM und YY; Visualisierung, HL; Überwachung, YD und YM; Projektverwaltung, HL; Mittelbeschaffung, YD. Alle Autoren haben die veröffentlichte **Version des Manuskripts gelesen und stimmen ihr zu .**

Finanzierung: Diese Arbeit wird hauptsächlich durch das Vehicle-road Cooperative Autonomous Driving Fusion Control Project unterstützt und von den Academic Research Projects der Beijing Union University (Nr. ZK80202003, ZK90202105) und dem Wissenschafts- und Technologieprogramm der Bildungskommission der Stadt Peking (Nr. KM202111417007, KM202211417006) gefördert.

Datenverfügbarkeitserklärung: Der verwendete COCO-Datensatz kann unter https://cocodataset abgerufen werden. org/#Startseite (abgerufen am 24. Juni 2024). Der VOC-Datensatz ist unter **http://host verfügbar . robots.ox.ac.uk/pascal/VOC/ (abgerufen am 24. Juni 2024). Der KITTI-Datensatz ist verfügbar unter https://www.cvlibs.net/datasets/kitti/eval_object.php?obj_benchmark=3d (abgerufen am 24. Juni 2024). Die Fusionsarchitektur mit YOLO-RedeCa und abnormalem Sprungänderungsfilter ist verfügbar**

Machine Translated by Google

auf GitHub unter https://github.com/Mr-Lv-BUU/The-Fusion-Architecture-Using-YOLO-RedeCa- and-Abnormal-Jumping-Change-Filter.git (abgerufen am 24. Juni 2024).

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Verweise

1. **Tan, K.; Wu, J.; Zhou, H.; Wang, Y.; Chen, J. Integration fortschrittlicher Computer Vision- und KI-Algorithmen für autonomes Fahren Systeme.** J. Theorie Praxis. Eng. Sci. **2024, 4, 41–48.**
2. **Haris, M.; Glowacz, A. Straßenobjekterkennung: Eine vergleichende Studie von Deep-Learning-basierten Algorithmen.** Electronics **2021,** 10, **1932.** [Querverweis]
3. **Li, Y.; Huang, B.; Chen, Z.; Cui, Y.; Liang, F.; Shen, M.; Liu, F.; Xie, E.; Sheng, L.; Ouyang, W.; Shao, J. Fast-BEV: Eine schnelle und starke Basislinie zur Wahrnehmung aus der Vogelperspektive.** IEEE Trans. Pattern Anal. Mach. Intell. **2024, 1–14.**
4. **Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Umfangreiche Merkmalshierarchien für genaue Objekterkennung und semantische Segmentierung. In Proceedings der IEEE-Konferenz zu Computer Vision und Mustererkennung, Columbus, OH, USA, 23.–28. Juni 2014; S. 580–587.**
5. **Diwan, T.; Anirudh, G.; Tembhurne, JV Objekterkennung mit YOLO: Herausforderungen, Architekturnachfolger, Datensätze und Anwendungen.** Multimed. Tools Appl. **2023,** 82, **9243–9275.** [CrossRef] [PubMed]
6. **Zhao, ZQ; Zheng, P.; Xu, ST; Wu, X. Objekterkennung mit Deep Learning: Eine Übersicht.** IEEE Trans. Neural Netw. Learn. Syst. **2019,** 30, **3212–3232.** [CrossRef] [PubMed]
7. **Poggi, M.; Tosi, F.; Batsos, K.; Mordohai, P.; Mattoccia, S. Über die Synergien zwischen maschinellem Lernen und binokularem Stereo zur Tiefenschätzung aus Bildern: Eine Untersuchung.** IEEE Trans. Pattern Anal. Mach. Intell. **2021,** 44, **5314–5334.** [CrossRef] [PubMed]
8. **Saxena, A.; Sun, M.; Ng, AY Make3d: Lernen der 3D-Szenenstruktur aus einem einzelnen Standbild.** IEEE Trans. Pattern Anal. Mach. Intell. **2008,** 31, **824–840.** [CrossRef]
9. **Karsch, K.; Liu, C.; Kang, SB Tiefenübertragung: Tiefenextraktion aus Video mit nichtparametrischer Abtastung.** IEEE Trans. Pattern Anal. Mach. Intell. **2014,** 36, **2144–2158.** [CrossRef]
10. **Fu, C.; Yuan, H.; Xu, H.; Zhang, H.; Shen, L. TMSO-Net: Texturadaptive Multiskalenbeobachtung für Lichtfeldbildtiefe Schätzung.** J. Vis. Commun. Image Represent. **2023,** 90, **103731.** [CrossRef]
11. **Zou, Z.; Chen, K.; Shi, Z.; Guo, Y.; Ye, J. Objekterkennung in 20 Jahren: Eine Untersuchung.** Proc. IEEE **2023,** 111, **257–276.** [CrossRef]
12. **Krizhevsky, A.; Sutskever, I.; Hinton, GE Imagenet-Klassifizierung mit tiefen Convolutional Neural Networks.** Adv. Neural Inf. Prozess. Syst. **2012,** 25, **1097–1105.** [CrossRef]
13. **Xiao, Y.; Tian, Z.; Yu, J.; Zhang, Y.; Liu, S.; Du, S.; Lan, X. Eine Überprüfung der Objekterkennung basierend auf Deep Learning.** Multimed. Tools Appl. **2020,** 79, **23729–23791.** [CrossRef]
14. **He, K.; Zhang, X.; Ren, S.; Sun, J. Räumliches Pyramidenpooling in tiefen Faltungsnetzwerken zur visuellen Erkennung.** IEEE Trans. Pattern Anal. Mach. Intell. **2015,** 37, **1904–1916.** [CrossRef] [PubMed]
15. **Girshick, R. Fast r-cnn. In Proceedings der IEEE International Conference on Computer Vision, Santiago, Chile, 7.–13. Dezember 2015; S. 1440–1448.**
16. **Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Auf dem Weg zur Echtzeit-Objekterkennung mit Region Proposal Networks.** Adv. Neural Inf. Process. Syst. **2015,** 28, **91–99.** [CrossRef] [PubMed]
17. **Martin, D.; Fowlkes, C.; Tal, D.; Malik, J. Eine Datenbank mit von Menschen segmentierten natürlichen Bildern und ihre Anwendung zur Auswertung von Segmentierungsalgorithmen und zur Messung ökologischer Statistiken. In Proceedings der 8. IEEE International Conference on Computer Vision, ICCV 2001, Vancouver, BC, Kanada, 7.–14. Juli 2001.**
18. **Dai, J.; Li, Y.; He, K.; Sun, J. R-fcn: Objekterkennung über regionenbasierte, vollständig konvolutionelle Netzwerke.** Adv. Neural Inf. Process. Syst. **2016,** 29, **379–387.**
19. **Sermanet, P.; Eigen, D.; Zhang, X.; Mathieu, M.; Fergus, R.; LeCun, Y. Overfeat: Integrierte Erkennung, Lokalisierung und Detektion unter Verwendung von Faltungsnetzwerken.** arXiv **2013, arXiv:1312.6229.**
20. **Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. Man schaut nur einmal: Einheitliche Objekterkennung in Echtzeit. In Proceedings der IEEE-Konferenz zu Computer Vision und Mustererkennung, Las Vegas, NV, USA, 26. Juni–1. Juli 2016; S. 779–788.**
21. **Han, X.; Chang, J.; Wang, K. Echtzeit-Objekterkennung basierend auf YOLO-v2 für winzige Fahrzeugobjekte.** Procedia Comput. Sci. **2021,** 183, **61–72.** [CrossRef]
22. **Redmon, J.; Farhadi, A. YOLO9000: Besser, schneller, stärker. In Proceedings der IEEE-Konferenz über Computer Vision und Pattern Recognition, Honolulu, HI, USA, 21.–26. Juli 2017; S. 7263–7271.**
23. **Redmon, J.; Farhadi, A. Yolov3: Eine schrittweise Verbesserung.** arXiv **2018, arXiv:1804.02767.**
24. **Bochkovskiy, A.; Wang, CY; Liao, HYM Yolov4: Optimale Geschwindigkeit und Genauigkeit der Objekterkennung.** arXiv **2020, arXiv:2004.10934.**
25. **Jocher, G.; Chaurasia, A.; Stoken, A.; Borovec, J.; Kwon, Y.; Michael, K.; Fang, J.; Wong, C.; Yifu, Z.; Montes, D.; et al.** ultralytics / yolov5: v6.2 – YOLOv5-Klassifizierungsmodelle, Apple M1, Reproduzierbarkeit, ClearML- und Deci.ai-Integrationen, **v6.2; Zenodo: Genf, Schweiz, 2022.**
26. **Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, CY; Berg, AC Ssd: Single Shot Multibox Detector. In Proceedings of the Computer Vision–ECCV 2016: 14. Europäische Konferenz, Amsterdam, Niederlande, 11.–14. Oktober 2016; Proceedings, Teil I 14; Springer: Berlin/Heidelberg, Deutschland, 2016; S. 21–37.**

Machine Translated by Google

27. Fu, CY; Liu, W.; Ranga, A.; Tyagi, A.; Berg, AC Dssd: Dekonvolutioneller Einzelschussdetektor. arXiv **2017, arXiv:1701.06659.**

28. Li, Z.; Yang, L.; Zhou, F. FSSD: Feature-Fusion-Einzelschuss-Multibox-Detektor. arXiv **2017, arXiv:1712.00960.**

29. Ren, D.; Yang, M.; Wu, J.; Zheng, N. Oberflächennormal- und Gaußsche Gewichtsbeschränkungen für die Vervollständigung von Tiefenstrukturen in Innenräumen. Mustererkennung. **2023,** 138, **109362. [CrossRef]**

30. Favaro, P.; Soatto, S. Ein geometrischer Ansatz zur Formbestimmung aus Defokussierung. IEEE Trans. Pattern Anal. Mach. Intell. **2005,** 27, **406–417. [Querverweis] [PubMed]**

31. Liu, B.; Gould, S.; Koller, D. Tiefenschätzung einzelner Bilder aus vorhergesagten semantischen Labels. In Proceedings der 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13.–18. Juni 2010; S. 1253–1260.

32. Eigen, D.; Puhrsch, C.; Fergus, R. Tiefenkartenvorhersage aus einem einzelnen Bild mithilfe eines mehrskaligen tiefen Netzwerks. Adv. Neural Inf. Prozess. Syst. **2014,** 27, **2366–2374.**

33. Kuznietsov, Y.; Stuckler, J.; Leibe, B. Halbüberwachtes Deep Learning zur monokularen Tiefenkartenvorhersage. In Proceedings der IEEE-Konferenz zu Computer Vision und Mustererkennung, Honolulu, HI, USA, 21.–26. Juli 2017; S. 6647–6655.

34. Eigen, D.; Fergus, R. Vorhersage von Tiefe, Oberflächennormalen und semantischen Bezeichnungen mit einer gemeinsamen mehrskaligen Faltungsarchitektur. In Proceedings der IEEE International Conference on Computer Vision, Santiago, Chile, 7.–13. Dezember 2015; S. 2650–2658.

35. Liu, C.; Yang, G.; Zuo, W.; Zang, T. DPDFormer: Ein Grob-Fein-Modell zur monokularen Tiefenschätzung. ACM Trans. Multimed. Comput. Commun. Appl. **2024,** 20, **1–21. [CrossRef]**

36. Li, Z.; Chen, Z.; Liu, X.; Jiang, J. Depthformer: Nutzung von Fernkorrelation und lokalen Informationen für genaue monokulare Tiefenschätzung. Mach. Intell. Res. **2023,** 20, **837–854. [CrossRef]**

37. Jiao, J.; Cao, Y.; Song, Y.; Lau, R. Blicken Sie tiefer in die Tiefe: Monokulare Tiefenschätzung mit semantischem Booster und aufmerksamkeitsgesteuertem Verlust. In Proceedings der European Conference on Computer Vision (ECCV), München, Deutschland, 8.–14. September 2018; S. 53–69.

38. Miclea, VC; Nedevschi, S. Monokulare Tiefenschätzung mit verbesserter Langstreckengenauigkeit für die UAV-Umgebungswahrnehmung. IEEE Trans. Geosci. Remote Sens. **2021,** 60, **1–15. [CrossRef]**

39. Garg, R.; Bg, VK; Carneiro, G.; Reid, I. Unüberwachtes CNN zur Tiefenschätzung einzelner Ansichten: Geometrie als Rettung. In Proceedings of the Computer Vision–ECCV 2016: 14. Europäische Konferenz, Amsterdam, Niederlande, 11.–14. Oktober 2016; Proceedings, Teil VIII 14; Springer: Berlin/Heidelberg, Deutschland, 2016; S. 740–756.

40. Godard, C.; Mac Aodha, O.; Brostow, GJ Unüberwachte monokulare Tiefenschätzung mit Links-Rechts-Konsistenz. In Proceedings der IEEE-Konferenz über Computer Vision und Mustererkennung, Honolulu, HI, USA, 21.–26. Juli 2017; S. 270–279.

41. Zhou, T.; Brown, M.; Snavely, N.; Lowe, DG Unüberwachtes Lernen von Tiefe und Ego-Motion aus Video. In Proceedings der IEEE-Konferenz zu Computer Vision und Mustererkennung, Honolulu, HI, USA, 21.–26. Juli 2017; S. 1851–1858.

42. Mahaur, B.; Mishra, K. Kleinobjekterkennung basierend auf YOLOv5 in autonomen Fahrsystemen. Mustererkennung. Lett. **2023,** 168, **115–122. [CrossRef]**

43. Aharon, N.; Orfaig, R.; Bobrovsky, BZ BoT-SORT: Robuste Assoziationen bei der Verfolgung mehrerer Fußgänger. arXiv **2022, arXiv:2206.14651.**

44. Lin, TY; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollar, P.; Zitnick, CL Microsoft coco: Gemeinsame Objekte im Kontext. In Proceedings of the Computer Vision–ECCV 2014: 13. Europäische Konferenz, Zürich, Schweiz, 6.–12. September 2014; Proceedings, Teil V 13; Springer: Berlin/Heidelberg, Deutschland, 2014; S. 740–755.

45. Everingham, M.; Eslami, SA; Van Gool, L.; Williams, CK; Winn, J.; Zisserman, A. Die Herausforderung der visuellen Objektklassen von Pascal: Eine Retrospektive. Int. J. Comput. Vis. **2015,** 111, **98–136. [CrossRef]**

46. Geiger, A.; Lenz, P.; Stiller, C.; Urtasun, R. Vision meets robotics: Der Kitti-Datensatz. Int. J. Robot. Res. **2013,** 32, **1231–1237. [Querverweis]**

47. Li, C.; Kleine.; Jiang, H.; Weng, K.; Geng, Y.; Kleine.; Ke, Z.; Li, Q.; Cheng, M.; Nie, W.; et al. YOLOv6: Eine einstufige Objekterkennung Framework für industrielle Anwendungen. arXiv **2022, arXiv:2209.02976.**

48. Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-YM YOLOv7: Trainierbares Bag-of-Freebies setzt neue Maßstäbe für Echtzeit-Objektdetektoren . In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung, Vancouver, BC, Kanada, 17.–24. Juni 2023; S. 7464–7475.

49. Lou, H.; Duan, X.; Guo, J.; Liu, H.; Gu, J.; Bi, L.; Chen, H. DC-YOLOv8: Kamerabasierter Algorithmus zur Erkennung kleiner Objekte Sensor. Elektronik **2023,** 12, **2323. [CrossRef]**

50. Wang, C.-Y.; Yeh, I.-H.; Liao, H.-YM YOLOv9: Lernen, was Sie lernen möchten, mithilfe programmierbarer Gradienteninformationen. arXiv **2024, arXiv:2402.13616.**

51. Wang, A.; Chen, H.; Liu, L.; Chen, K.; Lin, Z.; Han, J.; Ding, G. YOLOv10: End-to-End-Objekterkennung in Echtzeit. arXiv **2024, arXiv:2405.14458.**

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultieren.

Machine Translated by Google

Elektronik

MDPI

Artikel

Ein Ansatz zur Klassifizierung von verschlüsseltem Datenverkehr basierend auf dem Pfad Signaturfunktionen und LSTM

Yihe Mei [1], Nurbol Luktarhan 2,*, Guodong Zhao [1] und Xiaotong Yang [2]

[1] School of Software, Universität Xinjiang, Urumqi 830091, China; 107552204789@stu.xju.edu.cn (YM); 107552204851@stu.xju.edu.cn (GZ)
[2] Hochschule für Informationswissenschaft und Ingenieurwesen, Universität Xinjiang, Urumqi 830046, China;
107552204103@stu.xju.edu.cn * Korrespondenz: nurbol@xju.edu.cn

Zusammenfassung: Die Klassifizierung von verschlüsseltem Datenverkehr ist ein entscheidender Aspekt der Netzwerksicherheit. Gängige Methoden sind jedoch mit mehreren Einschränkungen konfrontiert, wie z. B. der Abhängigkeit von Feature Engineering und der Notwendigkeit komplexer Modellarchitekturen, um eine effektive Klassifizierung zu gewährleisten. Um diese Herausforderungen zu bewältigen, schlagen wir eine Methode vor, die Pfadsignaturmerkmale mit Long Short-Term Memory (LSTM)-Modellen kombiniert, um Servicetypen innerhalb von verschlüsseltem Datenverkehr zu klassifizieren. Unser Ansatz konstruiert Verkehrspfade anhand von Paketgröße und Ankunftszeiten . Wir generieren Pfadsignaturmerkmale in verschiedenen Maßstäben mithilfe einer innovativen mehrskaligen kumulativen Merkmalsextraktionstechnik. Diese Merkmale dienen als Eingaben für LSTM-Netzwerke zur Durchführung der Klassifizierung. Insbesondere durch die Verwendung von nur 24 sequenziellen Paketmerkmalen in Verbindung mit LSTM-Modellen hat unsere Methode Die experimentellen Ergebnisse unterstreichen die Überlegenheit unserer vorgeschlagenen Methode gegenüber führenden Ansätzen auf diesem Gebiet.

Schlüsselwörter: Klassifizierung verschlüsselten Datenverkehrs; Pfadsignaturfunktionen; langes Kurzzeitgedächtnis; Deep Learning

check for updates

Zitat: Mei, Y.; Luktarhan, N.; Zhao, G.; Yang, X. Ein Ansatz zur Klassifizierung verschlüsselten Datenverkehrs basierend auf Pfadsignaturfunktionen und LSTM. Electronics **2024,** 13, **3060. https:// doi.org/ 10.3390/electronics13153060**

Wissenschaftlicher Redakteur: Martin Reisslein

Empfangen: 27. Juni 2024
Überarbeitet: 29. Juli 2024
Akzeptiert: 31. Juli 2024
Veröffentlicht: 2. August 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz.
Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/ 4.0/).

1. Einleitung

Im Zeitalter der allgegenwärtigen Internetkonnektivität hat sich die Klassifizierung des Netzwerkverkehrs zu einem Eckpfeiler der Cybersicherheit entwickelt. Die Bedeutung dieser Disziplin wird durch ihre kritischen Anwendungen bei der Erkennung von Angriffen, der Analyse von Verkehrstrends, dem Qualitätsmanagement von Diensten und der Verbesserung der Netzwerktransparenz unterstrichen **[1]. Mit der weit verbreiteten Einführung von Verschlüsselungstechnologien hat sich die Landschaft des Netzwerkverkehrs dramatisch verändert, da immer mehr Benutzer und Dienstanbieter Informatione**

Verschlüsselter Datenverkehr unterscheidet sich von unverschlüsseltem Datenverkehr dadurch, dass er Verschlüsselungsprotokolle verwendet, um die zwischen Geräten übertragenen Daten zu verschlüsseln und so den Zugriff durch Unbefugte zu verhindern . Diese Verschlüsselung bietet erhebliche Vorteile in Bezug auf Datenschutz und Sicherheit, indem sie vertrauliche Informationen vor dem Abfangen und Ausnutzen durch **böswillige Akteure schützt. Verschlüsselter Datenverkehr wahrt die Vertraulichkeit und Integrität von Benutzerdaten, Kommunikationen und Transaktionen und macht sie manipulationssicher. Im Gegensatz dazu überträgt unverschlüsselter Datenverkehr Daten im Klartext, der von jedem mit Netzwerkzugriff leicht abgefangen und gelesen werden kann. Während unverschlüsselter Datenverkehr eine einfachere Überwachung und Analyse durch Netzwerkadministratoren ermöglicht, setzt er Benutzer einem höheren Risiko von Datenlecks und Datenschutzverletzungen aus . Die Hauptvorteile von unverschlüsseltem Datenverkehr sind seine Einfachheit und der geringere Rechenaufwand , was zu einer schnelleren Datenübertragung führen kann. Die Sicherheitsvorteile von verschlüsseltem Datenverkehr überwiegen diese Vorteile jedoch bei weitem, sodass Verschlüsselung in der heut**

Obwohl verschlüsselter Datenverkehr die Privatsphäre und Anonymität der meisten Benutzer schützt, stellt er gleichzeitig erhebliche Herausforderungen für die Cybersicherheit dar [3]. Kriminelle nutzen zunehmend ausgefeilte Verschlüsselungstechniken, um illegale Aktivitäten durchzuführen,

Machine Translated by Google

Sie untergraben die Stabilität des Cyberspace und stellen eine Bedrohung für die nationale Sicherheit dar. Um diese gewaltige Herausforderung zu bewältigen, benötigen Forscher dringend neue Klassifizierungsmethoden.

Die Einführung der Verschlüsselung hat die traditionellen **Methoden zur Klassifizierung des Datenverkehrs erheblich in Frage gestellt [4]. Techniken, die sich einst auf die Überprüfung der Nutzlast und die Klassifizierung auf Port-Basis stützten, sind heute gegen den Schleier der Verschlüsselung wirkungslos [5,6].**

In der Vergangenheit stützte sich das Fachgebiet auf statistische Methoden und in jüngerer Zeit auf Algorithmen des maschinellen Lernens, um die Komplexität der Klassifizierung von verschlüsseltem Datenverkehr zu bewä Diese Methoden bringen jedoch ihre eigenen Herausforderungen mit sich. Sie erfordern oft umfangreiche Feature-Engineering-Anstrengungen und sind stark von der Qualität und Quantität der verfügbaren Daten abhängig. Darüber hinaus erfordern die komplexen Designs von Deep-Learning-Modellen zwar praktische, aber enorme Rechenressourcen und große Datensätze, um eine zufriedenstellende Genauigkeit zu erreichen **[4].**

Zu den heutigen realen Anwendungen für die Klassifizierung von verschlüsseltem Datenverkehr zählen Cisco Encrypted Traffic Analytics (ETA), Darktrace, Symantec Encrypted Traffic Management (ETM) und Vectra AI. Cisco ETA identifiziert Malware in verschlüsseltem Datenverkehr mithilfe von Netzwerktelemetrie und maschinellem Lernen. Darktrace verwendet KI-gesteuerte Anomalieerkennung, um Bedrohungen in verschlüsselten Daten zu identifizieren. Symantec ETM entschlüsselt, prüft und verschlüsselt den Datenverkehr erneut und nutzt zur Klassifizierung Deep Packet Inspection. Die Cognito-Plattform von Vectra AI verwendet maschinelles Lernen und Verhaltensanalyse, um Bedrohungen in verschlüsseltem Datenverkehr zu erkennen. Diese Lösungen unterstreichen die Abhängigkeit der Branche von maschinellem Lernen, KI und erweiterter Netzwerkanalyse für ein sicheres und ef

Die von Chen **[11] vorgeschlagene Pfadsignaturfunktion wurde in den letzten Jahren erfolgreich in Bereichen wie maschinellem Lernen, Mustererkennung und Pfadanalyse eingesetzt. Diese Funktion wurde kürzlich auch im Bereich der Klassifizierung von verschlüsseltem Datenverkehr eingeführt, wo sie ihre Durchführbarkeit anhand mehrerer Datensätze unter Beweis gestellt hat [12]. Die inhärente Komplexität der Pfadsignaturfunktionen führt jedoch dazu, dass die Klassifizierungsergebnisse nicht immer zufriedenstellend sind. Bei der Klassifizierung von verschlüsseltem Datenverkehr bleibt die Herausforderung bestehen, die Pfadsignaturfunktionen voll auszunutzen, um zufriedenstellende Ergebnisse zu erzielen, was weitere Untersuchungen und Innovationen erforderlich macht .**

In diesem Artikel stellen wir eine Methode zur Extraktion kumulativer Pfadsignaturmerkmale auf mehreren Skalen vor, um den Informationsgehalt von Pfadsignaturmerkmalen zu maximieren. Um das Potenzial dieser Merkmale voll auszuschöpfen, haben wir ihre Integration mit tiefen neuronalen Netzwerken untersucht. Diese Integration hat zu zufriedenstellenden Klassifizierungsergebnissen geführt und die Wirksamkeit der Kombination von Pfadsignaturfunktionen mit Deep Learning bei der Klassifizierung von verschlüsseltem Datenverkehr d

Die wichtigsten Beiträge dieses Papiers sind wie folgt:

1. **Wir schlagen einen Klassifizierer für verschlüsselten Datenverkehr vor, der Pfadsignaturmerkmale mit einem LSTM-Modell kombiniert. Dieser Klassifizierer extrahiert die komplexen Dynamiken und geometrischen Strukturen des verschlüsselten Datenverkehrs mithilfe von Pfadsignaturmerkmalen und verwendet ein LSTM-Modell zur Durchführung der Klassifizierungsaufgabe, wodurch die für die Verkehrsklassifizierung erforderlichen Rechenressourcen erheblich reduziert werden.**
2. **Wir führen eine mehrskalige kumulative Merkmalsextraktionsmethode ein, um Pfadsignaturmerkmale zu generieren , die optimal für LSTM-Modelle geeignet sind. Mit dieser Methode kann die Klassifizierungsgenauigkeit des LSTM-Modells um 5–30 % verbessert werden, was die Wirksamkeit unseres neuen Merkmalsextraktionsansatzes bestätigt.**
3. **Unsere vorgeschlagene Methode erzielt wettbewerbsfähige Klassifizierungsergebnisse mit nur 24 aufeinanderfolgenden Paketlängen- und Zeitintervallmerkmalen. Sie weist Klassifizierungsgenauigkeiten von 94,74 %, 90,53 %, 93,86 % und 95,03 % für die Datensätze ISCX-VPN, ISCX-nonVPN, ISCX-Tor und ISCX-nonTor auf und beweist damit die Wirksamkeit unseres Ansatzes.**

2. Verwandte Arbeiten

2.1. Klassifizierung des verschlüsselten Datenverkehrs

Die Klassifizierung des Internetverkehrs kann in drei Hauptbereiche unterteilt werden:

1. Unterscheidung zwischen verschlüsseltem und unverschlüsseltem Datenverkehr [3]; 2. Identifizierung der spezifischen Anwendungen, die mit dem Datenverkehr verbunden sind [13];

Machine Translated by Google

3. Klassifizierung des Datenverkehrs basierend auf verschiedenen Diensttypen. Diensttypen beziehen sich auf den Zweck des Datenverkehrs, d. h. auf den Datenverkehr, der generiert wird, um bestimmte Benutzeranforderungen zu erfüllen. Beispielsweise wird Dateiübertragungsverkehr generiert, wenn Benutzer Upload- oder Downloadvorgänge im Netzwerk durchführen, und Streaming-Datenverkehr wird generiert, wenn Benutzer Streaming-Mediendienste verwenden , z. B. um Musik zu hören oder Videos anzus

In diesem Abschnitt werden wichtige Methoden zur Klassifizierung des Netzwerkverkehrs zusammengefasst, kategorisiert nach den verwendeten Techniken: portbasiert, nutzlastbasiert, Fingerabdruckkonstruktion, statistische Methoden und Deep-Learning-Modelle.

Portbasiert: Diese Methode, eine der ältesten zur Klassifizierung von Netzwerkverkehr, basierte auf festen Portnummern, die von der Internet Assigned Numbers Authority (IANA) während der frühen Entwicklung des Internets zugewiesen wurden. Dieser Ansatz erleichterte die Zuordnung von Netzwerkanwendungen zu ihrem entsprechenden Datenverkehr durch Port-Scanning und ermöglichte so die Verkehrsidentifizierung [15]. Angesichts des exponentiellen Wachstums von Netzwerkanwendungen und dynamischen Porttechniken ist Port-Scanning jedoch für die aktuelle Netzwerkumgebung nicht mehr anwendbar [5]. Unsere Methode ist insbesondere nicht von Portnummern abhängig.

Payload-basiert: Um die Einschränkungen des Port-Scannings zu beheben, haben Forscher Deep Packet Inspection (DPI)-Methoden [6] zur Verkehrsidentifizierung vorgeschlagen. DPI untersucht den Datenteil (Payload) eines Pakets und verwendet bekannte Signaturen, um den Verkehrstyp zu identifizieren. Das Kernprinzip von DPI besteht darin, dass verschiedene Datenpakete von Kommunikationsprotokollen bestimmte Paketformate haben und durch Extrahieren der Payload von Paketen bestimmte Zeichen oder Zeichenfolgen mit aussagekräftigen Informationen gefunden werden können . Diese spezifischen Payload-Inhalte dienen als Indikatoren für die Verkehrstyperkennung [6]. Diese Methode birgt jedoch das Risiko, die Privatsphäre des Benutzers zu verletzen, hohe Rechenkosten zu verursachen und bei verschlüsseltem Verkehr ineffektiv zu sein **[16]. Unser Ansatz umgeht die Abhängigkeit von Payload-Informationen.**

Fingerabdruck-Erstellung: Um die Einschränkungen von nutzlastbasierten Verkehrsklassifizierungsmethoden zu beheben, haben mehrere Studien unverschlüsselte Protokollfeldinformationen zur Verkehrsklassifizierung verwendet [3]. Beispielsweise verwenden T Van Ede et al. [17] Merkmale wie Gerätetyp, Zertifikatsdetails, Paketgröße und Zeitinformationen, um jeden Datenfluss darzustellen, und erstellen durch Clustering- und Kreuzkorrelationstechniken eine Fingerabdruckbibliothek zur effizienten Verkehrsklassifizierung. Diese Methoden basieren jedoch stark auf Klartextinformationen, die während der Übertragung leicht manipuliert werden können, was zum Verlust der genauen Bedeutung führt. Im Gegensatz dazu arbeitet unser Modell unabhängig von Klartextinformationen innerhalb der Datenpake

Statistische Methoden: In dieser Kategorie werden Merkmale wie Verkehrsgröße und -zeitpunkt manuell extrahiert, wobei maschinelle Lernalgorithmen zur Verkehrsklassifizierung eingesetzt werden. Der Random Forest-Klassifikator ist ein klassischer maschineller Lernalgorithmus, der während des Trainings mehrere Entscheidungsbäume erstellt und das Klassenmuster für die Klassifizierungsaufgabe ausgibt . Diese Ensemble-Methode verbessert die Robustheit und Genauigkeit des Modells, indem sie das häufige Problem der Überanpassung in einzelnen Entscheidungsbäumen abmildert. Beispielsweise verwenden VF Taylor et al. [9] einen Random Forest-Klassifikator mit Paketgrößenmerkmalen für Klassifizierungsaufgaben. Andererseits verwendete **Vinai George Biju [18] den Support Vector Machine (SVM)-Algorithmus, um verschlüsselten Verkehr zu klassifizieren. SVM ist ebenfalls ein klassischer maschineller Lernalgorithmus, der funktioniert, indem er die optimale Hyperebene findet, die den Abstand zwischen verschiedenen Klassen im Merkmalsraum maximiert, und der in hochdimensionalen Räumen außergewöhnlich gut funktioniert. Obwohl diese Methoden die Notwendigkeit einer Überprüfung des Paketinhalts umgehen und sich daher für die Klassifizierung von verschlüsseltem Verkehr eignen, sind sie auch mit Einschränkungen verbunden. Einfache statistische Merkmale reichen oft nicht für eine genaue Klassifizierung aus, und die Ableitung komplexer Merkmale erfordert erheblichen Aufwand und Ressourcen. Unsere Methodik basiert auf einfachen statistischen Merkmalen zur Ableitung von Pfadsignaturmerkmalen und bereichert so die Aussagekraft der Merkmale o**

Deep-Learning-Modelle: Das Aufkommen des Deep Learning hat die Verwendung dieser Modelle bei der Klassifizierung von verschlüsseltem Datenverkehr erheblich beeinflusst. DF [19] verwendet zunächst ein Convolutional Neural Network (CNN), um Merkmale aus verschlüsseltem Datenverkehr zu extrahieren, während FS-NET [20] ein rekursives neuronales Netzwerk zur Klassifizierung von verschlüsseltem Datenverkehr verwendet. DeepPacket [6] und TSCRNN [21] erstellen Modelle, indem sie Merkmalsextraktion und Klassif

Machine Translated by Google

automatische Handhabung der Merkmalsextraktion und -klassifizierung, wenn die Nutzlast der Originalpakete in das Modell eingegeben wird. Zhang et al. [22] verwendeten Graph-Neural-Netzwerke, um Merkmale auf Paketebene für die Verkehrsklassifizierung zu extrahieren. MT-FlowFormer [23] verwendet einen halbüberwachten Flusstransformator, um die verschlüsselte Flussklassifizierung abzuschließen. Lin et al. [16] erfassten robuste und allgemeine Assoziationen aus Rohverkehrspaketen mithilfe großer vorab trainierter Modelle wie Bidirectional Encoder Representation from Transformers (BERT) und erzielten hervorragende Klassifizierungsergebnisse. Dennoch erfordern Deep-Learning-Modelle normalerweise komplizierte Netzwerkarchitekturen und riesige Datensätze für eine optimale Leistung. Im Gegensatz dazu ist unser Modell auf Effizienz ausgelegt und benötigt weniger Ressourcen.

2.2. Pfadsignaturmerkmale

Pfadsignaturen wurden ursprünglich von Chen [11] als Kernkonzept der Theorie der groben Pfade vorgeschlagen, wobei „grob“ sich auf kontinuierliche Pfade bezieht, die starke Schwankungen in groben Pfaden aufweisen. Diese Pfade sind kontinuierlich, aber nicht überall differenzierbar, und die kontinuierlichen Paketgrößensequenzen im verschlüsselten Datenverkehr können als grobe Pfade angesehen werden. In den letzten Jahren haben Pfadsignaturen die Aufmerksamkeit der Forscher wieder auf sich gezogen und wurden erfolgreich in den Bereichen maschinelles Lernen, Mustererkennung und Datenanalyse angewendet. Graham [24] verwendete gleitende Fenster-basierte Pfadsignaturmerkmale mit einem CNN zur Erkennung handschriftlicher Zeichen und erzielte damit hervorragende Ergebnisse. LG Gyurkó [25] verwendete Pfadsignaturen, um Merkmale aus Finanzdatenflüssen für die Klassifizierung von Finanzdaten zu extrahieren. IP Arribas [26] wandte Pfadsignaturen im medizinischen Bereich an, um psychische Störungen zu differenzieren. In [12] wurde eine neuartige Methodik zur Beschreibung von Signaturen durch längennormalisierte Pfadsignaturen eingeführt. Die Studie untersuchte auch Pfadsignaturmerkmale und setzte sie in Verbindung mit dem Random-Forest-Algorithmus zur Klassifizierung von verschlüsseltem Datenverkehr ein. Insgesamt demonstrieren diese Anwendungen die Effektivität der Merkmalsextraktion mit Pfadsignaturfunktionen und den Wert der Darstellung informationsreicher Vektoren. Der nächste Abschnitt stellt die Definition von Pfadsignaturen, ihre geometrische Interpretation niedriger Ordnung und relevante Eigenschaften vor. Weitere mathematische Beweise finden Sie in [11].

2.2.1. Definition von Pfadsignaturen Eine

Pfadsignatur ist eine transformative Abbildungsfunktion, die Rohdaten eines Pfades in eine Folge reeller Zahlen umwandelt **[12]. Jede Zahl in dieser Folge wird eindeutig durch die Datenpunkte entlang des ursprünglichen Pfades abgeleitet und erfasst eine eindeutige geometrische Eigenschaft dieses Pfades. Für einen Pfad** X **im d-dimensionalen euklidischen Raum Rd , der als kontinuierliche Abbildung vom Intervall [a, b] nach Rd** definiert ist, beschreibt der Parameter t **[a, b] eine Trajektorie innerhalb dieses d-dimensionalen euklidischen Raums:**

$$Xt = X \quad {}^{1} \quad {}^{d\,2\,3} \quad \cdot \tag{1}$$

wobei X_t **einen beliebigen kontinuierlichen Pfad in der d-ten Dimension darstellt. Die Pfadsignatur , bestehend aus der Signatur nullter ist eine unendliche Folge eines Weges X: [a, b] ÿ Ordnung (1) R und verschiedener Ordnungssignaturterme (auch iterierte Integrale genannt):**

$$S(X)_{a,b} = (1, S(X) \quad {}^{1}_{a,b}, \ldots, S(X) \quad {}^{D}_{a,b}, S(X) \quad {}^{1,1}_{a,b}, S(X) \quad {}^{1,2}_{a,b}, \ldots, S(X) \quad {}^{d,d}_{a,b} \ldots) \tag{2}$$

Für einen d-dimensionalen Weg X lässt sich sein k-faches iteriertes Integral wie folgt ausdrücken:

$$S(X)^{i_1,\ldots,i_k}_{bei} = \int_{a<t_k<t} \cdots \int_{a<t_1<t_2} dX^{i_1}_{t_1} \ldots dX^{i_k}_{t_1} \tag{3}$$

In Formel (3) beziehen sich t1, **t2, ... ,** tk **auf Integrationsvariablen, die sich von t unterscheiden. Diese Unterscheidung wird vorgenommen, um Verwechslungen mit dem Pfadparameter t bei Mehrfachintegrationen zu vermeiden. Am Beispiel des zweidimensionalen Pfades** Xt = X **in Abbildung 1 sind seine Pfadsignaturmerkmale dritter Ordnung wie folgt: {57,17, 54,67, 23,67, 30,33, 44,17, 34,33, 22,83, 20,83}.**

Machine Translated by Google

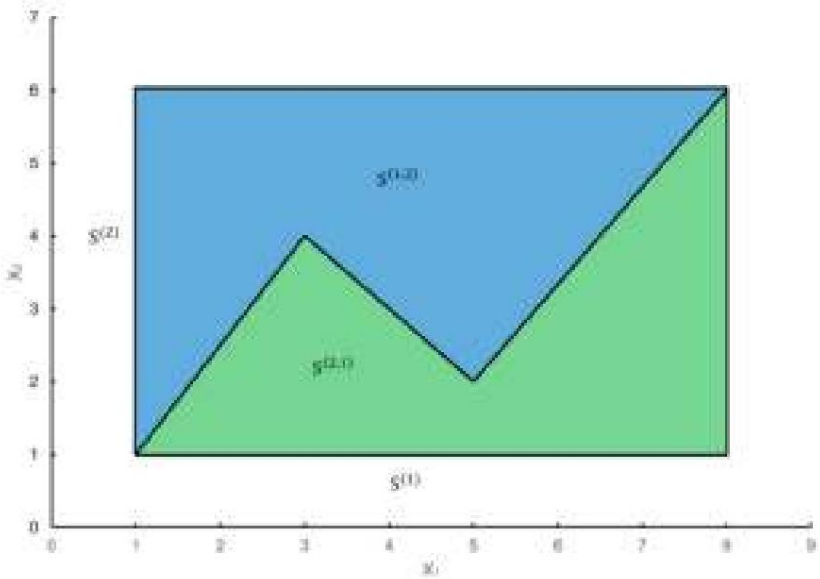


Abbildung 1. Geometrische Interpretation eines zweidimensionalen Pfades.

2.2.2. Geometrische Interpretation von

Pfadsignaturen Um die geometrische Interpretation von Pfadsignaturmerkmalen weiter zu erklären, verwenden wir **als Beispiel einen zweidimensionalen Pfad** Xt = X , wobei t **ÿ [0,4], (a = 0,b = 4). Der Pfad ist dafür gegeben durch** X = **{1, 4, 2, 6}. Es ist wichtig zu beachten, dass** t **ein kontinuierlicher Parameter ist, der Einfachheit halber haben wir vier diskrete Punkte gewählt, um den Pfad zu konstruieren. Die Form des Pfades ist in Abbildung 1 dargestellt. Die berechneten Pfadsignaturen nullter, erster und zweiter Ordnung lauten wie folgt:**

$$S = \mathbf{1, S(1), S(2), S(1,1), S(1,2), S(2,1), S(2,2)} = \mathbf{\{1, 7, 5, 24,5, 19, 16, 12,5\}. (,) (,) (,) (,)} \quad \mathbf{(4)}$$

Im Rahmen der Pfadsignatur ist das erste Element, das mit 1 bezeichnet wird, eine Konstante, die die nullte Ordnungseigenschaft des Pfades angibt. S 1) **und** S **(2)** stellen die **Pfadsignaturkomponenten erster Ordnung und ()** dar, die den Projektionsabständen des Pfades entlang **der X1- bzw. X2- Achse entsprechen. Die Ausdrücke** S **(1,2) und** S **(2,1) bezeichnen die Bereiche, die durch den Pfad und die geraden Linien parallel Koordinatenachsen** begrenzt **(,) und (,) sind** und durch die **Start- und Endpunkte des Pfades verlaufen. Unterdessen** spiegeln S **(1,1) und** S **(2,2)** jeweils **die Hälfte der quadrierten Projektionsabstände (,) und (,) der X1- bzw. X2- Achse wider. Wir erklären nur die Pfadsignaturmerkmale nullter, erster und zweiter Ordnung, da Pfadsignaturen höherer Ordnung visuell schwer zu beschreiben sind und unsere Methode nur Signaturen bis zur zweiten Ordnung verwendet. ffi**

2.2.3. Eigenschaften von Pfadsignaturen

In diesem Abschnitt werden wir auf die entscheidenden Eigenschaften von Pfadsignaturfunktionen eingehen Diese Eigenschaften bilden die theoretische Grundlage und die bemerkenswerten Vorteile ihrer Anwendung bei der Klassifizierung von verschlüsseltem Datenverkehr. Sie werden im Folgenden

detailliert beschrieben: • Eindeutigkeit: Hambly et al. [27] haben festgestellt, dass jeder grobe Pfad eine eindeutige ffi- Pfadsignatur besitzt, wodurch eine Eins-zu-eins-Entsprechung zwischen nicht baumartigen Pfaden und ihren Signaturen sichergestellt wird. Diese grundlegende Eigenschaft besagt, dass Pfadsignaturen die geometrischen Merkmale von Pfaden genau zusammenfassen können. Die Einbeziehung der Zeit als monoton zunehmende Dimension in die Sequenz des verschlüsselten Datenverkehrs wandelt diese Sequenzen in nicht baumartige Pfade um und bietet so einen robusten theoretischen Rahmen für die Ersetzung ursprünglicher Pfade durch ihre Pfadsignaturmerkmale als Eingabe für analytische Modelle. • Invarianz bei Parameteränderungen: Verschiedene Stichprobenstrategien führen zu unterschiedlichen Parametern für denselben Pfad, doch die Pfadsignatur bleibt über diese Variationen hinweg konsistent [27]. Diese Invarianz impliziert, dass Kla

Machine Translated by Google

Die spezifischen Anwendungstypen bleiben von der Parametervielfalt unberührt und stabil. Die Nutzung dieses Attributs ermöglicht die Beseitigung von Diskrepanzen, die durch verschiedene Neuparametrisierungen des Datenverkehrs innerhalb der gleichen Kategorie entstehen, und unterstreicht einen entscheidenden Vorteil der Verwendung von Pfadsignaturfunktionen.

• Dimensionsinvarianz: Die Dimensionalität der Pfadsignaturmerkmale wird ausschließlich durch die gewählte Kürzungstiefe bestimmt, unabhängig von der tatsächlichen Pfadlänge [27]. Betrachtet man **beispielsweise den zuvor diskutierten zweidimensionalen Pfad** Xt = X **1 t ÿ [0,4], (a = 0,b = 4) mit** X = **{1, 3, 5, 8} und** X = **{1, 4, 2, 6}, so führt die Kürzung auf eine Dimension zu einer Pfadsignaturlänge von 3, während die** X^1_t, X^2_t **,** **Kürzung auf zwei Dimensionen eine Länge von 7 ergibt. Diese Merkmalsextraktion mit fester Länge aus Pfaden unterschiedlicher Länge vereinfacht den Merkmalsextraktionsprozess erheblich, insbesondere für Modelle, die Eingabemerkmale mit fester Länge erfordern.**

2.3. LSTM-Modell

LSTM-Netzwerke, eine Unterklasse rekurrierender neuronaler Netzwerke (RNNs), zeichnen sich durch ihre Fähigkeit aus, langfristige Abhängigkeiten aufrechtzuerhalten [28]. Shen et al. verwendeten LSTM, um den Luftwiderstandsbeiwert von Karosserien zu analysieren [29]. Ref. [30] verwendete LSTM, um den Grundwasserspiegel vorherzusagen , und demonstrierte die Fähigkeit des Modells, schwankende Umweltdaten zu verarbeiten und genaue Prognosen zu liefern. Ref. [31] erkundete die Aussichten der Anwendung von LSTM im Bereich des autonomen Fahrens und hob sein Potenzial hervor, riesige Mengen sequenzieller Daten zu verarbeiten und daraus zu lernen . Diese vielfältigen Anwendungen unterstreichen die Vielseitigkeit von LSTM in verschiedenen Bereichen und seine robuste Fähigkeit, Zeitreihendaten zu verarbeiten. Wie in Abbildung 2 dargestellt , unterscheiden sich LSTM -Einheiten von herkömmlichen RNNs durch die Integration von drei verschiedenen „Gates“, **die gemeinsam die Speicherverwaltungsfunktionen des Modells verbessern [32]: • Eingangs-Gate: Überwacht den Fluss eingehender Daten in die Zelle und entscheidet, wie viele der neuen Informationen gespeichert werden sollen;**

• Output Gate: Bestimmt das Ausmaß, in dem der aktuelle Zustand der Zelle andere Teile des Netzwerks, die den Ausgabefluss steuern;

• Forget Gate: Passt die selbstrekurrierenden Verbindungen der Zelle mithilfe von Sigmoidfunktionen an, um Werte zwischen 0 und 1 zu skalieren und zu bestimmen, welche Informationen verworfen oder beibehalten v

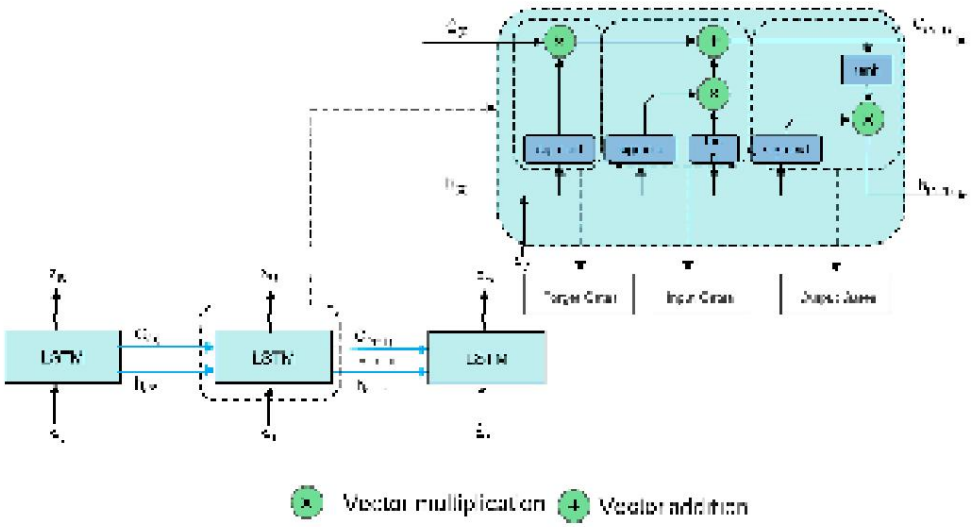


Abbildung 2. LSTM-Modellarchitektur.

Als typisches Zeitreihenmodell verarbeitet LSTM effektiv Pfadsignaturmerkmale innerhalb von Sequenzen. LSTM ist bekannt für seine Fähigkeit, langfristige Abhängigkeiten in sequenziellen Daten zu erfassen **[29–31], und eignet sich aufgrund der inhärenten sequenziellen Natur von Netzwerkverkehrsdaten besonders gut für diese Aufgabe . Im Vergleich zu anderen bekannten Deep-Learning-Modellen wie CNNs oder Transformers bietet LSTM bei dieser Aufgabe der verschlüsselten Verkehrsklassifizierung die folgenden Vorteile:**

ffi

ffi

tt

Machine Translated by Google

• Bessere Handhabung zeitlicher Abhängigkeiten: CNNs eignen sich zwar hervorragend für die Erfassung räumlicher Hierarchien und Muster, sind aber in der Regel effektiver bei der Verarbeitung von Daten mit räumlichen Korrelationen (wie Bildern) und können zeitliche Abhängigkeiten nicht so

effektiv erfassen wie LSTM [33]. • Einfachere und direktere Datenverarbeitung: Die Anwendung von CNNs auf Sequenzdaten erfordert die Umwandlung der Zeitreihen in ein für Faltungsoperationen geeignetes Format, was eine komplexere Datenverarbeitung erforderlich macht. Im Gegensatz dazu ermöglicht LSTM die direkte Eingabe von Zeitreihenmerkmalen in das Modell.

• Geringerer Bedarf an Rechenressourcen: Transformer-Modelle haben in letzter Zeit an Popularität gewonnen, da sie sich bei Sequenz-zu-Sequenz-Aufgaben, insbesondere bei der Verarbeitung natürlicher Sprache, als erfolgreich erwiesen haben. Transformer verwenden Selbstaufmerksamkeitsmechanismen, um Abhängigkeiten zwischen verschiedenen Teilen der Eingabeseque Dies ist zwar für die Erfassung globaler Abhängigkeiten von Vorteil, für das praktische Training erfordern Transformers jedoch in der Regel erhebliche Rechenressourcen und große Datensätze. LSTM-Netzwerke, die inhärente Gating-Mechanismen zur Verarbeitung sequentieller Daten verwenden, erfordern weniger Rechenressourcen für Training und Bereitstellung.

3. Methodik 3.1.

Überblick über den Ansatz

Der Prozessablauf unserer Methode zur Klassifizierung von verschlüsseltem Datenverkehr ist in Abbildung 3 dargestellt. Zunächst analysieren wir während der Datenvorverarbeitungsphase PCAP-Dateien aus öffentlich zugänglichen Datensätzen, indem wir das Fünf-Tupel-Kriterium anwenden. Anschließend verwerfen wir Pakete ohne gültige Nutzdaten. Dieser Schritt stellt die Erstellung von Datenflüssen In der Phase der Merkmalsextraktion leiten wir aus diesen einheitlichen Datenflüssen Attribute für Größe, Richtung und Ankunftszeit ab, um Verkehrspfade zu erstellen. Auf diese Verkehrspfade wird ein Transformationsprozess angewendet, um sie mit zusätzlichen Informationen anzureichern. In der nächsten Phase werden mehrskalige Akkumulationstechniken eingesetzt, um **aus diesen verbesserten Verkehrspfaden Pfadsignaturmerkmale zu extrahieren .** Nachdem wir diese Merkmale extrahiert haben, wenden wir Datenausgleichstechniken an , **um einen ausgewogenen Trainingssatz zu erhalten. Diese optimierten Pfadsignaturmerkmale werden in ein LSTM-Modell und vollständig verbundene Schichten eingegeben, um die endgültige Klassifizierungsaufgabe auszuführen . In den folgenden Unterabschnitten werden die einzelnen Phasen dieser Methodik ausführlicher behandelt .**

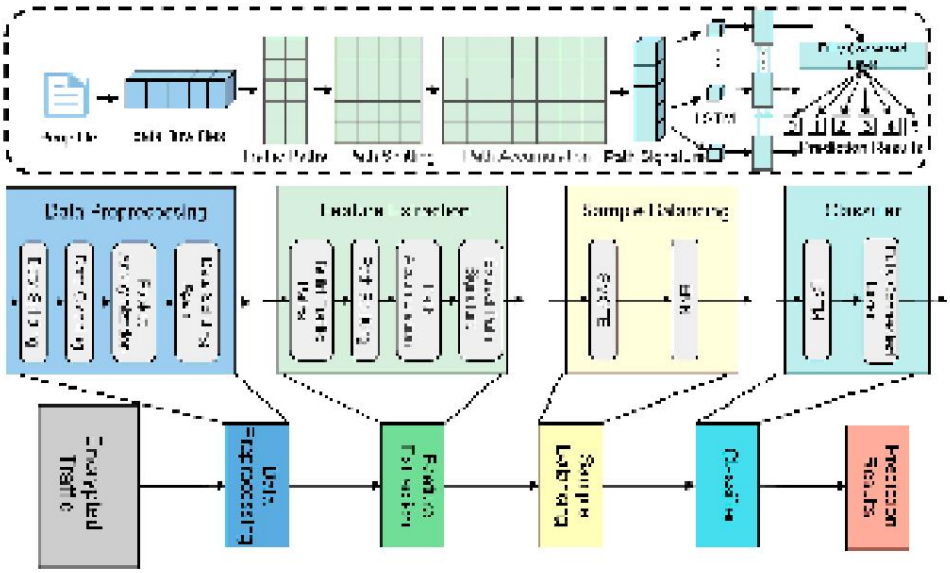


Abbildung 3. Der Gesamtrahmen unseres Ansatzes.

Machine Translated by Google

3.2. Vorverarbeitung der

Daten **Bei der Verarbeitung der Roh-PCAP-Dateien aus dem Datensatz verwenden wir SplitCap, um sie in einzelne Sitzungsdateien aufzuteilen. Dabei verwenden wir das Fünf-Tupel-Kriterium – Quell-IP, Quell-Port, Ziel-IP, Ziel-Port und das Transportschichtprotokoll [3]. In den folgenden Schritten werden Sitzungen ohne Daten und Duplikate eliminiert. Um das Risiko einer Überanpassung des Modells mit diesem Datensatz zu verringern, anonymisieren wir die MAC- und IP-Adressen in den Paketdaten, indem wir sie durch Nullen ersetzen** [6]. Anschließend verfeinern wir die Sitzungsdateien weiter, um sicherzustellen, dass jede genau 24 Datenpakete enthält – ein Parameter, dessen Bedeutung **im kommenden experimentellen Abschnitt näher erläutert wird .** Sitzungen, die das Kriterium von 24 Datenpaketen nicht erfüllen , **werden ausgeschlossen, wodurch Beispieldateien generiert werden, die unseren experimentellen Voraussetzungen entspr**

3.3. Verkehrspfade

erstellen **Nachdem wir die Sitzungsdateien auf eine einheitliche Länge standardisiert haben, kategorisieren wir Datenpakete, die vom Client zum Server gesendet werden, mit positiven Werten, während wir Pakete, die in die entgegengesetzte Richtung gesendet werden, mit negativen Werten kennzeichnen. Dieser Prozess führt zur Erstellung einer Datenpaketlängensequenz,** L = **(l1, l2, ...,** ln), **wobei** l **die Länge jedes Datenpakets darstellt [26]. Zusätzlich erfassen wir die Ankunftszeit für jedes Datenpaket. Um diese Zeitstempel zu synchronisieren, normalisieren wir die Daten, um eine Zeitsequenz für die Pakete zu erstellen, die als** T = **(t1, t2, ...,** tn) bezeichnet wird, **wobei** t **die Ankunftszeit für jedes Datenpaket da**

3.4. Verkehrspfade transformieren

Nach der Extraktion zweidimensionaler Pfadsignaturmerkmale, die die Zeit und Länge des Datenpakets kapseln, erkennen wir, dass die Darstellung der Informationen zwischen den Sitzungen unvollständig bleibt. Um die Nutzung der Merkmale zwischen den Sitzungen zu verbessern, wenden wir eine Methodik an, die von der Arbeit von Xu et al. [12] inspiriert ist. Dieser Ansatz beinhaltet die Transformation der ursprünglichen Verkehrspfade, um zusätzliche Informationsebenen freizulegen und so eine detailliertere und informativere Analyse zu ermöglichen.

3.4.1. Pfadaufteilung In

unserer Methode bezieht sich der Begriff „Client“ auf Benutzer, die Internetdienste benötigen. Die Nutzung verschiedener Dienste im Internet durch die Clients erzeugt unterschiedliche Arten von verschlüsseltem Datenverkehr. Wenn ein Benutzer beispielsweise Dateien online hoch- oder herunterlädt, entsteht Dateiübertragungsverkehr . Ebenso entsteht Streaming-Datenverkehr, wenn ein Benutzer Streaming-Dienste verwendet, um beispielsweise Musik zu hören oder Videos anzusehen. Um die Dynamik des Datenflusses zwischen Client und Server genauer darzustellen, unterteilen wir die Datenpaketlängensequenz in zwei unterschiedliche Sequenzen. Eine Sequenz erfasst Datenübertragungen vom Client zum Server, während die andere die Übertragungen vom Server zum Client detailliert beschreibt [12]. Diese Unterteilung wird im Folgenden wie folgt formalisiert.

Für Client-zu-Server-Übertragungen:

$$C = (c1, c2, \ldots, cn),\ ci = \begin{matrix} 0, & li < 0 \\ li, & li > 0 \end{matrix} \quad (5)$$

Umgekehrt gilt für Übertragungen vom Server zum Client:

$$S = (s1, s2, \ldots, sn), si = \begin{matrix} li, & li < 0 \\ 0, & li > 0 \end{matrix} \quad (6)$$

Indem wir die Datenpaketgrößen aus der entgegengesetzten Übertragungsrichtung auf Null setzen, unterstreichen wir die besonderen Flusseigenschaften, die jeder Richtung innewohnen, und legen damit den Grundstein für weitere Pfadtransformationen.

3.4.2. Pfadakkumulation

Akkumulationsmerkmale spielen eine entscheidende Rolle bei der Klassifizierung von verschlüsseltem Datenverkehr [35], vor allem weil die Nutzlastgröße bei gleicher Art von Datenverkehr tendenziel

Machine Translated by Google

Service, was zu ähnlichen Akkumulationsmustern führt. Um diese Konsistenz zu nutzen, um Informationen innerhalb verschlüsselter Verkehrsnutzlasten definieren wir die Akkumulationsfunktionen als folgt weiter unten.

Direkte Akkumulation von Paketlängen:

$$M^{\ddot{y}} = l \quad {}^{\ddot{y}}_{1}, {}^{\ddot{y}}_{2}, \cdots, \quad {}^{\ddot{y}}_{m_N}, \quad {}^{\ddot{y}}_{m_m} = \ddot{y}_{i=1}^{\ddot{y}} li, \tag{7}$$

Für clientgesteuerte Übertragungen:

$$C^{\ddot{y}} = c \quad {}^{\ddot{y}}_{1}, {}^{\ddot{y}}_{2}, {}^{\ddot{y}}_{in}, \cdots, C^{\ddot{y}} \quad C^{\ddot{y}} = \ddot{y}_{i=1}^{\ddot{y}} ci, \tag{8}$$

Für servergesteuerte Übertragungen:

$$S^{\ddot{y}} = s \quad {}^{\ddot{y}}_{1} S {}^{\ddot{y}}_{2}, \cdots, S {}^{\ddot{y}}_{N}, S^{\ddot{y}}_{m} = \ddot{y}_{i=1}^{k} ja. \tag{9}$$

Damit erreichen wir eine umfassende Darstellung des Verkehrsweges durch die folgende:

$$X^{l, c, s, \ddot{l}, C^{\ddot{y}}, S^{\ddot{y}}, T}_{t} = L, C, S, L \quad {}^{\ddot{y}}, C^{\ddot{y}}, S^{\ddot{y}}, T, \; t \ddot{y} [0, n]. \tag{10}$$

Diese Darstellung integriert die anfänglichen Datenpaketlängensequenzen und ihre akkumulierten Versionen neben der Ankunftszeitsequenz und bietet eine facettenreiche Ansicht de Verkehrsfluss für eine angemessene Klassifizierung. Unter Verwendung eines bestimmten Verkehrspfads L **= [194,ÿ83,ÿ32,53,86] und** T **= [0,2,5,8,10] als Beispiel, der entsprechende transformierte Verkehrspfad ist in dargestellt Tabelle 1.**

Tabelle 1. Beispiele für sieben verschiedene Verkehrspfade.

$X^{l, c, s, \ddot{l}, c, s^{\ddot{y}}, T}_{5}$						
M	**C**	**S**	$M^{\ddot{y}}$	$C^{\ddot{y}}$	$S^{\ddot{y}}$	**T**
194	**194**	**0**	**194**	**194**	**0**	**0**
ÿ83	**0**	**ÿ83**	**111**	**194**	**ÿ83**	**2**
ÿ32	**0**	**ÿ32**	**79**	**194**	**ÿ115**	**5**
53	**53**	**0**	**132**	**247**	**ÿ115**	**8**
86	**86**	**0**	**218**	**333**	**ÿ115**	**10**

3.5. Pfadsignatur extrahieren

Nach der Transformation der Verkehrswege besteht unser nächster Schritt darin, die Wegsignatur zu extrahieren. Wir definieren das Pfadmerkmal f **für jeden transformierten Pfad wie folgt:**

$$ft = \ddot{y}(X^{l, c, s, \ddot{l}, C^{\ddot{y}}, S^{\ddot{y}}, T}_{t}), \; t \ddot{y} [0, n]. \tag{11}$$

Dabei bezeichnet ÿ(X) **die Funktion, die für die Extraktion der Pfadsignatur zuständig ist und** ft **stellt die Pfadsignatur der ersten** t **Datenpakete dar. Es ist wichtig zu beachten, dass Die Dimensionalität der Pfadsignaturmerkmale hängt ausschließlich von der gewählten Pfadgröße ab. Kürzungstiefe, unabhängig von der tatsächlichen Länge jedes Pfades. Innerhalb einer Viele-zu-Eins-LSTM-Modell ist es üblich, die Merkmale des letzten Zeitschritts als Eingabe für die endgültige Einstufung.**

Durch die Nutzung der einzigartigen Eigenschaften der Pfadsignatur und des LSTM-Modells, In Abbildung 4 wird eine neuartige mehrskalige Akkumulations-Feature-Extraktion vorgestellt. Ansatz. Diese Methode erhöht den Informationswert, der im letzten Zeitschritt extrahiert wird Eingabe in das Modell. Die Formulierung und Struktur der Eingabemerkmalssequenz werden detailliert wie folgt:

$$F = (f2, f3, \ldots, fn). \tag{12}$$

Machine Translated by Google

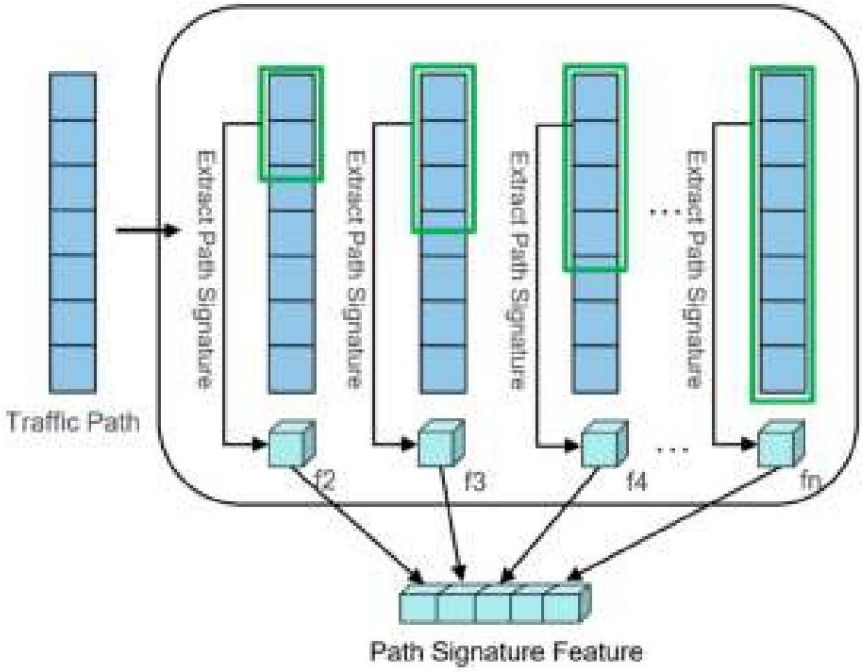


Abbildung 4. Methode zur Merkmalsextraktion mit mehrskaliger Akkumulation.

Die Anforderung, dass n ÿ **2 ist, ergibt sich aus der Überlegung, dass eine Pfadsignatur, die auf weniger als zwei Datenpaketen basiert, keine inhaltliche Relevanz hat. Wenn** n **= 1 ist, reduziert sich die Pfadsignatur darauf, lediglich die Länge und Ankunftszeit eines einzelnen Datenpakets widerzuspiegeln und erfüllt damit nicht die Kriterien für unseren Versuchsaufbau. Folglich kompilieren wir die Pfadsignatursequenz F , wobei jedes Element die Pfadsignaturmerkmale der ersten** n **Datenpakete** enthält **und alle Pfadsignaturen dieselbe Dimension haben.**

ff

3.6. Datenausgleich

Das Auffinden unausgewogener Daten über verschiedene Klassifizierungen hinweg ist ein häufiges Problem beim Sammeln experimenteller Daten aus Datensätzen. Die Implementierung von Datenausgleichstechniken ist entscheidend, um die Effektivität von Trainingsaufgaben zu steigern. Dieser Abschnitt beschreibt die Strategien, die wir zum Erreichen eines Datenausgleichs übernommen haben.

3.6.1. SMOTE (Synthetische Minderheiten-Oversampling-Technik)

SMOTE **wurde von Chawla [36] eingeführt und ist eine führende Oversampling-Methode zur Behebung von Datenungleichgewichten. Es erweitert den Datensatz, indem es neue, synthetische Instanzen für die Minderheitsklasse erstellt, die aus vorhandenen Beobachtungen abgeleitet werden. Innerhalb des** tt **Merkmalsraums erstellt SMOTE synthetische Proben, die der Minderheitsklasse ähneln. Insbesondere identifiziert SMOTE für jede Probe der Minderheitsklasse ihre n nächsten Nachbarn mithilfe der euklidischen Distanz. Anschließend wird ein Nachbar zufällig ausgewählt, wodurch ein Vektor generiert wird, der diesen Nachbarn und die Probe verbindet. Eine synthetische Instanz wird dann gebildet, indem ein zufälliger Bruchteil dieses Vektors zur ursprünglichen Probe hinzugefügt wird. Dieses Verfahren wird wiederholt, bis die Ziel-Oversampling-Rate erreicht** ist **[36]. Die Formel des SMOTE-Algorithmus**

lautet wie folgt: ff ff

$$n = s + d \ddot{y} s \ddot{y} s \quad R \quad , \ 0 \ddot{y} d \ddot{y} 1. \tag{13}$$

wobei s **die Stichproben der Minderheitsklasse symbolisiert,** s die durch k **nächste Nachbarn bestimmten Nachbarn bezeichnet ,** d **ein zufällig ausgewählter Wert im Bereich [0,1] ist und** n **die generierte synthetische Stichprobe darstellt. Wie in Abbildung 5 gezeigt, stellen zwei verschiedenfarbige Punkte zwei verschiedene Arten von Stichproben dar, die drei nächsten Nachbarn der S**

Bei der SMOTE-Methode wird für die Stichprobe s zufällig ein Nachbar s1 ausgewählt und durch Berechnung eine neue Stichprobe n ermittelt.

R

= + ÿ ( ), 0 1.

Machine Translated by Google

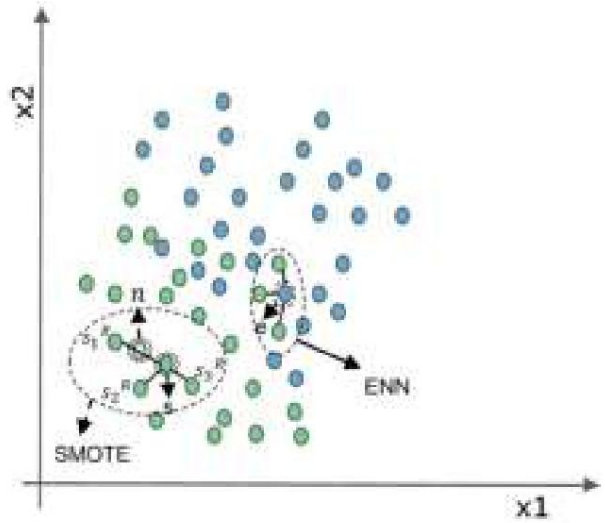


Abbildung 5. Das Beispiel von SMOTE und ENN.

3.6.2. ENN (Bearbeitete nächste Nachbarn)

ENN verwendet den k-Nearest-Neighbors-Algorithmus als sekundäre Sampling-Strategie [37]. Ziel ist es, den Datensatz zu bereinigen, indem wahrscheinlich falsch klassifizierte Instanzen entfernt werden. ENN berechnet die euklidische Distanz für jede Instanz, um ihre k nächsten Nachbarn zu bestimmen (wobei k üblicherweise auf 3 gesetzt wird). Wenn die meisten dieser Nachbarn nicht derselben Klasse angehören wie die Instanz, werden die Instanz und ihre Nachbarn aus dem Datensatz entfernt [37]. Auf diese Weise kann ENN die Präsenz von Mehrheitsklassenproben im Trainingssatz effektiv verringern. Nehmen wir als Beispiel die Probe e in Abbildung **5 : Ihre k (auf 3 gesetzt) Nachbarn stammen alle aus anderen Klassen. Mit der ENN-Methode werden alle vier Proben entfernt.**

3.6.3. SMOTE-ENN

SMOTE-ENN kombiniert die Stärken von SMOTE und ENN in einem einheitlichen Ansatz zum Datenausgleich, wie von Batista et al. [38] vorgeschlagen. Zunächst wird der Datensatz durch den SMOTE-Algorithmus verbessert, der die Darstellung der Minderheitsklasse durch synthetische Stichprobenerstellung verstärkt. Sobald eine ausreichende Anzahl von Stichproben der Minderheitsklasse erstellt wurde, wird der ENN-Algorithmus angewendet, um den Datensatz weiter zu verfeinern, indem Instanzen nahe der Klassifizierungsgrenze entfernt werden. Diese sequentielle Anwendung von SMOTE und ENN ermöglicht eine gleichmäßigere Verteilung des Stichprobensatzes über die K

3.7. Eingabe für das LSTM-Modell

Die als F bezeichnete Folge von Pfadsignaturen dient als Eingabe für das LSTM-Modell und bereitet die Bühne für den Klassifizierungsprozess. Die Konfiguration der Zeitschritte innerhalb des LSTM korreliert direkt mit der um eins reduzierten Länge des anfänglich Folglich kapselt das Feature fn+1 beim n-ten Zeitschritt die Pfadsignatur eines Verkehrspfads, der aus n + 1 Datenpaketen besteht. Diese Anordnung stellt sicher, dass mit jedem zusätzlichen Zeitschritt inkrementell mehr Informationen aufgenommen werden, was die Datendarstellung bereichert . Insbesondere integriert das Feature beim letzten Zeitschritt die Pfadsignatur des gesamten Verkehrspfads und enthält die umfassendsten v

Die inhärente Speicherfähigkeit des LSTM-Modells [28] spielt eine entscheidende Rolle, da sie es ermöglicht , **bei jedem Zeitschritt Informationen aus vorhergehenden Schritten zu nutzen und zu übermitteln. Dadurch aggregiert der abschließende Zeitschritt effektiv Informationen über die gesamte Sequenz hinweg und umgeht die potenziellen Nachteile, die sich ergeben, wenn man sich ausschließlich auf globale Merkmale verlässt. Diese umfassende** Integration von **Informationen steigert die Klassifizierungspräzision des Mod**

Machine Translated by Google

3.8. Vollständig verbundene Schicht

Nach der Verarbeitung der Eingabesequenz durch das LSTM wird die Ausgabe des letzten Zeitschritts wird extrahiert und dann zu einer vollständig verbundenen Schicht geleitet, wo die Klassifizierungsbemühungen werden verwirklicht. Das Wesen dieser Schicht liegt in ihrer Fähigkeit, die mit LSTM erfassten Datenmerkmale in den Bereich des Label-Raums. Die Operation innerhalb der Die vollständig verbundene Schicht wird mathematisch wie folgt dargestellt:

$$Y = B \times X + \mathbf{B.} \quad \textbf{(14)}$$

wobei Y **der Ausgabevektor,** W **das Gewicht der vollständig verbundenen Schicht und** X **der Eingabevektor ist Vektor, abgeleitet aus dem letzten Zeitschritt des LSTM, und** B **ist die Verzerrung.**

3.9. Arbeitsablauf abschließen

In diesem Abschnitt werden wir den Arbeitsablauf unserer vorgeschlagenen Methode im Detail erläutern. Um ein Verkehrssegment zu erfassen, extrahieren wir zunächst die beiden grundlegenden statistischen Merkmale der Paketlänge und Paketankunftszeit, um den Verkehrspfad zu bilden. Wir teilen dann die Paketlängensequenz auf in zwei Pfade, die die Richtung der Paketübertragung darstellen. Als nächstes führen wir eine Akkumulationsprozess, um drei Pfade zu erhalten, die die kumulative Paketlänge darstellen Übertragung. Dadurch ergibt sich ein siebendimensionaler Verkehrspfad.

Mit Hilfe der Pfadsignatur-Funktion transformieren wir die siebendimensionale Verkehrspfad in hochdimensionale Pfadsignaturmerkmale, die dann in das LSTM eingespeist werden Modell. Das LSTM-Modell verarbeitet die Eingabesequenz schrittweise, Schritt für Schritt. Bei jedem Zeitschritt interagieren die Eingabemerkmale mit dem verborgenen Zustand und dem Speicherzellenzustand aus dem vorherigen Zeitschritt. Durch eine Reihe von Gating-Mechanismen werden der verborgene Zustand und das Gedächtnis Der Zellzustand wird bei jedem Zeitschritt aktualisiert, sodass das Modell zeitliche Abhängigkeiten innerhalb der Verkehrsdaten.

Nach der Verarbeitung der gesamten Sequenz wird der verborgene Zustand des LSTM-Modells am Ende Zeitschritt wird als Feature-Ausgabe ausgewählt. Dieser versteckte Zustand hat integrierte Informationen aus der gesamten Eingabesequenz und stellt die globalen Eigenschaften der Sequenzdaten dar. Der ausgewählte verborgene Zustand wird dann in eine vollständig verbundene Schicht eingegeben. Die vollständig verbundene Schicht verarbeitet durch eine Reihe von Gewichten und Aktivierungsfunktionen die verborgenen Zustand, um eine Wahrscheinlichkeitsverteilung für das Klassifizierungsergebnis zu generieren. Tabelle 2 zeigt die Veränderungen in der Form in jeder Phase der Verarbeitung einer verschlüsselten Verkehrssequenz, die 24 Pakete.

Tabelle 2. Die Form der Daten ändert sich in jeder Verarbeitungsphase.

Bühne	Beschreibung	Form
Rohdaten	Erfassen Sie verschlüsselte Verkehrssequenzen	—
Verkehrswege erstellen	Extrahieren Sie Paketlänge und Ankunftszeit	(24, 2)
Pfadaufteilung	Teilen Sie die Längensequenz in zwei Pfade auf	(24, 4)
Pfadakkumulation	Akkumulieren Sie Sequenzen mit drei Längen	(24, 7)
Pfad extrahieren	Verwenden Sie die Funktion „Pfadsignatur“, um zu extrahieren	(24, 56)
Unterschrift	Pfadsignaturfunktionen	
LSTM-Eingabe	Verwenden Sie Pfadsignaturfunktionen als Eingabe für	(1, 24, 56)
Versteckte LSTM-Einheiten	LSTM 64 versteckte Einheiten der LSTM-Schichtprozesseingabe	(1, 24, 64)
	Sequenz, die versteckte Zustände zurückgibt für alle Zeitschritte	
Endgültiger verborgener Zustand	Wählen Sie den verborgenen Zustand des letzten Zeitschritts in	(1, 64)
	das LSTM-Modell und geben Sie die vollständig Verbundene Schicht	
Vollständig verbundene Layer-Ausgabe	Die vollständig verbundene Schicht gibt die	(1, 6)
	Wahrscheinlichkeitsverteilung der Klassifikationsergebnis	

Machine Translated by Google

4. Diskussion

In diesem Abschnitt werden wir auf verschiedene Aspekte unseres experimentellen Rahmens eingehen und verschiedene Facetten beleuchten, wie etwa die experimentelle Umgebung, Bewertungsmetriken, verwendete Datensätze, Datenverarbeitungstechniken und die Auswa
Um die Wirksamkeit und Wettbewerbsfähigkeit unserer Methode zu beweisen, führten wir außerdem **Vergleichsexperimente mit gängigen Mainstream-Techniken durch. Darüber hinaus wurden Ablationsstudien durchgeführt, um den Beitrag und die Auswirkungen einzelner Komponenten innerhalb unseres vorgeschlagenen Rahmens zu unterstreichen.**

4.1. Experimentelle Umgebung

Die Experimente wurden auf einem Windows 10-Betriebssystem mit Python Version 3.9 durchgeführt. Wir haben das LSTM-Modell mit PyTorch Version 1.13.1 implementiert. Wir haben die anfängliche Lernrate für das Modelltraining auf 0,0001 und die Trainingsepochen auf 500 festgelegt. Das für dieses Experiment verwendete LSTM-Modell hat 64 verborgene Dimensionen, und die Ausgabedimension der nachfolgenden vollständig verbundenen Schicht beträgt 6, was der Sechs-Klassen-Klassifizierungsaufgabe verschlüsselter Verkehrsdiensttypen entspricht.

4.2. Bewertungsmaßstäbe

Zur Beurteilung der Leistung der Methode haben wir vier Bewertungsmaßstäbe verwendet: Genauigkeit (AC), Präzision (PR), Trefferquote (RC) und F1-Score (F1). Wahre Positive (TP) sind korrekt identifizierte positive Proben, falsche Positive (FP) sind negative Proben, die fälschlicherweise als positiv klassifiziert wurden, Wahre Negative (TN) sind korrekt identifizierte negative Proben und falsche Negative (FN) sind positive Proben, die fälschlicherweise als negativ klassifiziert wurden.

$$\text{Genauigkeit} = \frac{TP + TN}{TP + TN + FP + FN}, \quad \textbf{(15)}$$

$$\text{Präzision} = \frac{TP}{TP + FP,} \quad \textbf{(16)}$$

$$\text{Rückruf} = \frac{TP}{TP + FN}, \quad \textbf{(17)}$$

$$\text{F1-Punktzahl} = \frac{\mathbf{2} \ddot{y} \text{ Präzision } \ddot{y} \text{ Rückruf}}{\text{Präzision + Rückruf}}. \quad \textbf{(18)}$$

Diese Kennzahlen bieten einen umfassenden Überblick über die Wirksamkeit des Modells und fassen seine Genauigkeit, Präzision und Sensibilität bei der Klassifizierung der Daten in die jeweiligen Kategorien.

4.3. Datensätze

Um die Leistungsfähigkeit unserer vorgeschlagenen Methode umfassend zu überprüfen, verwendeten wir zwei öffentlich verfügbare Datensätze: den ISCX VPN-nonVPN-Datensatz [39] und den ISCX Tor-nonTor- Datensatz [40]. Diese Datensätze wurden von der University of New Brunswick veröffentlicht.
Der ISCX VPN-nonVPN-Datensatz besteht aus zwei Teilen: ISCX-VPN und ISCX-nonVPN.
Der ISCX-VPN-Datensatz enthält in einer VPN-Umgebung erfassten Datenverkehr für sechs Diensttypen: Chat, E-Mail, Dateiübertragung, Streaming, P2P und VoIP. Der ISCX-nonVPN-Datensatz enthält dieselben Datenverkehrstypen, die jedoch außerhalb der VPN-Umgebung erfasst wurden.

Ebenso ist der ISCX Tor-nonTor-Datensatz in ISCX-Tor und ISCX-nonTor unterteilt. Der ISCX-Tor-Datensatz erfasst den Datenverkehr über das Tor-Netzwerk, was ein hohes Maß an Anonymität bietet. Der ISCX-nonTor-Datensatz enthält den Datenverkehr, der außerhalb der Tor-Umgebung erfasst wurde, wobei die Datenverkehrstypen mit denen im ISCX-VPN-Datensatz identisch sind. Tabelle 3 zeigt die Hauptanwendungen, die den einzelnen Diensttypen entsprechen, und bietet eine kurze Beschreibung.

Machine Translated by Google

Tabelle 3. Datensatzdetails.

Service	Anwendung	Beschreibung
Plaudern	AIM, ICQ, Skype, Facebook, Hangouts	Während des Online-Chats generierter Datenverkehr Kommunikation
Email	E-Mail, Gmail	Während des E-Mail-Versands generierter Datenverkehr Übertragung
Datei Übertragung	Skype SFTP, FTPS, SCP	Beim Hochladen von Dateien generierter Datenverkehr und Downloads
Streaming	Vimeo, YouTube, Netflix, Spotify	Der Datenverkehr bei Verwendung Streaming-Anwendungen
P2P	uTorrent, BitTorrent	Beim Teilen von Torrents generierter Datenverkehr Ressourcen mithilfe von P2P-Programmen
VoIP	Facebook, Skype, Hangouts, VoipBuster	Während Online-Sprachanrufen generierter Datenverkehr

Nach Erhalt des Datensatzes verwendeten wir die Paketlänge und die Ankunftszeit, um zu konstruieren unsere Pfadsignaturfunktionen. Die Paketlänge bezieht sich auf die Bytegröße jedes Pakets im Datenstrom. Da die Verkehrsübertragung in einem Netzwerk normalerweise bidirektional erfolgt, haben wir positive und negative Zeichen, um die Richtung des Datenverkehrs anzuzeigen: Datenverkehr vom Client zum Der Server wurde als positiv und der Datenverkehr vom Server zum Client als negativ eingestellt. Dadurch entsteht eine Paketlängensequenz, die sowohl die Übertragungsrichtung als auch Paketladegröße.

Die Paketankunftszeit bezieht sich auf die Zeit, zu der jedes Paket empfangen wird, beginnend mit der Empfang des ersten Pakets und die Aufzeichnung der Ankunftszeit jedes nachfolgenden Pakets. Daraus ergibt sich eine zunehmende Zeitreihe, die die Übertragungshäufigkeit der Pakete. Durch Segmentierung und Akkumulation der Paketlängensequenz und Normalisierung Die Paketankunftszeiten haben wir den endgültigen Verkehrspfad erhalten. Dieser Pfad wurde dann eingegeben in die Pfadsignaturfunktion zur Ableitung der Verkehrspfadsignaturfunktionen, die für unsere Methode.

4.4. Datenausgleich

Bei der Behebung des Stichprobenungleichgewichts im ISCX VPN-nonVPN-Datensatz für Bei der Klassifizierung des verschlüsselten Datenverkehrs verwendete unsere Methodik einen vielschichtigen Ansatz um eine gerechte Vertretung der verschiedenen Verkehrsdiensttypen zu gewährleisten. Zunächst wurde ein Schwellenwert 5000 Proben festgelegt, wobei Kategorien, die diese Zahl überschreiten, eine zufällige Teilmenge von 5000 Proben ausgewählt, während Kategorien unterhalb dieser Schwelle verwendet wurden alle verfügbaren Beispiele, um die Datennutzung zu maximieren. Nach dieser Auswahl wird die Kategorie mit dem Die niedrigste Probenanzahl wurde ermittelt und damit ein Maßstab für Einheitlichkeit gesetzt. Ein Teil Entsprechend einem Fünftel dieser Mindestanzahl wurde aus jeder Kategorie extrahiert, um einen umfassenden Testsatz zusammenzustellen und eine ausgewogene Darstellung sicherzustellen. Die SMOTE-ENN wurde angewendet, um das verbleibende Ungleichgewicht im Trainingsdatensatz zu beheben, indem neue Stichproben für Minderheitsklassen und Entfernen mehrdeutiger Stichproben, um die Qualität des Datensatzes zu verbessern. Abbildung 6 vergleicht die Anzahl der Proben für jede Kategorie vor und nach dem Ausgleich für die Trainingssatzdaten im ISCX-VPN-Datensatz.

Machine Translated by Google

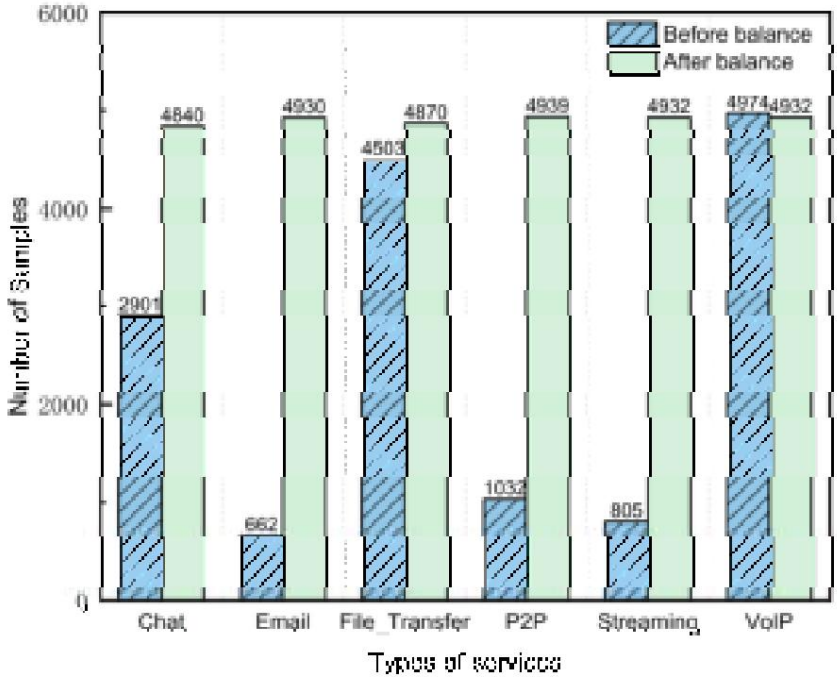


Abbildung 6. Die Anzahl der Proben.

4.5. Parameterauswahl

In unseren Experimenten spielten die Werte zweier entscheidender Hyperparameter, nämlich der Länge der Verkehrswege und der Dimensionen der Pfadsignaturen, eine bedeutende Rolle und standen in engem Zusammenhang mit der Leistung des Modells. Sie prägten direkt die Fähigkeit des Modells, aus den Daten zu lernen und sie zu interpretieren, und beeinflussten somit die Tiefe Die Länge des Verkehrspfads bestimmt die Eingabedaten für jede Analyseinstanz und beeinflusst das Verständnis des Modells für die Sitzungskontinuität. Gleichzeitig beeinflusst die Dimensionalität der Pfadsignaturen die Komplexität und Detailliertheit der aus den Verkehrspfaden extrahierten Merkmale und sorgt für ein Gleichgewicht zwischen Informationsreichtum und rechnerischer Machbarkeit. In diesem Abschnitt erläutern wir den Prozess zum Erhalten dieser Hyperpara

Die optimale Länge des Verkehrspfads wird durch die Notwendigkeit bestimmt, eine ausreichende Anzahl aufeinanderfolgender Datenpakete mit gültigen Nutzdaten in einer Sitzungsdatei zu speichern, die die Grundlage des ursprünglichen Verkehrspfads bildet . Eine zu kurze Pfadlänge kann dazu führen, dass Sitzungsdateien keine ausreichenden Daten für eine wirksame Verkehrsklassifizierung enthalten. Umgekehrt können übermäßig lange Verkehrspfade den Datenerfassungsaufwand unnötig erhöhen, ohne dass die Klassifizierungsgenauigkeit proportional zunimmt [41]. Die Ermittlung der optimalen Länge der Verkehrspfade war ein grundlegendes Element unseres Versuchsdesigns.

Bei der Klassifizierung von Diensttypen für verschlüsselten Datenverkehr unter VPN-Bedingungen wollten wir die optimale Pfadlänge ermitteln. Wie in Abbildung 7 dargestellt, schwankte die Klassifizierungsgenauigkeit bei relativ kurzer Pfadlänge und verbesserte sich mit zunehmender Pfadlänge für jede beliebige Dimension allmählich. Bei einer Pfadlänge von 20 erzielten Modelle mit mehreren Dimensionen eine relativ gute Klassifizierungsleistung. Dabei erreichte das zweidimensionale Pfadmodell bei einer Pfadlänge von 24 die höchste Klassifizierungsgenauigkeit (94,74 %). Mit zunehmender Pfadlänge verbesserte sich die Klassifizierungsgenauigkeit nicht mehr signifikant und nahm in einigen Dimensionen sogar ab. Daher wählten wir 24 als unseren optimalen Parameter für die Pfadlänge.

Darüber hinaus ist die Wahl der Pfadsignaturdimensionen ebenso kritisch. Pfadsignaturen höherer Ordnung kapseln von Natur aus mehr Informationen als ihre Gegenstücke niedrigerer Ordnung. Dennoch führt ein Anstieg der Pfadsignaturdimensionen zu einem exponentiellen Anstieg der Merkmaldimensionen [11]. Beispielsweise hätte ein Signaturmerkmal zweiter Ordnung in einem siebendimensionalen Verkehrspfad eine Dimensionalität von 56, die sich auf 399 für die dritte Ordnung und 2800 für die vierte Ordnung erhöht. Dieser Trend unterstreicht die erforder

Machine Translated by Google

zwischen dem Informationsgewinn durch erhöhte Pfadsignaturdimensionen und der daraus resultierenden Rechenkomplexität.

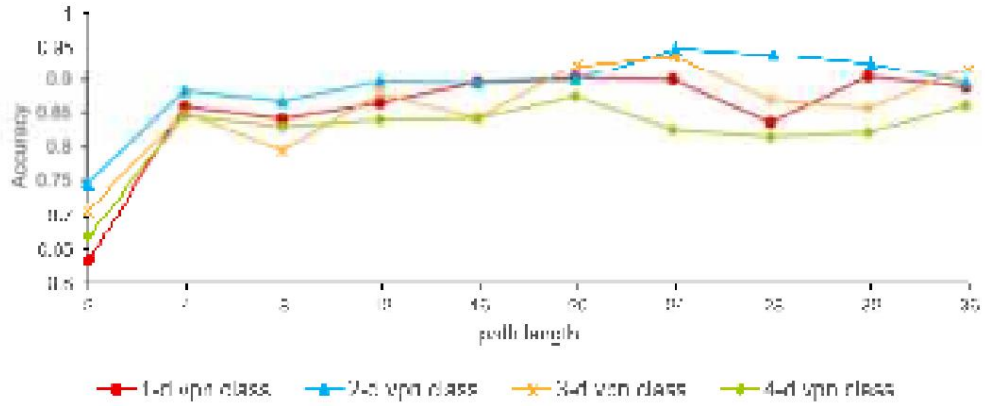


Abbildung 7. Auswirkung der Paketlänge und -dimension (Genauigkeit).

Wie in Abbildung 7 gezeigt, führte die Vergrößerung der Dimensionen auf Grundlage unserer Experimente nicht zu die Klassifizierungsleistung deutlich verbessern. Es gab einen leichten Leistungsabfall, als die Dimensionen auf vier erhöht wurden. Wir entschieden uns für einen zweidimensionalen Signatur als optimaler Parameter unter Berücksichtigung der Klassifizierungsleistung und der Rechenleistung.

4.6. Sensitivitätsanalyse

Um die Robustheit unserer Methode zu bewerten, haben wir eine Sensitivitätsanalyse durchgeführt. Die Sensitivitätsanalyse ist eine Technik, mit der die Reaktion eines Modells auf Parameteränderungen beurteilt wird. Wir haben die Auswirkungen auf die Modellleistung beobachtet, indem wir die Lernrate und die Regularisierungsparameter während des Modelltrainings variiert haben. Insbesondere bei der Bewertung der Aus Lernrate, wir setzen den L2-Regularisierungsparameter auf 0,0001 und bewerten die Um den Einfluss des Regularisierungsparameters zu berücksichtigen, setzen wir die Lernrate auf 0,0001. Die experimentelle Die Ergebnisse sind in den Tabellen 4 **und 5 dargestellt .**

Tabelle 4. Klassifizierungsgenauigkeit bei unterschiedlichen Lernraten.

Lernrate	0,1	0,01	0,001	0,0001	0,00001
Genauigkeit	0,8596	0,8918	0,9152	0,9474	0,9298

Tabelle 5. Klassifizierungsgenauigkeit unter verschiedenen Regularisierungsparametern.

Regularisierungsparameter	0,01	0,001	0,0001	0,00005	0,00001
Genauigkeit	0,8099	0,9006	0,9474	0,921	0,9152

Es kann beobachtet werden, dass sich mit der Änderung der Lernrate und der Regularisierungsparameter Die Klassifizierungsgenauigkeit unserer vorgeschlagenen Methode blieb trotz einiger Schwankungen. Dies zeigt die hohe Robustheit unserer Methode. Die Fähigkeit, eine Hohe Genauigkeit unter verschiedenen Bedingungen beweist die Zuverlässigkeit und Stabilität unseres Ansatzes.

4.7. Ablationsstudie

Um die Beiträge jeder Komponente innerhalb unserer vorgeschlagenen Methode zu unterstreichen für verschlüsselten Datenverkehrsklassifizierung haben wir eine Reihe von Ablationsstudien durchgeführt. Diese Studien waren entwickelt, um die Auswirkungen der Pfadsignaturfunktion einzeln zu bewerten, die mehrskalige kumulative Merkmalsextraktionsmethode und das LSTM-Modell auf die Gesamtleistung.

Machine Translated by Google

Ablation von Pfadsignaturmerkmalen: Für diesen Teil der Studie extrahieren wir keine Pfadsignaturmerkmale mehr aus den Verkehrspfaden. Stattdessen wurden die rohen siebendimensionalen Merkmalsproben direkt in das LSTM-Modell eingespeist, wobei der Schritt der Pfadsignaturverarbeitung umgangen wurde . Alle anderen experimentellen Bedingungen wurden unter Verwen

Ablation der Merkmalsextraktionsmethode: Bei der Bewertung der Rolle unserer mehrskaligen kumulativen Merkmalsextraktionsmethode haben wir diesen fortgeschrittenen Ansatz zugunsten der klassischen Methode, auf die sich Xu et al. [12] beziehen, beiseite gelegt. Die durch diesen traditionellen **Extraktionsprozess erhaltenen Merkmale** wurden dann zur Klassifizierung in das LSTM-Modell eingeführt. Auch diese **Anpassung wurde vorgenommen, während alle anderen experimentellen Parameter unverändert bliebe**

LSTM-Modellablation: Um den spezifischen Beitrag des LSTM-Modells zu untersuchen, haben wir **es die Klassifizierungsaufgabe durch einen Random-Forest-Algorithmus ersetzt, wobei wir Parametereinstellungen verwendeten, die denen von Xu [12] entsprechen. Diese Ersetzung ermöglichte es uns, die Wirksamkeit von LSTM direkt mit einer alternativen, gut etablierten Klassifizierungstechnik zu vergleichen.**

Die Ergebnisse dieser Ablationsexperimente, wie in Abbildung 8 dargestellt, weisen erhebliche ff-Abweichungen auf, wenn man sie mit den Ergebnissen vergleicht, die mit unserer gesamten Methode erzielt wurden. Solche Unterschiede unterstreichen die entscheidende Rolle jeder Komponente in unserer Methodik und bestätigen ihre gemeinsame Notwendigkeit und Wirksamkeit bei der Verbesserung der Klassifizierungs

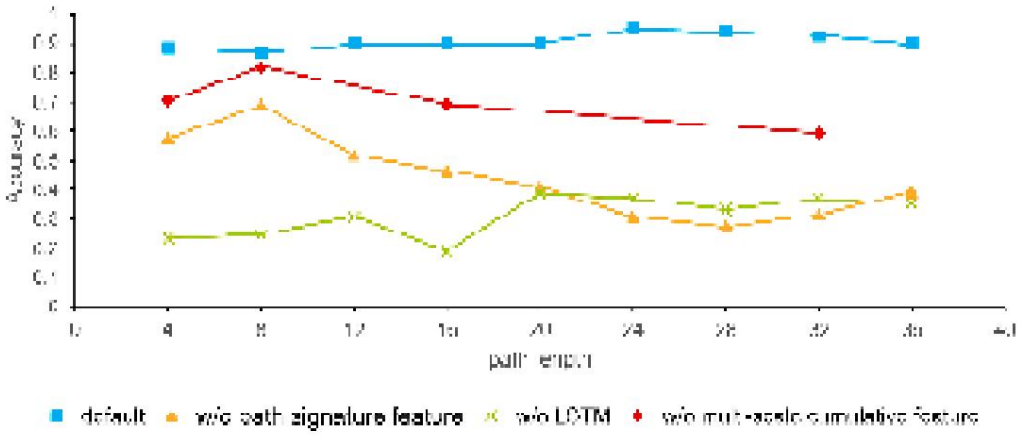


Abbildung 8. Ergebnisse der Ablationsstudien.

Durch diese Ablationsstudien wurden der integrale Wert und die synergetische Wirkung der Pfadsignaturmerkmale, der neuen Merkmalsextraktionsmethode und des LSTM-Modells demonstriert, was ihren unverzichtbaren Beitrag zur überlegenen Leistung unserer Methode bestätigte.

4.8. Vergleichsexperimente **4.8.1.**
Die Benchmark-Methoden

Um die Wirksamkeit unserer vorgeschlagenen Methode zur Klassifizierung von verschlüsseltem Datenverkehr zu ermitteln, wurde sie mit elf aktuellen Mainstream-Ansätzen verglichen. Diese Methoden umfassen Folgendes: 1. Fingerabdruck-Erstellung:

FlowPrint [17]; ff Statistische Methoden: AppScanner [9],
CUMUL [35], K-FP (K-Fingerprinting) [42], 2.
GRAIN [43] und ETC-PS [12];
3. Deep-Learning-Modelle: FS-net [20], DF [19], MT-FlowFormer [23], GraphDApp [44] und ET-BERT (flow) [16].

Besonders bemerkenswert ist, dass unsere vorgeschlagene Methode ausschließlich Flussebenen-Merkmale von verschlüsseltem Datenverkehr verwendet. Der Fairness halber haben wir uns für einen Vergleich mit ET-BERT (Fluss) entschieden, das ebenfalls ausschließlich auf Flussebenen-Merkmalen basiert, und nicht mit ET-BERT (Paket), das eine bessere Klassifizierungsleistung erzielt, aber Paketebenen-Merkmale erfordert.

Machine Translated by Google

4.8.2. Experimentelle Ergebnisse

Wir haben die Wirksamkeit verschiedener Methoden, einschließlich unserer eigenen, auf dem ISCX bewertet VPN-nonVPN und ISCX Tor-nonTor Datensätze, wie in den Tabellen 6 und 7 gezeigt. Tabelle 6 präsentiert die Klassifizierungsleistung von Verkehrsdiensttypen in VPN- und Nicht-VPN-Umgebungen, während Tabelle 7 die Klassifizierungsleistung in Tor- und Nicht-Tor-Umgebungen zeigt. Unsere Methode zeigte in den meisten Bereichen eine überlegene und wettbewerbsfähige Leistung im Vergleich zu den meisten Benchmarkmethoden.

Tabelle 6. Klassifizierungsleistung verschiedener Methoden für den ISCX VPN-Nicht-VPN-Datensatz.

Datensatz	ISCX-VPN				ISCX-nonVPN			
Methode	Genauigkeit	Präzision	Abrufen	F1-Ergebnisse	Genauigkeit	Präzision	Abrufen	F1-Ergebnisse
FlowPrint [17]	0,8538	0,7451	0,7917	0,7566	0,6944	0,7073	0,7310	0,7131
AppScanner [9]	0,8889	0,8679	0,8851	0,8722	0,7576	0,7594	0,7465	0,7486
KUMUL [35]	0,7661	0,7531	0,7852	0,7644	0,6187	0,5941	0,5971	0,5897
K-FP [42]	0,8713	0,8750	0,8748	0,8747	0,7551	0,7478	0,7354	0,7387
KORN [44]	0,8129	0,8077	0,8109	0,8027	0,6667	0,6532	0,6664	0,6567
ETC-PS [12]	0,8889	0,8803	0,8937	0,8851	0,7273	0,7414	0,7133	0,7208
FS-Netz [20]	0,9298	0,9263	0,9211	0,9234	0,7626	0,7685	0,7534	0,7355
DF [19]	0,8012	0,7799	0,8152	0,7921	0,6742	0,6857	0,6717	0,6701
MT-FlowFormer [23]	0,9327	0,9152	0,9243	0,9193	0,8549	0,8473	0,8268	0,8344
GraphDApp [44]	0,6491	0,5668	0,6103	0,5740	0,4495	0,4230	0,3647	0,3614
ET-BERT (Fluss) [16]	0,9532	0,9436	0,9507	0,9463	0,9167	0,9245	0,9229	0,9235
Vorgeschlagen	0,9474	0,9480	0,9474	0,9472	0,9053	0,9064	0,9053	0,9050

Tabelle 7. Klassifizierungsleistung verschiedener Methoden für den ISCX Tor-nonTor-Datensatz.

Datensatz	ISCX-Tor				ISCX-nichtTor			
Methode	Genauigkeit	Präzision	Abrufen	F1-Ergebnisse	Genauigkeit	Präzision	Abrufen	F1-Ergebnisse
FlowPrint [17]	0,2400	0,0300	0,1250	0,0484	0,5243	0,7590	0,6074	0,6153
AppScanner [9]	0,7543	0,6629	0,6042	0,6163	0,9153	0,8435	0,814	0,8273
KUMUL [35]	0,6686	0,5349	0,4899	0,4997	0,8605	0,8143	0,7393	0,7627
K-FP [42]	0,7771	0,7417	0,6209	0,6313	0,8741	0,8653	0,7792	0,8167
KORN [44]	0,6914	0,5253	0,5346	0,5234	0,7895	0,6714	0,6615	0,6613
ETC-PS [12]	0,7486	0,6811	0,5929	0,6033	0,9155	0,8710	0,8311	0,8486
FS-Netz [20]	0,8286	0,7487	0,7197	0,7242	0,9278	0,8368	0,8254	0,8285
DF [19]	0,6514	0,4803	0,4767	0,4719	0,8568	0,8003	0,7415	0,7590
MT-FlowFormer [23]	0,8750	0,8252	0,8217	0,8220	0,8941	0,8742	0,8651	0,8670
GraphDApp [44]	0,4286	0,2557	0,2509	0,2281	0,6936	0,5447	0,5398	0,5352
ET-BERT(Fluss) [16]	0,9543	0,9242	0,9606	0,9397	0,9029	0,8560	0,8217	0,8332
Vorgeschlagen	0,9386	0,9400	0,9386	0,9385	0,9503	0,9510	0,9503	0,9502

Bei diesen Klassifizierungsaufgaben zeigte unsere Methode erhebliche Verbesserungen in Bezug auf Genauigkeit, Präzision, Rückruf und F1-Score im Vergleich zu ETC-PS, das ebenfalls Pfadsignaturen verwend Merkmale. Im Vergleich zum leistungsstärksten Modell, ET-BERT (Flow), ist die Die Genauigkeit war in der VPN-Umgebung nur um 0,58 % und in der Tor-Umgebung um 1,57 % niedriger . Es ist jedoch wichtig anzuerkennen, dass ET-BERT (Flow) ein komplexes, groß angelegtes vorab trainiertes Modell, das in krassem Gegensatz zu unserem rationalisierten Ansatz steht unter Verwendung eines klassischen LSTM-Modells im kleinen Maßstab.

4.8.3. Leistungsanalyse

Um die Komplexität des Modells und den Rechenaufwand umfassender zu bewerten, Um die benötigten Ressourcen zu ermitteln, verwendeten wir Parameter (Params), Gleitkommaoperationen (FLOPs), Inferenzzeit und Latenzzeit als Bewertungsmaßstäbe. Parameter bezeichnen aller lernbaren Gewichte und Biase innerhalb eines neuronalen Netzes, während FLOPs sich hauptsächlich auf die Anzahl der Gleitkomma-Rechenoperationen, die während der Ausführung des neuronalen Netzwerks erforderlich sind Vorwärts- und Rückwärtsausbreitungsprozesse.

In unserem Experiment verwendeten wir eine Batchgröße von 32 Proben. Die Zeit für die Eingabe Die Zeit vom Modell bis zum Erhalt der Vorhersageergebnisse wird als Inferenzzeit bezeichnet, während Latenzzeit bezeichnet die Zeitspanne vom Laden der 32 Samples in den Speicher bis zum Erhalt die Vorhersageergebnisse. Es ist wichtig zu beachten, dass die Anzahl der Parameter und FLOPs nur zur Bewertung von Deep-Learning-Modellen verwendet werden. Daher haben wir nur Methoden bewertet, die Nutzen Sie Deep-Learning-Modelle.

Machine Translated by Google

Tabelle 8 zeigt die Parameter, FLOPs, Inferenzzeit und Latenzzeit für verschiedene Methoden. Die Daten in der Tabelle zeigen deutlich, dass unsere Methode erhebliche Vorteile aufweist in mehreren Aspekten für dieselbe Aufgabe der Klassifizierung verschlüsselten Datenverkehrs.

Tabelle 8. Die Komplexität verschiedener Methoden.

Methode	Parameter (M)	FLOPs (G)	Inferenzzeit (ms)	Latenzzeit (ms)
FS-Netz [20]	2,17	24,86	39,94	45,72
DF [19]	1,83	3,06	29,57	33,56
MT-FlowFormer [23]	0,26	1,07	57,79	63,08
GraphDApp [44]	0,22	0,59	78,43	84,72
ET-BERT (Fluss) [16]	85,70	10,87	104,04	110,94
Vorgeschlagen	0,03	7,32 × 10ÿ4	4,02	5,56

Erstens verwendet unsere Methode deutlich weniger Parameter (Params) als andere Methoden in Bezug auf die Modellkomplexität. Dies bedeutet, dass unsere Modellstruktur prägnanter und effizienter ist, was dazu beiträgt, das Risiko einer Überanpassung zu verringern und gleichzeitig die Interpretierbarkeit des Modells zu ve und Wartbarkeit.

Zweitens, was den Bedarf an Rechenressourcen betrifft, benötigt unsere Methode weniger Gleitkommaoperationen (FLOPs). Dies deutet darauf hin, dass unser Modell weniger Rechenleistung erfordert während der Vorwärts- und Rückwärtsausbreitung, was schnellere Trainings- und Inferenzprozesse ermöglicht. Dies ist besonders wichtig für ressourcenbeschränkte Anwendungen, wie z. B. eingebettete Systeme oder mobile Geräte.

Darüber hinaus weist unsere Methode klare Vorteile sowohl in der Inferenzzeit als auch Latenzzeit. Die Inferenzzeit ist kürzer und die Latenzzeit geringer. Dies deutet darauf hin, dass unser Modell nicht nur rechnerisch effizient, sondern auch sehr reaktionsschnell ist, was es geeignet für Echtzeit- oder nahezu Echtzeitanwendungen. Unter Berücksichtigung all dieser Faktoren halten wir Die leichte Diskrepanz in der Klassifizierungsgenauigkeit ist akzeptabel.

Die kurze Inferenzzeit und Latenzzeit unserer Methode führen wir auf Folgendes zurück: zwei Hauptgründe:

- Unsere Modellarchitektur hat eine geringere Komplexität und erfordert weniger Rechenleistung Ressourcen. Dadurch sind die Berechnungen schnell abgeschlossen, sobald die Proben Eingabe in das Modell, wodurch die Inferenzzeit reduziert wird.
- Unsere Stichprobengröße ist mit einer Eingabegröße von (1,23,56) relativ klein. Darüber hinaus Die Rechenlast der CPU ist minimal und ermöglicht eine schnelle Datenübertragung von von der CPU zur GPU, wodurch die Latenzzeit reduziert wird.

4.8.4. Statistische Analyse

Um die Wirksamkeit unserer Methode weiter zu validieren, haben wir den Wilcoxon-Vorzeichen-Rang-Test verwendet, um festzustellen, ob es statistisch signifikante Unterschiede zwischen unseren Klassifizierungsergebnissen und denen anderer Modelle gibt. Der Wilcoxon-Vorzeichen-Rang-Test ist ein weit ver verwendete statistische Methode zum Vergleich zweier Datensätze [45]. Unsere Nullhypothese geht davon aus folgendes: „Die beiden Erkennungsmethoden haben die gleiche Klassifizierungsleistung“. Wenn der p-Wert größer als 0,05 ist, akzeptieren wir die Nullhypothese. Wenn der p-Wert kleiner als 0,05, lehnen wir die Nullhypothese mit einem Konfidenzniveau von 95 % ab. Die p-Werte, die unsere Die für jedes Basismodell vorgeschlagene Methode ist in den Tabellen 9 und 10 aufgeführt.

Tabelle 9. Die p-Werte beim Vergleich unserer Methode mit der Fingerabdruckkonstruktion und statistischen Methoden.

Methode	FlowPrint [17]	AppScanner [9]	CUMUL [35]	K-FP [42]	KORN [44]	ETC-PS [12]
p-Wert	0,006956	0,036267	0,000186	0,005447	0,000104	0,041655

Tabelle 10. Die p-Werte beim Vergleich unserer Methode mit Deep-Learning-Modellen.

Methode	FS-Net [20]	DF [19]	MT-FlowFormer [23]	GraphDApp [44]	ET-BERT (Fluss) [16]
p-Wert	0,015764	0,000183	0,016351	0,000322	0,167007

Machine Translated by Google

Basierend auf den Ergebnissen können wir schlussfolgern, dass die p-Werte für alle Methoden im Vergleich zu unserer Methode, mit Ausnahme von ET-BERT, kleiner als 0,05 sind. Dies weist darauf hin, dass es, mit Ausnahme von ET-BERT, signifikante Unterschiede in der Klassifizierungsleistung zwischen unserer Methode und den anderen Methoden gibt. Insbesondere bedeutet ein p-Wert von weniger als 0,05, dass wir die Nullhypothese mit einem Konfidenzniveau von 95 % ablehnen können, was bedeutet, dass die Annahme, dass die beiden Methoden die gleiche Klassifizierungsleistung aufweisen, nicht zutrifft.

Daher übertrifft unsere Methode die anderen Basismodelle in Bezug auf die Klassifizierungsleistung deutlich, mit Ausnahme von ET-BERT. Für ET-BERT ist der p-Wert größer als 0,05, was darauf hinweist, dass es keinen statistisch signifikanten Unterschied in der Klassifizierungsleistung zwischen unserer Methode und ET-BERT gibt. Dies bestätigt die Wirksamkeit unserer Methode weiter und zeigt, dass die Vorteile unseres Modells in Bezug auf Komplexität und Rechenressourcen nicht auf Kosten der Klassifizierungsleistung gehen.

5. Schlussfolgerungen und zukünftige Arbeiten

In dieser Studie haben wir eine Methode vorgeschlagen, die Pfadsignaturmerkmale mit einem LSTM-Modell kombiniert, um verschlüsselte Verkehrsdiensttypen unter Verwendung von nur 24 aufeinanderfolgenden Inter- Traffic-Merkmalen zu klassifizieren. Der Kern unserer Methode besteht darin, grundlegende Merkmale (wie Paketgröße und Ankunftszeit) aus der Sequenz verschlüsselter Verkehrspakete zu extrahieren, um detaillierte Verkehrspfade zu erstellen. Diese Pfade durchlaufen einen Transformationsprozess und werden mithilfe mehrskaliger, kumulativer Merkmalsextraktionstechniken verbessert, wodurch letztendlich mehrskalige Pfadsignaturmerkmale generiert werden. Diese Merkmale werden dann effektiv als Eingaben für das LSTM- Modell verwendet, was zu präzisen Klassifizierungsergebnissen

Wir haben unsere Methode anhand der öffentlich verfügbaren Datensätze ISCX VPN-nonVPN und ISCX Tor-nonTor evaluiert. Die experimentellen Ergebnisse zeigten Klassifizierungsgenauigkeiten von 94,74 %, 90,53 %, 93,86 % und 95,03 % in den VPN-, Nicht-VPN-, Tor- und Nicht-Tor-Umgebungen . Diese Ergebnisse sind mit den modernsten Methoden vergleichbar. Darüber hinaus zeigten unsere Leistungsanalyse- und statistischen Analyseexperimente die Vorteile unserer Methode gegenüber groß angelegten neuronalen Netzwerkmodellen.

Da es sich jedoch um eine Klassifizierungsmethode handelt, die auf Paketlängen basiert, ist unser Ansatz, wie auch andere Methoden [12] , die auf Paketlängenmerkmalen basieren, in seiner Wirksamkeit gegen verschlüsselten Datenverkehr, der Paketlängenverschleierungstechniken wie Paket-Padding [46] verwendet, äußerst begrenzt. Paket-Padding ist eine wirksame Verteidigung gegen Verkehrsanalysen, da es alle Paketlängen identisch macht. Die Untersuchung unterscheidbarerer Merkmale im verschlüsselten Datenverkehr könnte dieses Problem lösen. Darüber hinaus verwendet unser Modell eine relativ einfache LSTM-Struktur; es kann daher bei der Bearbeitung komplexerer Klassifizierungsaufgaben in der realen Welt eine schlechte Leistung erbringen, was auf die Notwendigkeit robuste

In zukünftigen Arbeiten planen wir, unseren Ansatz durch die Nutzung weiterer Datensätze zu erweitern, um ihn auf komplexere Netzwerkumgebungen anzuwenden. Darüber hinaus bleibt die Erforschung der Anwendbarkeit unserer Methodik im Bereich der Erkennung von anormalem Datenverkehr eine zentrale Säule unserer zukünftigen Forschungsziele.

Beiträge der Autoren: Konzeptualisierung, YM und NL; Methodik, YM; Software, YM; Validierung , YM, GZ und XY; formale Analyse, YM; Untersuchung, GZ; Ressourcen, YM; Datenkuratierung, XY; Schreiben – Vorbereitung des Originalentwurfs, YM; Schreiben – Überprüfung und Bearbeitung, YM; Visualisierung, GZ; Aufsicht, NL; Projektverwaltung, NL; Mittelbeschaffung, NL. Alle Autoren haben die veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu.

Finanzierung: Diese Arbeit wurde teilweise vom National Social Science Fund of China im Rahmen des Zuschusses 20&ZD293 finanziert.

Datenverfügbarkeitserklärung: In dieser Studie wurden öffentlich verfügbare Datensätze analysiert. Diese Daten finden Sie hier: https://www.unb.ca/cic/datasets/vpn.html (abgerufen am 1. Juni 2024).

Danksagung: Die Autoren möchten den anonymen Gutachtern für ihren Beitrag zu diesem Artikel danken.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte vorliegen.

Machine Translated by Google

Verweise

1. **Bader, O.; Lichy, A.; Hajaj, C.; Dubin, R.; Dvir, A. MalDIST: Von der Klassifizierung verschlüsselten Datenverkehrs zur Erkennung und Klassifizierung von Malware-Datenverkehr. In Proceedings der 2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 8.–11. Januar 2022; S. 527–533.**
2. **Wang, Y.; He, H.; Lai, Y.; Liu, AX Ein zweiphasiger Ansatz zur schnellen und genauen Klassifizierung von verschlüsseltem Datenverkehr.** IEEE/ACM Trans. Netw. **2023,** 31, **1071–1086. [CrossRef]**
3. **Wang, W.; Zhu, M.; Wang, J.; Zeng, X.; Yang, Z. End-to-End-Klassifizierung verschlüsselten Datenverkehrs mit eindimensionalen Faltungsneuronalen Netzwerken. In Proceedings der 2017 IEEE International Conference on Intelligence and Security Informatics (ISI), Peking, China, 22.–24. Juli 2017; S. 43–48.**
4. **Shen, M.; Liu, Y.; Zhu, L.; Xu, K.; Du, X.; Guizani, N. Optimierung der Merkmalsauswahl für eine effiziente Klassifizierung von verschlüsseltem Datenverkehr: Ein systematischer Ansatz.** IEEE Netw. **2020,** 34, **20–27. [CrossRef]**
5. **Karagiannis, T.; Papagiannaki, K.; Faloutsos, M. BLINC: Mehrstufige Verkehrsklassifizierung im Dunkeln.** ACM SIGCOMM Comput. Commun. Rev. **2005,** 35, **229–240. [CrossRef]**
6. **Lotfollahi, M.; Jafari Siavoshani, M.; Shirali Hossein Zade, R.; Saberian, M. Deep Packet: Ein neuartiger Ansatz zur Klassifizierung von verschlüsseltem Datenverkehr mittels Deep Learning.** Soft Comput. **2020,** 24, **1999–2012. [CrossRef]**
7. **Al-Naami, K.; Chandra, S.; Mustafa, A.; Khan, L.; Thuraisingham, BM Adaptives Fingerprinting von verschlüsseltem Datenverkehr mit bidirektionaler Abhängigkeit. In Proceedings der Conference on Computer Security Applications, Los Angeles, CA, USA, 5.–9. Dezember 2016.**
8. **Taylor, VF; Spolaor, R.; Conti, M.; Martinovic, I. Robuste Smartphone-App-Identifizierung durch verschlüsselte Netzwerkverkehrsanalyse** IEEE Trans. Inf. Forensics Secur. **2018,** 13, **63–78. [CrossRef]**
9. **Taylor, VF; Spolaor, R.; Conti, M.; Martinovic, I. AppScanner: Automatisches Fingerprinting von Smartphone-Apps aus verschlüsseltem Netzwerkverkehr. In Proceedings des 2016 IEEE European Symposium on Security and Privacy (EuroS&P), Saarbrücken, Deutschland, 21.–24. März 2016; S. 439–454.**
10. **Xie, G.; Li, Q.; Jiang, Y. Selbstaufmerksame Deep-Learning-Methode zur Online-Verkehrsklassifizierung und deren Interpretierbarkeit.** Comput. Netw. **2021,** 196, **108267. [CrossRef]**
11. **Chen, KT Integration von Pfaden – Eine getreue Darstellung von Pfaden durch nichtkommutative formale Potenzreihen.** Trans. Am. Math. Chem. Soc. **1958,** 89, **395–407. [CrossRef]**
12. **Xu, S.-J.; Geng, G.-G.; Jin, X.-B.; Liu, D.-J.; Weng, J. Verkehrspfade erkennen: Klassifizierung verschlüsselten Verkehrs mit Pfadsignaturfunktionen.** IEEE Trans. Inf. Forensics Secur. **2022,** 17, **2166–2181. [CrossRef]**
13. **Yamansavascilar, B.; Guvensan, MA; Yavuz, AG; Karsligil, ME Anwendungsidentifikation durch Netzwerkverkehrsklassifizierung. In Proceedings der 2017 International Conference on Computing, Networking and Communications (ICNC), Silicon Valley, CA, USA, 26.–29. Januar 2017; S. 843–848.**
14. **Chen, RTQ; Rubanova, Y.; Bettencourt, J.; Duvenaud, DK Neuronale gewöhnliche Differentialgleichungen.** Adv. Neural Inf. Process. Syst. **2018,** 31, **07366.**
15. **Dainotti, A.; Pescape, A.; Claffy, KC Probleme und zukünftige Richtungen in der Verkehrsklassifizierung.** IEEE Netw. **2012,** 26, **35–40. [CrossRef]**
16. **Lin, X.; Xiong, G.; Gou, G.; Li, Z.; Shi, J.; Yu, J. Et-bert: Eine kontextualisierte Datagrammdarstellung mit vorab trainierten Transformatoren zur Klassifizierung von verschlüsseltem Datenverkehr. In Proceedings der ACM Web Conference 2022, Lyon, Frankreich, 25.–29. April 2022; S. 633–642.**
17. **Van Ede, T.; Bortolameotti, R.; Continella, A.; Ren, J.; Dubois, DJ; Lindorfer, M.; Choffnes, D.; van Steen, M.; Peter, A. Flowprint: Halbüberwachtes Fingerprinting von mobilen Apps im verschlüsselten Netzwerkverkehr. In Proceedings des Network and Distributed System Security Symposium (NDSS), San Diego, CA, USA, 23.–26. Februar 2020.**
18. **Biju, VG; Prashanth, C. Friedman und Wilcoxon Auswertungen vergleichen SVM, Bagging, Boosting, K-NN und Entscheidungsbaum Klassifikatoren.** J. Appl. Comput. Sci. Methods **2017, 9, 23–47. [CrossRef]**
19. **Sirinam, P.; Imani, M.; Juarez, M.; Wright, M. Deep Fingerprinting: Unterminierung von Website-Fingerprinting-Abwehrmechanismen mit Deep Learning. In Proceedings der 2018 ACM SIGSAC-Konferenz zu Computer- und Kommunikationssicherheit, Toronto, ON, Kanada, 15.–19. Oktober 2018; S. 1928–1943.**
20. **Liu, C.; He, L.; Xiong, G.; Cao, Z.; Li, Z. Fs-net: Ein Flusssequenznetzwerk zur Klassifizierung von verschlüsseltem Datenverkehr. In Proceedings der IEEE INFOCOM 2019-IEEE Conference On Computer Communications, Paris, Frankreich, 29. April–2. Mai 2019; S. 1171–1179.**
21. **Lin, K.; Xu, X.; Gao, H. TSCRNN: Ein neuartiges Klassifizierungsschema für verschlüsselten Datenverkehr basierend auf räumlich-zeitlichen Flussmerkmalen für effizientes Management des IIoT.** Comput. Netw. **2021,** 190, **107974. [CrossRef]**
22. **Zhang, H.; Yu, L.; Xiao, X.; Li, Q.; Mercaldo, F.; Luo, X.; Liu, Q. TFE-GNN: Ein Temporal Fusion Encoder mit Graph Neural Networks zur feinkörnigen Klassifizierung von verschlüsseltem Datenverkehr. In Proceedings der ACM Web Conference 2023, Austin, TX, USA, 30. April–1. Mai 2023; S. 2066–2075.**
23. **Zhao, R.; Deng, X.; Yan, Z.; Ma, J.; Xue, Z.; Wang, Y. Mt-flowformer: Ein halbüberwachter Flusstransformator zur Klassifizierung verschlüsselten Datenverkehrs. In Proceedings der 28. ACM SIGKDD-Konferenz zu Knowledge Discovery und Data Mining, Washington, DC, USA, 14.–18. August 2022; S. 2576–2584.**
24. **Graham, B. Sparse Arrays von Signaturen für die Online-Zeichenerkennung.** arXiv **2013, arXiv:1308.0371.**
25. **Gyurkó, LG; Lyons, T.; Kontkowski, M.; Field, J. Extrahieren von Informationen aus der Signatur eines Finanzdatenstroms.** arXiv **2013, arXiv:1307.7244.**

Machine Translated by Google

26. Perez Arribas, I.; Goodwin, GM; Geddes, JR; Lyons, T.; Saunders, KE Ein signaturbasiertes maschinelles Lernmodell zur Unterscheidung zwischen bipolarer Störung und Borderline-Persönlichkeitsstörung. Transl. Psychiatry **2018, 8, 274.** [CrossRef] [PubMed]

27. Hambly, B.; Lyons, T. Eindeutigkeit für die Signatur eines Pfades mit beschränkter Variation und der reduzierten Pfadgruppe. Ann. Math. **2010,** 171, **109–167.** [CrossRef]

28. Hochreiter, S.; Schmidhuber, J. Langzeit-Kurzzeitgedächtnis. Neural Comput. **1997, 9, 1735–1780.** [CrossRef]

29. Shen, S.; Han, T.; Pang, J. Vorhersage des Luftwiderstandsbeiwerts von Autos mithilfe eines neuronalen Netzwerks mit Langzeit-Kurzzeitgedächtnis und LASSO. Measurement **2024,** 225, **113982.** [CrossRef]

30. Yeganeh, A.; Ahmadi, F.; Wong, YJ; Shadman, A.; Barati, R.; Saeedi, R. Shallow vs. Deep Learning Modelle zur Vorhersage des Grundwasserspiegels: Ein Multi-Piezometer-Datenintegrationsansatz. Wasser Luft Boden. Verschmutzung. **2024,** 235, **441.** [CrossRef]

31. Zhao, L.; Farhi, N.; Valero, Y.; Christoforou, Z. Neuronale Netzwerke mit langem Kurzzeitgedächtnis zur Modellierung des menschlichen Fahrverhaltens. Transp. Res. Procedia **2023,** 72, **2589–2596.** [CrossRef]

32. Eswarsai. Verschiedene **Arten von LSTMs erkunden. Online verfügbar: https://medium.com/analytics-vidhya/exploring-different-types-of-lstms-6109bcb037c4 (abgerufen am 28. Juli 2024).**

33. Sun, Y.; Wang, X.; Tang, X. Tief erlernte Gesichtsdarstellungen sind spärlich, selektiv und robust. In Proceedings der IEEE- Konferenz über Computer Vision und Mustererkennung 2015, Boston, MA, USA, 7.–12. Juni 2015; S. 2892–2900.

34. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, AN; Kaiser, ÿ.; Polosukhin, I. Aufmerksamkeit ist alles, was Sie brauchen. In Proceedings of the Advances in Neural Information Processing Systems 30, Long Beach, CA, USA, 4.–9. Dezember 2017; S. 5998–6008.

35. Panchenko, A.; Lanze, F.; Pennekamp, J.; Engel, T.; Zinnen, A.; Henze, M.; Wehrle, K. Website-Fingerprinting im Internet-Maßstab. In Proceedings der NDSS 2016, San Diego, CA, USA, 21.–24. Februar 2016.

36. Chawla, NV; Bowyer, KW; Hall, LO; Kegelmeyer, WP SMOTE: Synthetische Minderheiten-Oversampling-Technik. J. Artif. Intell. Res. **2002,** 16, **321–357.** [CrossRef]

37. Wilson, DL Asymptotische Eigenschaften von Regeln für nächste Nachbarn unter Verwendung bearbeiteter Daten. IEEE Trans. Syst. Man Cybern. **1972, 2, 408–421.** [Querverweis]

38. Batista, GE; Prati, RC; Monard, MC Eine Studie über das Verhalten mehrerer Methoden zum Ausgleichen von Trainingsdaten für maschinelles Lernen. ACM SIGKDD Explor. Newsl. **2004, 6, 20–29.** [CrossRef]

39. Draper-Gil, G.; Lashkari, AH; Mamun, MSI; Ghorbani, AA Charakterisierung von verschlüsseltem und VPN-Verkehr mithilfe zeitbezogener Daten. In Proceedings der 2. Internationalen Konferenz zu Sicherheit und Datenschutz von Informationssystemen (ICISSP), Rom, Italien, 19.–21. Februar 2016; S. 407–414.

40. Lashkari, AH; Gil, GD; Mamun, MSI; Ghorbani, AA Charakterisierung von Tor-Verkehr mithilfe zeitbasierter Funktionen. In Proceedings der International Conference on Information Systems Security and Privacy, Rom, Italien, 29.–31. Mai 2017; S. 253–262.

41. Shapira, T.; Shavitt, Y. FlowPic: Eine generische Darstellung zur Klassifizierung von verschlüsseltem Datenverkehr und zur Identifizierung von Anwendungen. IEEE Trans. Netw. Serv. Manag. **2021,** 18, **1218–1232.** [CrossRef]

42. Hayes, J.; Danezis, G. k-fingerprinting: Eine robuste, skalierbare Website-Fingerprinting-Technik. In Proceedings der 25. USENIX Security Symposium (USENIX Security 16), Austin, TX, USA, 10.–12. August 2016; S. 1187–1203.

43. Zaki, F.; Afifi, F.; Abd Razak, S.; Gani, A.; Anuar, NB GRAIN: Granulare Klassifizierung von verschlüsseltem Datenverkehr mit mehreren Labels mithilfe eines Klassifikators Kette. Comput. Netw. **2022,** 213, **109084.** [CrossRef]

44. Shen, M.; Zhang, J.; Zhu, L.; Xu, K.; Du, X. Genaue dezentrale Anwendungsidentifizierung durch verschlüsselte Verkehrsanalyse mit Graph-Neuralnetze. IEEE Trans. Inf. Forensics Secur. **2021,** 16, **2367–2380.** [CrossRef]

45. Sun, P.; Li, S.; Xie, J.; Xu, H.; Cheng, Z.; Yang, R. GPMT: Generieren von praktischem bösartigem Datenverkehr auf der Grundlage feindlicher Angriffe mit geringe Vorkenntnisse. Comput. Secur. **2023,** 130, **103257.** [CrossRef]

46. Yu, S.; Zhao, G.; Dou, W.; James, S. Vorhergesagte Paketauffüllung für anonymes Surfen im Internet gegen Angriffe durch Verkehrsanalyse. IEEE Trans. Inf. Forensics Secur. **2012, 7, 1381–1393.** [CrossRef]

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultieren.

Machine Translated by Google

Elektronik

MDPI

Artikel

Netzwerk-Intrusion-Detection basierend auf Deep-Belief-Network Umfassendes Lernsystem zur Gleichsetzung

Miaolei Deng [1,2,*], **Chuanchuan Sonne** [1,2], **Yupei Kan** [1,2], **Haihang Xu** [1,2], **Xin Zhou** [1,2] und **Shaojun Fan** [1,2]

[1] Hochschule für Informationswissenschaft und Ingenieurwesen, Technische Universität Henan, Zhengzhou 450001, China; scc@stu.haut.edu.cn (CS); 2022930951@stu.haut.edu.cn (YK); 2022930927@stu.haut.edu.cn (HX); 2023930889@stu.haut.edu.cn (XZ); 2023930813@stu.haut.edu.cn (SF)

[2] Henan Internationales Gemeinsames Labor für Getreideinformationsverarbeitung, Zhengzhou 450001, China

* Korrespondenz: dengmiaolei@haut.edu.cn

Zusammenfassung: Netzwerk-Intrusion-Detection-Systeme sind eine wichtige Verteidigungstechnologie, um die Informationssicherheit zu gewährleisten und ein Netzwerk vor Angriffen zu schützen. In den letzten Jahren hat das Broad-Learning-System viel Aufmerksamkeit auf sich gezogen und wurde mit einigem Erfolg in Intrusion-Detection-Systeme eingeführt. Da das traditionelle Broad-Learning-System jedoch eine einfache lineare Struktur hat, ignoriert es beim Umgang mit unausgewogenen Datensätzen häufig das Merkmalslernen von Minderheitsklassenproben, was zu einer schlechteren Erkennungsrate von Minderheitsklassenproben führt. Zweitens beeinträchtigen die hohe Dimensionalität und die redundanten Merkmale in Intrusion-Detection-Datensätzen auch die Trainingszeit und die Erkennungsleistung des traditionellen Broad-Learning-Systems erheblich. Um die oben genannten Probleme zu lösen, schlagen wir ein Broad-Equalization-Learning-System mit einem Deep-Belief-Network vor. Das Modell lernt den groß angelegten hochdimensionalen Datensatz vollständig über ein Deep-Belief-Network und stellt ihn als optimalen niedrigdimensionalen Datensatz dar. Anschließend führt es die Idee der Neugewichtung des Equalization-Loss-v2 in das Broad-Learning-System ein und lernt, den niedrigdimensionalen Datensatz über ein Broad-Equalization-Learning-System zu klassifizieren. Das Modell wurde experimentell mit dem CICIDS2017-Datensatz getestet und mit dem CICIDS2018-Datensatz vollständig validiert. Im Vergleich zu anderen Algorithmen im gleichen Bereich verkürzt da

Schlüsselwörter: Netzwerk-Intrusion-Detection-System; breites Lernsystem; unausgewogene Datensätze; Ausgleichsverlust v2; CICIDS2017; CICIDS2018

check for updates

Zitat: Deng, M.; Sun, C.; Kan, Y.; Xu, H.; Zhou, X.; Fan, S. Netzwerk Intrusion Detection basierend auf Deep Belief Network Broad Equalization Learning System. Electronics **2024**, 13, **3014**. https://doi.org/10.3390/Elektronik13153014

Wissenschaftlicher Herausgeber: Rajendra V. Boppana

Empfangen: 17. Juni 2024
Überarbeitet: 26. Juli 2024
Akzeptiert: 28. Juli 2024
Veröffentlicht: 31. Juli 2024

Copyright: © 2024 bei den Autoren. Lizenznehmer MDPI, Basel, Schweiz. Dieser Artikel ist ein Open Access-Artikel vertrieben unter den Bedingungen und Bedingungen der Creative Commons Namensnennungslizenz (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Einleitung

Mit der rasanten Entwicklung moderner Computertechnologie und Netzwerkinformationstechnologie beeinflussen Netzwerke jeden Aspekt unseres Lebens, Netzwerkangriffe sind immer komplexer geworden und Netzwerksicherheit war noch nie so wichtig. Obwohl das Internet den Menschen viele Annehmlichkeiten gebracht hat, bringt es auch Die Netzwerksicherheitsprobleme werden immer schwerwiegender. So bedrohen beispielsweise Phishing-Websites, Distributed-Denial-of-Service-Angriffe und Netzwerkwürmer die Ne Netzwerk-Intrusion-Detection-Systeme (NIDS) wurden 1980 von Anderson [1] vorgeschlagen . Das Funktionsprinzip eines Intrusion Detection Systems besteht hauptsächlich darin, das Netzwerkverhalten durch die Analyse von Host-Audit-Daten, Netzwerkverkehrsdaten und anderen Merkmalen zu überwachen und zu filtern , abnormale Zugriffe bei der Netzwerkkommunikation zu identifizieren und die Netzwerkadministratoren umgehend zu benachrichtigen, um die Sicherheit der Netzwerki

Aufgrund der rasanten Entwicklung der künstlichen Intelligenz können maschinelle Lernalgorithmen in den letzten Jahren im Vergleich zu herkömmlichen Angriffserkennungssystemen komplexe Daten für automatisiertes Lernen besser verarbeiten und verfügen über eine bessere Anpassungsfäh Aus diesem Grund haben immer mehr Forscher maschinelle Lerntechnologie in das Intrusion Detection System eingeführt [2] und damit Erfolge erzielt [3]. Traditionelle Modelle des maschinellen Lernens schneiden jedoch schlecht ab, wenn es um komplexe und hochdimensionale

Machine Translated by Google

Daten [4], und die Robustheit des maschinellen Lernmodells schränkt die Leistung des Modells ein **[5]. Deep Learning kann die komplexen Merkmale in den Daten automatisch extrahieren und vollständig erlernen, um eine hohe Genauigkeit und Robustheit zu erreichen. Daher werden Deep-Learning-Techniken häufig in Intrusion Detection-Systemen verwendet [6]. Der Trainingsprozess eines Deep-Learning-Modells ist jedoch rechenintensiv, mit langen Trainingszeiten und hohem Energieverbrauch [7]. Außerdem sind Deep-Learning-Modelle üblicherweise Black-Box-Modelle, denen es an Interpretierbarkeit mangelt, was es schwierig macht, ihren internen Entscheidungsprozess zu verstehen und zu erklären , was die Leistung von Intrusion Detection-Systemen beeinträchtigt. Darüber hinaus schränkt die Robustheit von Deep-Learning-Modellen auch die Sicherheit von Intrusion Detection -Systemen ein. Abbas et al. [8] haben eine ausführliche Studie zur Robustheit der Modelle durchgeführt und einige Erfolge erzielt. Immer mehr Wissenschaftler [9] beginnen, neue Techniken im Bereich der Intrusion Detection zu übernehmen.**

Kürzlich schlugen Chen et al. [10] ein neues randomisiertes neuronales Netzwerk namens Broad Learning System (BLS) vor, das auf einer flachen Netzwerkarchitektur basiert. Dieses Modell unterscheidet sich von der Extend-to-Depth-Deep-Learning-Struktur dadurch, dass es keinen Gradientenabstieg zur Aktualisierung der Gewichte benötigt und die Netzwerkstruktur klein und daher rechnerisch schneller ist **[11]. Darüber hinaus basiert die BLS-Gewichtsaktualisierung auf pseudoinversen Operationen mit besserer Modellinterpretierbarkeit und Skalierbarkeit. Daher haben einige Leute BLS mit guten Ergebnissen in Intrusion Detection Systems eingeführt [12]. Einerseits neigen NIDS-Datensätze jedoch dazu, stark unausgewogen zu sein, und herkömmliche BLS-Modelle können mit unausgewogenen Datensätzen nicht gut umgehen, was zu einer sehr schlechten Erkennungsleistung für Minderheitsklassen von Angriffen führt, was für Intrusion Detection Systems, die Angriffe erkennen, fatal ist. Andererseits neigen NIDS-Datensätze dazu, eine hohe Dimensionalität zu haben und mehr redundante Informationen zu enthalten ,** was zu einem höheren Zeitaufwand und einer geringeren Erkennungsleistung **des BLS-Modells führen kann . Daher müssen bestimmte Maßnahmen ergriffen werde**

Um das Problem der Unausgewogenheit von Datensätzen bei der Einbruchserkennung zu lösen, besteht die traditionelle Lösung darin, die Anzahl der Stichproben der Minderheitsklassen oder der Mehrheitsklassen durch Über- oder Unterabtastung zu verändern, um die Stichproben anzugleichen. Die synthetische Minderheits- Oversampling-Technik (SMOTE) [13] beispielsweise **erhöht die Anzahl der Stichproben der Minderheitsklassen durch Synthese neuer Stichproben der Minderheitsklassen, um die Anzahl der Stichproben unterschiedlicher Klassen im Datensatz auszugleichen. Die von SMOTE generierten Stichproben können jedoch den ursprünglichen Stichprobenraum erweitern und die Verteilung der Stichprobendaten verändern, was zu einer schlechten Klassifizierung oder sogar Verwirrung des Modells im Grenzbereich führen kann und häufig durch die Einführung neuer Stichproben einen höheren Zeitaufwand verursacht . Um das Problem der hohen Dimensionalität von Datensätzen** bei der Einbruchserkennung zu lösen, basiert die traditionelle Methode auf der **Dimensionsreduzierung durch Hauptkomponentenanalyse** (PCA) **[14]. Obwohl es die Dimensionalität des Datensatzes reduzieren kann, kann es als lineare Methode zur Dimensionsreduzierung nur die lineare Korrelationsstruktur in den Daten finden. Bei der nichtlinearen Struktur von Einbruchserkennungsdatensätzen kann PCA diese Informationen nicht effektiv erfassen. Seine Dimensionsreduzierung wird durch die Projektion der Daten in einen neuen niedrigdimensionalen Raum erreicht, was häufig zu einem erheblichen Informationsverlust in den Originaldaten füh**

Um die oben genannten Probleme zu lösen, schlagen wir ein Deep Belief Network Broad Equalization Learning System (DBELS) vor, bei dem das Deep Belief Network (DBN) eine Datendimensionalitätsreduzierung am NIDS-Datensatz durchführt, um redundante Informationen zu eliminieren. Die Idee des Equalization Loss v2 (EQL v2) wird in das BLS eingeführt, und das Modell gleicht die verschiedenen Klassen von Proben aus, indem es die positiven und negativen Gradientenfaktoren anpasst, um die Leistung beim Erkennen von Angriffsproben von Minderheitsklassen zu verbessern, und klassifiziert niedrigdimensionale Datensätze, um das Ziel der Klassifizierung und Erkennung von Angriffen zu erreichen. Die wichtigsten Beiträge dieses Papiers sind wie folgt. • Wir führen die

Idee des EQL v2 in das BLS ein und schlagen ein neues Modell vor, indem wir positive und negative Gradientenfaktoren hinzufügen und seine Gewichte neu berechnen, was die schlechte Lernfähigkeit des BLS-Modells für Minderheitsklassenproben durch Anpassen der positiven und negativen Gradientenfaktoren verbessert und den Defekt des BLS-Modells in

Machine Translated by Google

dass es nicht gut im Umgang mit unausgewogenen Datensätzen ist, um die Leistung beim Erkennen von Minderheitsklassenproben zu verbessern. • Wir haben zwei Arten von DBN-basierten Modellen zur Dimensionsreduktion ausgewertet – das traditionelle DBN-Modell und das DeBN-Modell, das die Idee von EQL v2 einführt – und die Auswirkungen der beiden Modelle zur Dimensionsreduktion auf den Klassifizierungseffekt experimentell verglichen.

• Wir haben viele Experimente mit dem DBELS-Modell auf Grundlage des öffentlich verfügbaren Benchmark-Datensatzes CICIDS2017 durchgeführt, detaillierte Versuchsaufbauten einschließlich binärer und Mehrfachklassifizierung erstellt und das Modell hinsichtlich Genauigkeit, Rückruf, Falsch- Positiv-Rate, Zeit und ROC-Kurven bewertet, wobei wir offensichtliche Vorteile gegenüber anderen Modellen festgestellt haben. Wir haben auch die Auswirkungen von Hyperparametern auf das Modell in vielen Experimenten getestet und vergleichende Analysen mit anderen hochmodernen Modellen durchgeführt. • Wir haben die Eignung und Skalierbarkeit des vorgeschlagenen Modells mit dem CI-CIDS2018-Datensatz weiter validiert. Die Leistung und Nützlichkeit des vorgeschlagenen Modells wurden durch vergleichende Analysen mit anderen Modellen bewertet.

Der Rest dieses Dokuments ist wie folgt gegliedert: Abschnitt 2 befasst sich mit der aktuellen Forschung zu Intrusion Detection Systems auf Basis von BLS, Datendimensionalitätsreduzierung und Datenungleichgewicht. In Abschnitt 3 beschreiben wir die Details des vorgeschlagenen DBELS. Detaillierte experimentelle Ergebnisse werden in den Abschnitten 4 und 5 berichtet und analysiert. Abschnitt 6 schließt dieses Dokument ab.

2. Verwandte Werke

Dieser Abschnitt konzentriert sich auf den aktuellen Stand der Forschung und die Mängel des auf Broad Learning System (BLS) basierenden Network Intrusion Detection System (NIDS). Wir bieten traditionelle Lösungen für die Einschränkungen des traditionellen BLS und analysieren die Vor- und Nachteile dieser Methoden.

2.1. NIDS basierend auf BLS

Das BLS-Modell ist ein flaches Netzwerkmodell, das Features durch Mapping-Knoten und Verbesserungsknoten transformiert und lernt und die Modellgewichte schnell durch pseudoinverse Operationen berechnet, wodurch es einfach und schnell zu trainieren ist. Als es vorgeschlagen wurde, erregte das breite Lernen großes Interesse und seine Anwendungen werden in vielen Forschungsfeldern untersucht, wie etwa Computersehen, Bildverarbeitung, medizinische Datenanalyse und Verarbeitung natürlicher Sprache. Es gibt auch einige Arbeiten auf dem Gebiet der Netzwerk-Intrusion-Detection. Li et al. [15] schlugen ein hybrides Intrusion-Detection-Modell basierend auf dem rekurrierenden neuronalen Netzwerk (RNN) und BLS vor und die experimentellen Ergebnisse zeigen, dass das auf BLS basierende Modell einen besseren Trainingseffekt und eine kürzere Trainingszeit hat. Laura et al. [16] implementierten die Kaskade von Feature-Mapping-Knoten, eine Kaskade von Verbesserungsknoten und eine Kaskade von Feature-Mapping-Knoten und Verbesserungsknoten.
Die Autoren kamen zu dem Schluss, dass die Kaskade von Verbesserungsknoten eine längere Trainingszeit erfordert als andere BLS-Varianten. Anschließend schlugen sie ein DDoS-Erkennungssystem basierend auf breitem Lernen für Kommunikationsnetzwerke vor [17]. Die Autoren kamen zu dem Schluss, dass die beste Erkennungsleistung normalerweise durch die Verwendung eines kaskadierten BLS erreicht werden kann. Li et al. [18] schlugen ein dreifach breites Lernsystem (TBLS) basierend auf dem BLS-Modell vor, das Merkmale aus drei Dimensionen der zeitlichen Granularität, des Dateninhalts und der räumlichen Granularität des Datensatzes lernt. Experimentell wurde nachgewiesen, dass das TBLS-Modell durch das Lernen von Merkmalen aus den drei Dimensionen eine bessere Erkennungsleistung erzielen kann .

Obwohl das BLS-Modell im Bereich der Einbruchserkennung einige Erfolge erzielt hat, verlässt es sich als flaches neuronales Netzwerk zu stark auf die Topologie, die Anzahl der Stichproben und die Klasseninformationen der Trainingsstichproben und kann mit unausgewogenen Datensätzen nicht umgehen. Dies führt dazu, dass den Stichproben der Mehrheitsklasse zu viel Aufmerksamkeit geschenkt wird und das Lernen der Stichproben der Minderheitsklasse beim Lernen der Klassifizierung fast ignoriert wird, was zu einer schlechten Erkennungsleistung bei unausgewogenen Datensätzen führt, insbesondere bei der Minderh

Machine Translated by Google

Klassenproben, die schlechter oder sogar nicht erkennbar sind. Zweitens erhöhen hochdimensionale Datensätze die Komplexität von BLS-Modellen und erhöhen die Rechen- und Sp Darüber hinaus ist der Einfluss von Rauschen und redundanten Merkmalen umso größer, je höher die Datendimensionalität ist. Diese können nützliche Informationen überdecken und die Genauigkeit und Stabilität des BLS-Modells verringern. Daher sind bestimmte Techniken erforderlich, um das BLS an unausgewogene und hochdimensionale Datensätze für die Einbruchserkennung anzupassen.

2.2. NIDS aufgrund von Datenungleichgewicht

Datenungleichgewicht bezieht sich auf die Tatsache, dass in der realen Welt unterschiedliche Datenklassen unterschiedliche Verteilungen aufweisen und bestimmte Datentypen deutlich unterrepräsentiert sind, was sich schwerwiegend negativ auf die Modellklassifizierung auswirkt. Die Datensätze zur Einbruchserkennung sind häufig stark unausgewogen, was sich negativ auf die Trainings- und Erkennungsleistung des Modells auswirkt. Im Fall von CICIDS2017 beispielsweise macht das normale Verhalten von „Benign“ 82,248 % aus und das abnormale Verhalten von „Attack“ nur 17,752 %. Auf DoS-/DDoS-Angriffsproben entfallen 14,571 %, auf Port Scan 2,607 %, auf Brute Force 0,389 %, auf Web Attack 0,096 % und auf Botnet 0,088 %. Es ist daher offensichtlich, dass es viel mehr normale Probenklassen als Angriffsprobenklassen gibt, was dazu führt, dass das Modell zu viele Proben der Mehrheitsklasse lernt. Als Reaktion auf das Ungleichgewicht der Datensätze zur Wu T et al. [13] befassten sich mit unausgewogenen Daten bei der Erkennung von Netzwerkangriffen mithilfe von K-Means - Clustering und der Synthetic Minority Over-Sampling Technique (SMOTE). Sie clusterten Daten, identifizierten Cluster von Minderheitsklassen und verwendeten SMOTE, um synthetische Stichproben zu erzeugen, und demonstrierten die Wirksamkeit in Experimenten. Ahmad T et al. [19] entwickelten ein Hybridmodell, das Merkmalsauswahl und Mustererkennung kombiniert. Sie verwendeten regelbasierte Analysen zur Dimensionsreduzierung , SMOTE zum Ausgleichen der Daten und adaptives Boosting (AdaBoost) zum Lernen und erkannten so effektiv Stichproben von Minderheitsklassen. Mit Fortschritten beim Deep Learning verwendeten **Hao X et al. [20] GANs, um synthetische Datensätze zu erstellen und verbesserten die Klassifikatorleistung bei Minderheitsklassen durch realistische Datengenerierung.**

Die meisten Methoden zur Lösung des Problems des Datenungleichgewichts bei der Einbruchserkennung basieren jedoch auf der Generierung von Daten, sei es durch Überabtastung von Minderheitsklassenproben oder durch Reduzierung von Mehrheitsklassenproben durch Unterabtastung, wodurch die verschiedenen Proben durch Ändern der Anzahl der Datensätze ausgeglichen werden. Dies ändert jedoch die Datenverteilung der Originaldaten, was sich auf die Lernwirkung des Modells zur Klassifizierung des Datensatzes auswirken kann. Darüber hinaus erhöht die Überabtastungsmethode die Anzahl der Datensätze, was den zusätzlichen Rechenaufwand des Modells erhöht und zusätzlichen Zeitaufwand verursacht. In den letzten Jahren haben Jin et al. [21] den Ausgleichsverlust v2 (EQL v2) vorgeschlagen, der sich an positiven und negativen Gradienten orientiert und die Gewichte neu anpasst, um das Lernen jeder Art von Aufgabe auszugleichen und das Lernen von Minderheitsklassenproben zu verbessern, und bestimmte Ergebnisse erzielt. Da die Verteilung und Anzahl der ursprünglichen Stichproben nicht geändert werden und die Gewichte angepasst werden, sodass das Modell jede Klasse von Stichproben auf ausgewogenere Weise lernt, ist dies für die Lösung des Ungleichgewichtsproblems von Intrusion Detection-Systemen von großer Bedeutung. Die Einführung dieser Idee in das BLS trägt dazu bei, die Erkenn

2.3. NIDS basierend auf der Reduzierung der Datendimensionalität

Forscher haben einige Lösungen vorgeschlagen, um das Problem der hohen Dimensionalität im Bereich der Intrusion Detection zu bewältigen. Zhang B et al. [14] schlugen eine Intrusion Detection-Methode vor, die eine verbesserte Hauptkomponentenanalyse (PCA) in Kombination mit dem einfachen Bayes-Algorithmus von Gauss verwendet. Durch Gewichtung der primären Merkmalsvektoren in der PCA reduziert die Methode die Datenkontamination. Dieser Ansatz, gefolgt vom Bayes-Algorithmus zur Erkennung, verkürzt die Erkennungszeit im Vergleich zur Verwendung von Klassifikatoren allein erheblich. Shen Z et al. [22] führten einen verbesserten naiven Bayes-Klassifizierungsalgorithmus ein, der die Hauptkomponentenanalyse mit der linearen Diskriminanzanalyse integriert, um die Dimensionalität des Stich

Machine Translated by Google

Korrelation. Dieser Ansatz ermöglicht es dem Klassifikator, sowohl Attributhäufigkeits- als auch Korrelationsprobleme zu berücksichtigen und so Echtzeit- und Genauigkeitsprobleme bei der Erkennung von Eindringlingen mit zahlreichen Merkmalen zu lösen. Salo F et al. [23] entwickelten eine Integrationstechnik für die Merkmalsverarbeitung , die Informationsgewinn und Hauptkomponentenanalyse kombiniert, um einen niedrigdimensionalen optimalen Unterraum zu extrahieren. Anschließend fusionierten sie mehrere Klassifikatoren mithilfe von Support-Vektor -Maschinen, Instanzlernalgorithmen und mehrschichtigen Perzeptron-Entscheidungsstrategien und verwendeten eine Wahrscheinlichkeitsmittelwertkombinationsregel für die Abstimmung. Tests mit verschiedenen Datensätzen zeigten, dass diese Methode nicht nur eine hohe Genauigkeit erreicht, sondern auc

Die meisten aktuellen Lösungen für das Problem der hohen Dimensionalität von NIDS-Datensätzen sind jedoch Algorithmen maschinellen Lernens, die die Dimensionalität des Datensatzes bis zu einem gewissen Grad reduzieren, aber angesichts nichtlinearer, hochdimensionaler Merkmale immer noch größere Einschränkungen aufweisen, was zu einer niedrigeren Erkennungsrate des Modells führt. Da Deep- Learning-Algorithmen beim Lernen hoher Dimensionalität effektiv sind und nichtlineare Merkmale das Lernen und die Merkmalsextraktion von Rohdaten ohne allzu großes menschliches Eingreifen autonom abschließen können, wurden Deep-Learning-Algorithmen in groß angelegte Intrusion-Detection-Systeme eingeführt , die die Extraktion und Dimensionsreduzierung von Datenmerkmalen besser abschließen können. Das Deep Belief Network (DBN) als repräsentatives Deep- Learning-Modell wurde 2006 von Geoffrey Hinton et al. [24] vorgeschlagen. Es kann **aus den Rohdaten die mehrschichtigen abstrakten Merkmalsdarstellungen lernen, die die komplexe Struktur und die sich ändernden Muster der Daten effektiv erfassen. Dieser hierarchische Merkmalsextraktionsprozess hilft dabei, hochdimensionale Eingabedaten einem niedrigdimensionalen Darstellungsraum zuzuordnen und gleichzeitig die wichtigen Informationen der Daten zur Dimensionsreduzieru**
Im Vergleich zu herkömmlichen Methoden zur Dimensionsreduzierung behält die Dimensionsreduzierung von DBN nicht nur die Struktur und Relevanz der Daten besser bei, sondern lernt auch adaptiv die nichtlinearen Beziehungen in den Daten, was die Ausdrucksfähigkeit und Klassifizierungsleistung der reduzierten Daten verbessert. Daher ist DBN bei der Dimensionsreduzierung für Intrusion-Detection-Modelle üb

3. Methodik

Dieser Abschnitt konstruiert das Deep Belief Network Broad Equalization Learning System (DBELS) basierend auf BLS, EQL v2 und DBN detailliert in drei Phasen: Datenvorverarbeitung, Reduzierung der Datendimensionalität und Modellklassifizierung.

3.1. DBELS-Architektur

Der Rahmen von DBELS ist in Abbildung 1 dargestellt und in drei Phasen unterteilt.
Die erste Phase ist die Datenvorverarbeitung, bei der hauptsächlich die Originaldaten in eine Datenform verarbeitet werden, die für die Dimensionsreduzierung des DBN geeignet ist. Der CICIDS2017-Datensatz enthält 79 Merkmale, darunter viele redundante Merkmale, die das Training und die Klassifizierung des Modells erschweren. In diesem Artikel umfasst die Datenvorverarbeitung zwei Teile: Datenbereinigung und Datenstandardisierung. Der Zweck der Datenbereinigung besteht darin, redundante Merkmale zu reduzieren, indem der Beitrag der Merkmale zur Klassifizierung durch Korrelationskoeffizienten berechnet wird, die schlechten Werte der Daten zu eliminieren und ihre negativen Inte
Der Zweck der Datenstandardisierung besteht darin, das Modell auf ausgewogene Weise auf verschiedene Merkmale der Daten zu konzentrieren und gleichzeitig den Einfluss von Merkmalen mit unterschiedlichen Gewichten auf das Modell aufgrund der Größe des Werts zu ignorieren. Aufgrund des schwerwiegenden Datenungleichgewichts im CICIDS2017- Datensatz kann die Oversampling-Methode SMOTE verwendet werden, um die Minderheitsangriffsproben zu erhöhen, wenn das DBN die Dimensionalität reduziert, um das Problem zu geringer Proben für die Minderheitsklasse zu lindern.

Die zweite Stufe ist die DBN-Dimensionsreduzierung, die darauf abzielt, die **ursprünglichen hochdimensionalen Daten als optimale niedrigdimensionale Daten darzustellen .** Das **von uns verwendete DBN-Dimensionsreduzierungsmodell verwendet eine klassische Struktur, die die Beibehaltung der Kernfunktionen maximiert und gleichzeitig die Merkmalsdimensionen so weit wie möglich reduziert. Das Modell trainiert jede eingeschränkte Boltzmann-Maschine (RBM) vollständig durch Vortr**

Machine Translated by Google

tz

Struktur durch Feinabstimmung der Gewichte und lernt die Merkmale des standardisierten Datensatzes vollständig, der als optimaler niedrigdimensionaler Datensatz dargestellt wird. Die dritte Stufe verwendet das Broad Equalization Learning System (BELS), um die optimalen niedrigdimensionalen Daten zu klassifizieren. Das BELS-Modell gleicht das Lernen der Minderheitsklassenproben durch das Modell aus, indem es die positiven und negativen Gradientenfaktoren anpasst und seine Gewichte erhöht, um die Erkennungs- und Erkennungsklasse der Min

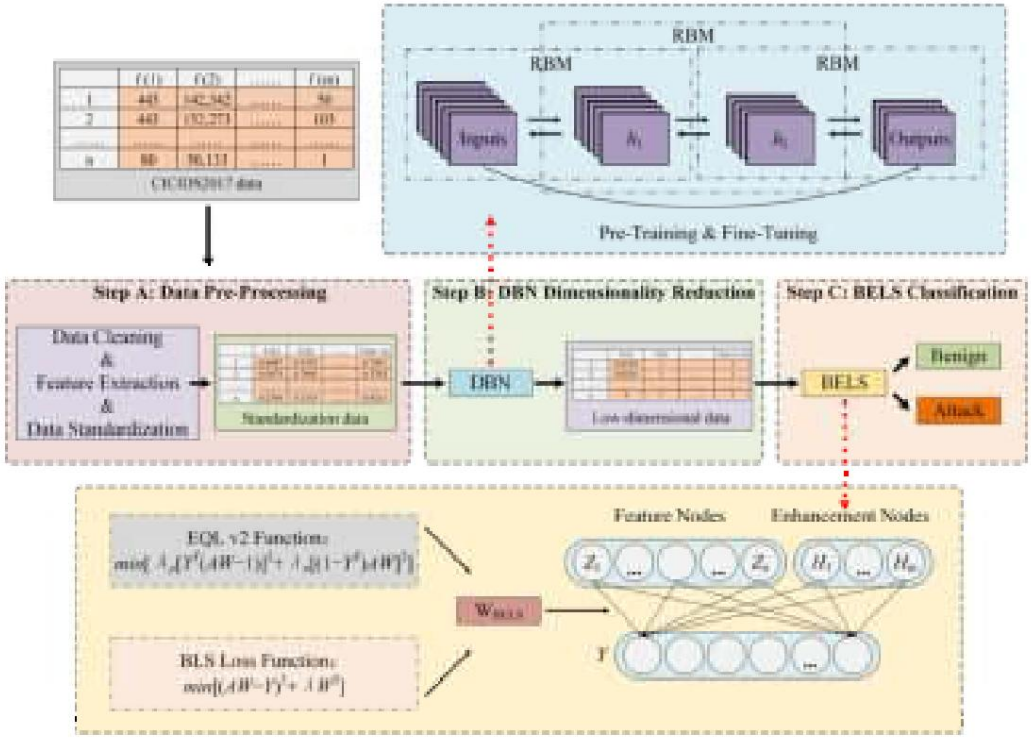


Abbildung 1. Die Struktur von DBELS.

3.2. Deep Belief Network

3.2.1. Eingeschränkte Boltzmann-Maschine

tz

Die eingeschränkte Boltzmann-Maschine [24] ist die Grundstruktur des DBN, das ein flaches neuronales Netzwerk ist und aus einer zweischichtigen Struktur aus sichtbaren und verborgenen Schichten b Verbindungen in einem RBM sind ungerichtet und es gibt keine Verbindungen zwischen Neuronen in derselben Schicht; seine Struktur ist in Abbildung 2 zu sehen. Häufig verwendete RBMs sind im Allgemeinen binär. Unabhängig davon, ob sie sich in der verborgenen oder der sichtbaren Schicht befinden, nimmt ihr Neuron nur den Wert 0 oder 1 an. Das Funktionsprinzip eines RBM basiert auf der Energiefunktion, und die Energiefunktion der sichtbaren und der verborgenen

tz

$$\ddot{y}\, E(v, h) = \ddot{y}\ddot{y} \text{ bivi } \ddot{y}\ddot{y} \text{ cjhj } \ddot{y}\ddot{y} \text{ vihjwij} \quad (1)$$

Schicht ist gegeben durch , ÿÿ

wobei i **zur sichtbaren Schicht gehört und** j **zur verborgenen Schicht,** b **und** c **sind Modellverzerrungen , w ist das Gewicht des Modells, der erste Teil stellt den Beitrag der Knoten der sichtbaren Schicht zur Energie des Systems dar, der zweite Teil stellt den Beitrag der Knoten der verborgenen Schicht zur Energie des Systems dar und der letzte Teil stellt den Beitrag der Energie des Systems aufgrund der Interaktion zwischen der sichtbaren Schicht und der verborgenen Schicht dar.**

Machine Translated by Google

(ÿ)

ÿ

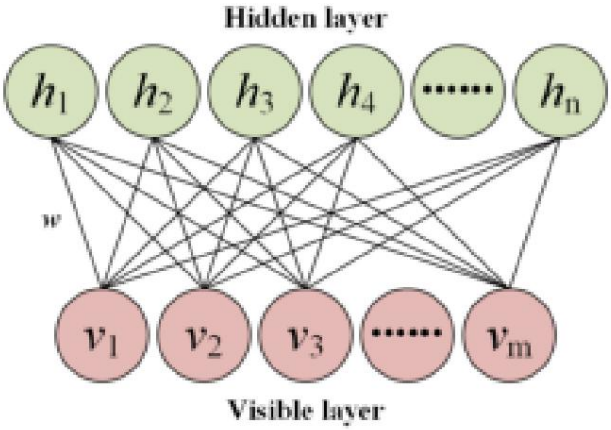


Abbildung 2. Die Struktur von RBM.

Das RBM definiert die gemeinsame Wahrscheinlichkeitsdichte der sichtbaren und verborgenen Schichten in Form einer Energiefunktion. Die gemeinsame Wahrscheinlichkeit der sichtbaren Schicht und der verborgenen Schicht ist gege

$$p(v, h) = \frac{1}{Z} t^{\ddot{y}E(v,h)} \quad \textbf{(2)}$$

wobei Z **die Zustandssumme ist, die verwendet wird, um die Wahrscheinlichkeitsverteilung zu normalisieren und sicherzustellen, dass die Summe der Wahrscheinlichkeitsverteilungen gleich 1 ist:**

$$Z = \ddot{y}_{v,h} \; t^{\ddot{y}E(v,h)} \quad \textbf{(3)}$$

tt

Mit der gemeinsamen Wahrscheinlichkeitsdichte der sichtbaren und verborgenen Schichten, summiert über h **ergibt die Randdichteverteilung der sichtbaren Schicht:**

$$p(v) = \frac{1}{Z\ddot{y}_H} t^{\ddot{y}E(v,h)} \quad \textbf{(4)}$$

Während des Trainings des RBM-Modells wird die Variable h1 der verborgenen Schicht **abgetastet und mit der Posterior-Wahrscheinlichkeit** p(h1|v) berechnet, **und die neue Variable v1** der sichtbaren Schicht **wird mit der Posterior-Wahrscheinlichkeit** p(v1|h1) abgetastet; **anschließend wird die neue Variable h2** der verborgenen Schicht **erneut wiederholt abgetastet, um die sichtbare Schicht v2 zu erhalten. Die obigen Schritte werden mehrmals wiederholt, bis die Parameter konvergieren oder die vordefinierte Anzahl von Iterationen erreichen, die gemeinsame Wahrscheinlichkeitsverteilung nahe an der glatten Verteilung liegt und das RBM-Training abgeschlossen ist.**

3.2.2. Training des Deep Belief Network

Das Deep Belief Network [24] ist ein tiefes neuronales Netzwerk, das aus mehreren RBMs besteht. Die DBN-Struktur ist in Abbildung 3 dargestellt. Während des Trainings verwendet das DBN zunächst einen unüberwachten, gierigen Schicht-für-Schicht-Trainingsalgorithmus für jede Schicht des RBM, mit Ausnahme der ersten und letzten Schicht der DBN-Struktur, bei denen jede Schicht der verborgenen RBM-Schicht als sichtbare Schicht der nächsten RBM-Schicht dient, um Merkmalsinformationen und Lerninformationen zu übertragen (diese Phase wird auch als Vortrainingsphase des DBN bezeichnet). Nach Abschluss des Vortrainings wird eine kleine Menge markierter Daten an die letzte Schicht des DBN-Modells angehängt, und auf jeder Schicht des RBM wird ein überwachtes Training durchgeführt. Der Rückwärtsausbreitungsalgorithmus überträgt die Fehlerinformationen auf jede Schicht des RBM und verwendet die Maximum-Likelihood-Funktion für die Zielfunktion, wodurch das gesamte DBN optimiert werden kann, wodurch die optimale niedrigdimensionale Darstellung der Originaldaten erhalten wird.

Machine Translated by Google

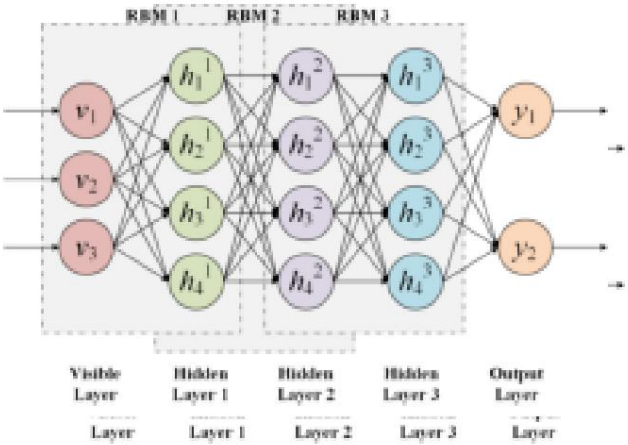


Abbildung 3. Die Struktur von DBN.

3.3. Breites Lernsystem

Der Träger eines breiten Lernsystems [10] ist eine zufällige Vektorfunktion, die neuronale Netzwerk. Im Vergleich zur Deep-Learning-Methode verfügt das BLS nur über zwei Schichten neuronaler Netzwerk, und seine Struktur hat die Form einer Breitenerweiterung. Seine Modellstruktur ist dargestellt in Abbildung 4. Zunächst bildet das BLS die Eingabedaten in eine Feature-Knotenmatrix ab, und zwar durch Abbildungsfunktion. Zweitens bildet die Feature-Knotenmatrix eine Enhancement-Knotenmatrix durch mithilfe der Erweiterungsfunktion. Anschließend werden die Feature-Mapping-Knoten und die Erweiterung Knoten werden zusammen als Eingabe der versteckten Schicht verwendet. Schließlich wird die Pseudoinverse Operation wird verwendet, um die Gewichtsmatrix zwischen der verborgenen Schicht und der Ausgabe zu berechnen Schicht, um den Zweck des Trainings des BLS-Modells zu erreichen.

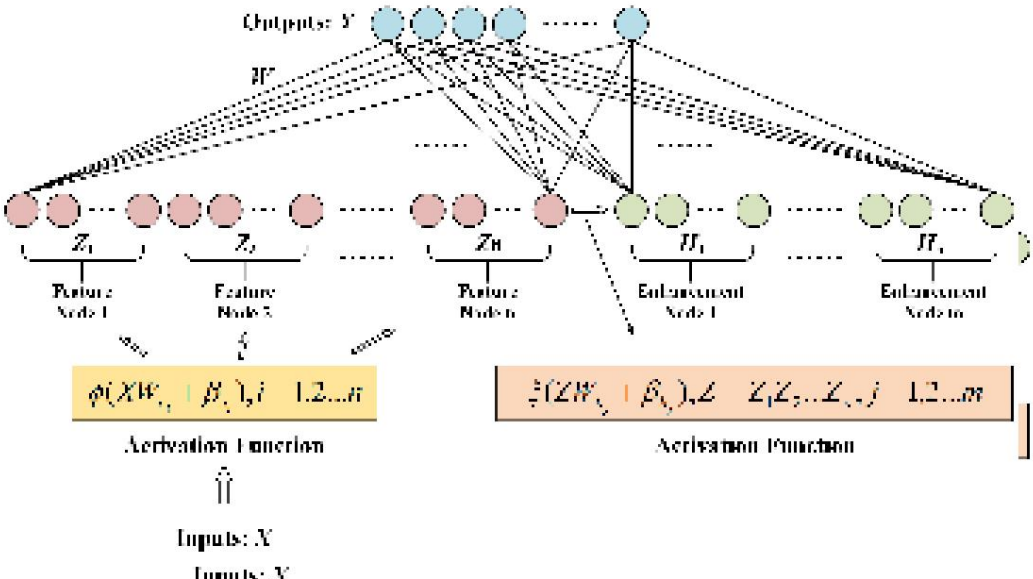


Abbildung 4. Die Struktur von BLS.

Angenommen, die Eingabedaten X = **[x1, x2, . . . ,** xn] **gehören zu** R(N ÿ M) **und die Ausgabe Matrix,** Y = **[y1, y2, . . . ,** yn] **gehört zu** R(N ÿ C), **wobei** N **die gesamte Datenstichprobe ist,** M **ist die Stichprobendimension,** C **ist die Anzahl der Kategorien der Datenstichprobe, Z1, Z2, . . . ,** Zn **H1, H2, . . . ,** Hm **ist die Enhancement-Knotenmatrix und** W **ist die Ausgabegewichtsmatrix zwischen der verborgenen Schicht und der Ausgabeschicht.**

Das spezifische Vorgehen beim BLS ist wie folgt. Zunächst werden die Stichprobendaten X **abgebildet durch die Merkmale** der Gruppe i , um die **Merkmalsknotenmatrix** Zi der Gruppe i zu erhalten , **wobei** i ÿ **(1, 2, ...,** n):

Zi = ÿi(XWei + ÿei) ÿ R N×ÿ i = **1, 2,** (1)

Machine Translated by Google

wobei ÿi **die Aktivierungsfunktion ist,** q **die Anzahl der Merkmalsknoten, die jeder Merkmalszuordnung entsprechen,** Wei **die Gewichtsmatrix der i-ten Merkmalszuordnung und** ÿei **die Bias-Matrix der i-ten Merkmalszuordnung. Sowohl** Wei **als auch** ÿei **werden zufällig generiert. Anschließend werden alle Merkmalsknotenmatrizen integriert, um die gesamte Merkmalsknotenmatrix zu erhalten:**

$$Z^N = [Z1, Z2, \cdots, Zn] \ddot{y} R^{N\times nq} \quad (6)$$

Zweitens wird die Gesamtmerkmalsknotenmatrix Zn **einer Verbesserungsabbildung** der Gruppe j **unterzogen ,** um **die Verbesserungsknotenmatrix** Hj der Gruppe j zu erhalten , **wobei** j ÿ **(1, 2, ...,** m) **wie folgt ist:**

$$Hj = \ddot{y}j\left(Z^{i}Whj + \ddot{y}hj\right) \ddot{y} R^{N\times r} \quad j = 1, 2, \cdots, M \quad (7)$$

wobei ÿj **die Aktivierungsfunktion ist,** r **die Anzahl der Verbesserungsknoten darstellt, die jeder Gruppe von Verbesserungstransformationen entsprechen,** Whj die **Verbesserungszuordnungsgewichtungsmatrix** der Gruppe j **ist und** ÿhj **die Verbesserungszuordnungsbiasmatrix** der Gruppe j ist. Die **gesamte Verbesserungsknotenmatrix wird durch Integration aller Verbesserungsknotenmatrizen erhalten:**

$$H^M = [H1, H2, \cdots, Hm] \ddot{y} R^{N\times mr} \quad (8)$$

Anschließend werden die Feature-Mapping-Knoten mit den Erweiterungsknoten als Eingaben für das BLS zusammengeführt, definiert als A:

$$EIN = [Z^N | H\ m] \ddot{y} R^{N\times(nq+mr)} \quad (9)$$

Schließlich kann der vorhergesagte Wert des breiten Lernalgorithmus wie folgt angegeben werden:

$$\hat{\mathbf{Y}} = [Z\ n | H\ m]W = AW \quad (10)$$

wobei W **die Gewichtsmatrix der Merkmalsknoten und Verbesserungsknoten zur Ausgabeebene ist.**

Um den Fehler zwischen dem vorhergesagten Wert Yˆ und dem wahren Wert Y zu minimieren und **ein geeignetes** W **zu finden , wird das BLS-Modell mit der folgenden Funktion optimiert:**

$$\arg\min_{B} \quad FBLS = \ddot{y}\left(Y - \ddot{y}\hat{\mathbf{Y}}\right)^{2}_{2} + \ddot{y}\left(\ddot{y}W\right)^{2}_{2} = \ddot{y}\left(Y - \ddot{y}AW\right)^{2}_{\ddot{y}} + \ddot{y}\left(\ddot{y}Wm\right)^{2}_{\ddot{y}} \quad (11)$$

wobei der erste Term zur Minimierung des Trainingsfehlers verwendet wird, der zweite Term dazu dient, eine Überanpassung des Modells zu verhindern, und der Wert *ÿ* **die weiteren Einschränkungen für die Summe der quadrierten Gewichte bezeichnet; wir nehmen die Ableitung der obigen Funktion und lassen sie den Wert Null annehmen, wodurch wir** W **als** erhalten

$$W = \left(A^{T}A + \ddot{y}I\right)^{\ddot{y}1} \quad (12)$$

wobei A **die Abbildungs- und Verbesserungsmatrix,** *ÿ* **ein Straffaktor,** I **eine Identitätsmatrix und** Y **die wahre Bezeichnung ist.**

3.4. Equalization Loss v2

Die Grundidee von Equalization Loss v2 [21] besteht darin, den Lerngrad des Modells für verschiedene Klassen von Proben auszugleichen, indem die positiven und negativen Gradienten jedes Klassifikators entsprechend den kumulativen Gradientenverhältnissen der positiven und negativen Gradienten des Klassifikators gewichtet werden. Bei der Lösung des Problems unausgewogener Datensätze ist die Focal-Strategie eine klassische Strategie, und Eqlv2 wurde auf dieser Grundlage verbessert. EQL v2 bietet jedoch deutliche Vorteile gegenüber dem Focal Loss, insbesondere beim Umgang mit Klassenungleichgewichten, ohne eine Verzerrung zugunsten von Minderheitsklassen einzuführen. Im Gegensatz zum Focal Loss, der den Verlustbeitrag basierend auf der Vorhersageschwierigkeit anpasst, gleicht EQL v2 die Gradienten während des Trainings dynamisch aus, wodurch Überanpassungen bei seltenen Kategorien reduziert werden, während der Fokus auf häufigen Klas

Machine Translated by Google

bietet eine differenziertere Antwort auf Ungleichgewichtsprobleme. Die positiven und negativen Gradienten für die Ausgabe Z **jedes Klassifikators in Bezug auf den Verlust** L **sind**

$$\ddot{y}_{zj}^{pos}(L) = \frac{1}{|Ich|} \ddot{y}_{j\,ich\ddot{y}Ich} {}_{ich_y}(P_J^{\sim} \ddot{y} 1) \quad (13)$$

$$\ddot{y}_{zj}^{neg}(L) = \frac{1}{|Ich|} \ddot{y}_{ich\ddot{y}Ich} (1 \ddot{y} y \quad {}_J^{\sim})P_J^{\sim} \quad (14)$$

wobei l **die Gesamtzahl der Proben ist,** p j **ist die Wahrscheinlichkeit, dass die i-te Stichprobe vorhergesagt wird zur Klasse j zu gehören , und ist der Wert der One-Hot-Matrix für wahrheitsgetreue Beschriftungen, in der die i-te** die yj - **Probe ist zur Klasse j.**

Die Gewichte von EQL v2 werden aktualisiert, indem zunächst gj **als Verhältnis des kumulierten positiven Gradienten zum negativen Gradienten von Aufgabe** j **bis zur Iteration t definiert wird. In dieser Iteration können die Gewichte des positiven Gradienten** qt **und des negativen Gradienten** rt **wie folgt berechnet werd**

$$(t)\,(t) = 1 + \ddot{y}\,1\,\ddot{y}\,f(g)\,q\,jj \quad (15)$$

$$r_j^{(t)} = f(g\,j\quad^{T}) \quad (16)$$

wobei f **eine Abbildungsfunktion ist als**

$$f(x) = \frac{1}{\ddot{y}\ddot{y}(x\ddot{y}\mu)\,1 + e} \quad (17)$$

Der positive Gradient q **und der negative Gradient** r **werden erhalten und dann werden sie auf die positiven bzw. negativen Gradienten der aktuellen Charge angewendet, und der** neu **gewichtete Gradient ist**

$$pos\ddot{y}\,\ddot{y}\,zj(L\,(t)) = q_J^{(t)\,pos\,\ddot{y}\,zj}(L\,(t)) \quad (18)$$

$$neg\ddot{y}\,(L\,(t)) = r\,\ddot{y}\,zj^{(t)\,neg\,\ddot{y}\,j}\,(L\,(t)) \quad (19)$$

Schließlich wird das des akkumulierten positiven Gradienten zum negativen Gradienten von Verhältnis g j **der Iteration** t + **1 Update berechnet und bezeichnet als**

$$g_j^{t+1} = \frac{\ddot{y}_{t=0}^{T}\ \ddot{y}\,zj^{posy}(L\,(t))}{\ddot{y}_{t=0}^{T}\ neg\ddot{y}\,\ddot{y}\,zj(L\,(t))} \quad (20)$$

EQL v2 wählt die Gradientenstatistik als Maß dafür, ob sich eine Aufgabe in einem ausgeglichenen Trainingszustand befindet, und gleicht das Modell danach aus, wie einfach oder schwierig jede Klassifizierung für das Modell zu trainieren ist, anstatt einfach die Anzahl seiner positiven und negativen Stichproben zu berücksichtigen. Die Einführung der Idee der gradientengesteuerten Gewichtung in DBN und BLS hilft dabei, das Modell mit mehr Aufmerksamkeit auf die Merkmale von Minderheitsklassen von Stichproben zu trainieren, wodurch das Modell bei der Bewertung jeder Klasse von Stichproben umfassender und a Das Modell passt den optimalen Gewichtungsbeitrag der Iteration basierend auf den Auswirkungen verschiedener Kategorien von Proben auf die Leistung der Klassifizierung an.

3.5. Breites Ausgleichslernsystem

Das W **des breiten Ausgleichslernsystems wird wie folgt gelöst; basierend auf der Positive und negative Gradienten in EQL v2. Die Beispiele für positive und negative Gradienten können wie folgt angegeben werden:**

$$p(W) = Ja\quad^{T}(AW \ddot{y} 1) \quad (21)$$

$$n(W) = (1 \ddot{y} Y\quad^{T})AW \quad (22)$$

Machine Translated by Google

wobei A **die Merkmalsmatrix der zugeordneten Knoten und der Erweiterungsknoten,** Y **die wahre Bezeichnung und** W **die Gewichtsmatrix ist.**

Die Gradientenfunktion kann aus den positiven und negativen Gradienten wie folgt konstruiert werden:

$$G(W) = \ddot{y}p[Y^{T}(AW \ddot{y} 1)]2 + \ddot{y}n[(1 \ddot{y} Y^{T})AW]^{2} \quad (23)$$

wobei ÿp **der positive Gradientenfaktor und** ÿn **der negative Gradientenfaktor ist. Je kleiner** G(W) **ist, desto größer ist die Wahrscheinlichkeit positiver Proben und desto kleiner ist die Wahrscheinlichkeit negativer Proben. Durch Anpassen der Parameter** ÿp **und** ÿn **kann der Lerneffekt des Modells auf verschiedene Proben angepasst werden, um den Beitragsgrad jeder Probenklasse zur Klassifizierung des Modelllernens anzugleichen.**

Basierend auf der BLS-Verlustfunktion,

$$L(W) = (AW \ddot{y} Y)^{2} + \ddot{y}W2 \quad (24)$$

wobei *ÿ* **der Straffaktor ist. Das Optimierungsziel des BELS ist dann die EQL v2- Gradientenfunktion und die BLS-Verlustfunktion:**

$$\text{argmin} : F(W) = (AW \ddot{y} Y)^{2} + \ddot{y}W2 + \ddot{y}p[Y^{T}(AW \ddot{y} 1)]2 + \ddot{y}n[(1 \ddot{y} \mathbf{Y}^{T})AW]^{2} \quad (25)$$

Der erste Teil davon ist die Verlustfunktion, die zur Kontrolle des Klassifizierungsfehlers des Modells verwendet wird. Der zweite Teil ist der Strafterm, der hauptsächlich eine Überanpassung des Modells verhindert. Der dritte Teil ist der Gradiententerm, der zum Ausgleich des Lerngrads verschiedener Stichproben und zur Linderung der schlechten Leistung des Modells beim Erkennen von Stichproben von Minderheitsklassen verwendet wird. ÿF(W)

Die vorhandene F(W) **wird durch die partielle Ableitung von W wie folgt ausgedrückt:**

$$W = [\ddot{y}pA\,TYYTA + \ddot{y}nA^{T}(J1 \ddot{y} Y)(J2 \ddot{y} Y^{T})A + A\,TA + \ddot{y}I]^{\ddot{y}1}(\ddot{y}pATYYT\,J3 + ATYY) \quad (26)$$

wobei J **eine Matrix ist, bei der alle Elemente gleich 1 sind,** ÿp **ein positiver Gradientenfaktor ist,** ÿn **ein negativer Gradientenfaktor ist, die positiven und negativen Gradientenfaktoren verwendet werden, um den Lerngrad des Modells anhand verschiedener Stichproben zu steuern, und** *ÿ* **ein Straffaktor ist, um eine Überanpassung des Modells aufgrund eines zu großen W** zu vermeiden.

4. Experimente

In diesem Abschnitt werden die detaillierten Experimente vorgestellt. Wir analysieren detailliert die Auswirkungen von Hyperparametern auf das Modell. Wir verwenden eine Rastersuche, um die positiven und negativen Gradientenfaktoren innerhalb geeigneter Bereiche zu fixieren und suchen nach der optimalen Parameterkombination, die die Erkennungsleistung des Modells maximiert. Um die Wirksamkeit der vorgeschlagenen Methode zu bewerten, führen wir vollständige Ablationsexperimente durch. Abschließend geben wir eine vergleichende Analyse des vorgeschlagenen Modells mit anderen hochmode

4.1. CICIDS2017-Datensätze

Der vom kanadischen Institut für Cybersicherheit vorgeschlagene CICIDS2017-Datensatz [25] ist einer der wichtigsten Benchmark-Datensätze für die Bewertung von Intrusion-Detection-Modellen. Die Autoren verwendeten das Verhaltensprofilsystem, um das abstrakte Verhalten menschlicher Interaktionen auf der Grundlage verschiedener Protokolle zu analysieren, 25 abstrakte Verhaltensweisen von Benutzern zu konstruieren und natürlichen, freundlichen Hintergrundverkehr zu generieren. Der Datensatz enthält harmlose und neue gängige Angriffe und umfasst sieben Arten von Netzwerkangriffen, nämlich Denial-of-Service- Angriffe, Brute-Force-Angriffe mit Secure Shell, Botnet-Angriffe, Distributed-Denial-of-Service-Angriffe, Angriffe auf Webanwendungen, Heartbleed-Exploits und Penetrationstest-Angriffe. Die Beispielverteilung und Merkmale des Datensatzes sind in Tabelle 1 dargestellt.

Machine Translated by Google

Tabelle 1. Klassenverteilung des CICIDS2017-Datensatzes.

Kategorie	Etikett	Probe	Besonderheit
Gutartig	Gutartig	2.273.097	Zielhafen, Durchflussdauer, Gesamte Fwd|Bwd-Pakete, Gesamte Len Fwd|Bwd Pkts, Fwd|Bwd Pkt Len Max|Min|Mittelwert|Std, Min|Max Pkt.Länge, Pkt.Länge Mittelwert|Standard|Var, Durchschnittliche Pkt-Größe, Durchschnittliche Fwd|Bwd-Segment Größe, Fluss Bytes|Pkts, Fwd|Bwd Pkts, Fwd|Bwd Durchschnittliche Bulk-Rate, Fwd|Bwd Durchschnittliche Bytes|Pkts, Flow IAT Mittelwert|Standard|Max|Min, Fwd|Bwd IAT Gesamt|Mittelwert|Standard|Max|Min, Down-Up-Verhältnis, Aktiv|Leerlauf Mittelwert|Std|Max|Min, Vorwärts|Rückwärts PSH Flags, Fwd|Bwd URG-Flags, Init Win Bytes Fwd|Bwd, Fwd|Bwd Header-Länge, Act-Datenpaket Weiter, FIN|SYN|RST|PSH|ACK|URG|CWE|ECE Flags, Mindestsegmentgröße Weiterleitung, Subflow Weiterleitung|Weiterleitung Pakete|Bytes, Label
DoS/DDoS	DoS Hulk, DDoS, DoS GoldenEye, DoS Slowloris, DoS Slowhttptest, Herzbluten	380.699	
PortScan	PortScan	158.930	
Rohe Gewalt	FTP-Patator, SSH-Patator	13.835	
Web-Angriff	Web-Angriff – Brute Force, Web-Angriff – XSS, Webangriff – SQL-Injection	2180	
Botnetz	Bot	1966	
Gesamt	-	2.830.707	79

4.2. Einzelheiten zur Implementierung

Die DBELS-Modellparameter wurden wie in Tabelle 2 festgelegt, einschließlich der Trainingsparameter des DBN-Dimensionsreduktionsmodells, der BELS-Aktivierungsfunktion, der Anzahl der Knoten in jeder der Zuordnungsgruppen, die Anzahl der Knoten in jeder der Erweiterungsgruppen Gruppen und dem Strafkoeffizienten ÿ. Das DBN-Modell zur Dimensionsreduktion bestand aus ein Stapel von drei RBMs, (49,64), (64,32) und (32,16). Alle Experimente wurden durchgeführt mit eine 64-Bit Intel(R) Core (TM) i7-11700 CPU mit 32 GB RAM in der Windows 11-Umgebung. Die Modelle wurden in Python v3.9.16 unter Verwendung der Bibliothek PyTorch v2.1.0 implementiert.

Tabelle 2. Die Parameter von DBELS.

Parameter	Vortraining	Feinabstimmung	Beschreibung
Epochen	30	100	-
Lernrate	0,0001	0,00001	-
Batchgröße	64	128	-
Optimierer	SGD	Adam	-
Gibbs-Schritt	-	-	5
Zuordnung|Anzahl der Erweiterungsgruppenknoten	-	-	16
Mapping|Erweiterungsaktivierungsfunktion	-	-	Relu
ÿ	-	-	0,001

4.3. Leistungsmetriken

Zu den Bewertungskriterien für die Erkennung von Angriffen zählen vor allem Genauigkeit, Rückruf und Falsch-Positiv-Rate (FPR). **In der Formel stellt das True Positive** (TP) **die Anzahl der Proben, die richtig als positiv identifiziert wurden, das wahre Negativ** (TN) **ist die Anzahl der Proben richtig als negativ identifiziert, das falsch positive** (FP) **ist die Anzahl der negativen Proben fälschlicherweise als positiv identifiziert, und das falsche Negativ** (FN) **ist die Anzahl der positiven Proben fälschlicherweise als negativ identifiziert. Die** Genauigkeit **bezeichnet das Verhältnis der Anzahl der Vom Modell richtig klassifizierte Proben im Verhältnis zur Gesamtzahl der Proben:**

$$\text{Genauigkeit} = \frac{TP + TN}{TP + TN + FP + FN} \quad (27)$$

Machine Translated by Google

Der Recall **bezeichnet den Anteil der positiven Fälle in der Stichprobe, die richtig vorhergesagt, auch als Erkennungsrate bezeichnet:**

$$\text{Rückruf} = \frac{TP}{TP + FN} \quad (28)$$

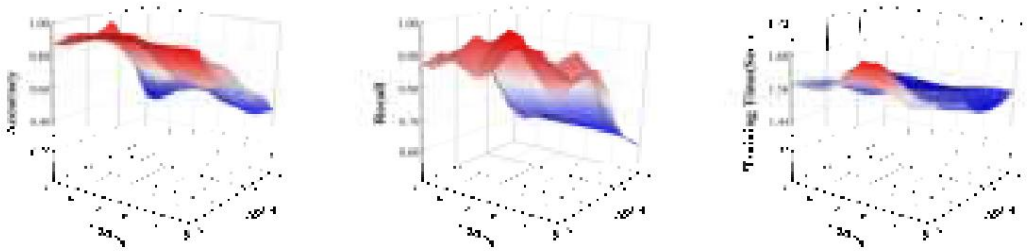
Die Falsch-Positiv-Rate ist der Anteil der wirklich negativen Stichproben, die das Modell fälschlicherweise als positiv klassifiziert:

$$FPR = \frac{FP}{FP + TN} \quad (29)$$

4.4. Analyse von Hyperparametern

4.4.1. Einfluss von ÿp **und** ÿn **auf die binäre Klassifizierung**

Für die beiden Parameter ÿp **und** ÿn **im DBELS-Modell sind ihre Auswirkungen auf die Leistung des Modells in verschiedenen Klassifizierungsaufgaben untersucht. Ihre Auswirkungen auf die Leistung Genauigkeit, Rückruf und Zeit bei der binären Klassifizierungsaufgabe sind in Abbildung 5** dargestellt. **Genauigkeit und Rückruf, die Leistung tendiert dazu, mit abnehmendem** ÿp **und** ÿn zu steigen. **Unterschiedliche Parameterkombinationen führen zu größeren Unterschieden, da das Modell empfindlicher auf die zwei Parameter in der binären Klassifizierungsaufgabe, und wenn sich die Größenordnung verbessert,** ändert **sich die Leistung des Modells drastisch. Die Experimente zeigen, dass, wenn** ÿp **etwa** 10ÿ5 **Größenordnungen und** ÿn **etwa** 10ÿ4 **Größenordnungen beträgt, kann eine höhere Genauigkeit und einen höheren Rückruf beibehalten, bei etwa 0,99 bzw. 0,98. Für die Trainingszeit des Modells mit der Änderung von** ÿp **und** ÿn **schwankt die Trainingszeit leicht über und unter 1,6 s, was darauf hinweist, dass die Modelltrainingszeit in der binären Klassifizierung Die Aufgabe reagiert weder empfindlich auf positive noch auf negative Faktoren.**



(A) **(B)** **(C)**

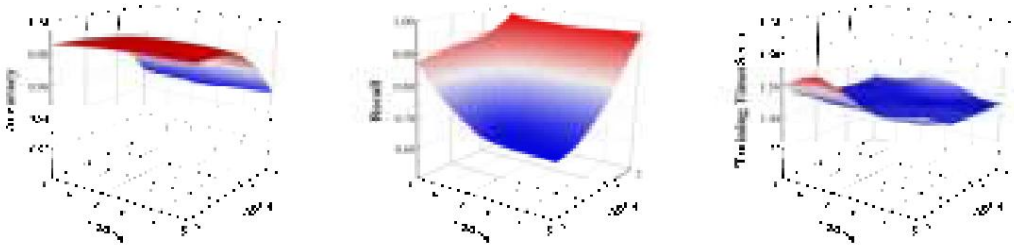
Abbildung 5. Die Auswirkung von ÿp **und** ÿn **auf die binäre Klassifizierung. Die Auswirkungen auf die Leistung: (a) die Auswirkungen auf die Genauigkeit; (b) Auswirkungen auf das Erinnerungsvermögen; (c) Auswirkungen auf die Trainingszeit.**

4.4.2. Einfluss von ÿp **und** ÿn **auf die Mehrfachklassifizierung**

Bei der Multiklassifizierungsaufgabe, seine Auswirkung auf die Genauigkeit, den Rückruf und die Zeitleistung sind in Abbildung 6 dargestellt . Was die Genauigkeit betrifft, ist das Modell nahezu unempfindlich gegenüber ÿn **und tendiert fast stabil zu sein, wenn** sich ÿn **ändert; wenn** 10ÿ3 < ÿp < 10ÿ2 **, die Genauigkeit nimmt mit** zunehmendem ÿp zu . **Für die Rückrufquote gilt, dass mit** zunehmenden ÿp **und** ÿn **auch die Rückrufquote des Modells zunimmt. Wenn** ÿp **Die Größenordnung des ÿn-Wertes beträgt** 10ÿ2 **und die Größenordnung des** ÿn- Wertes beträgt etwa 10ÿ3 **Der Rückruf steigt auf maximal 0,95. Für die Modelltrainingszeit mit der Änderung in** ÿp **und** ÿn **schwankt die Zeit des Modells zwischen 1,4 s und 1,6 s und stabilisiert sich fast, was zeigt, dass die Trainingszeit nicht empfindlich auf die positiven und negativen Faktorparameter reagiert.**

Um eine bessere Gesamtleistung zu erreichen, **liegt die Größenordnung von** ÿp bei etwa 10ÿ4 **Und Die Größenordnung von** ÿn **liegt bei etwa** 10ÿ3 **wodurch die Genauigkeit auf 0,96 stabilisiert werden kann und der Rückruf stabilisiert sich bei 0,95.**

Machine Translated by Google

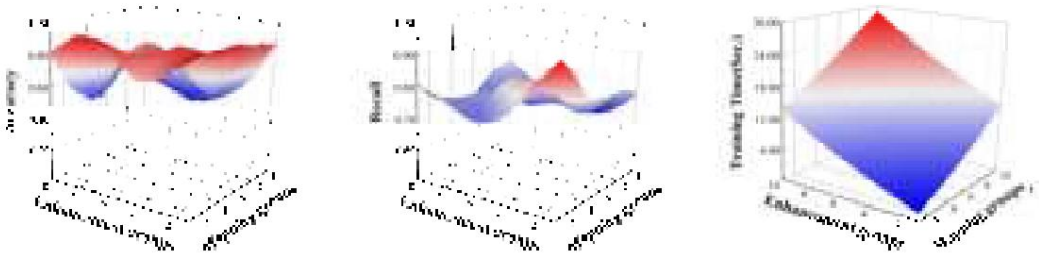


(A) **(B)** **(C)**

Abbildung 6. Die Auswirkung von ýp und ýff auf die Mehrfachklassifizierung. Die Auswirkungen auf die Leistung: (a) die Auswirkung auf die Genauigkeit; (b) die Auswirkung auf das Erinnerungsvermögen; (c) die Auswirkung auf die Trainingszeit.

4.4.3. Einfluss von Mapping- und Erweiterungsgruppen auf die binäre Klassifizierung

Die Experimente untersuchten die Auswirkungen von DBELS-Modellverbesserungsgruppen und Zuordnungsgruppen zur Leistung des Modells für verschiedene Klassifizierungsaufgaben. Die binäre Klassifizierungsaufgabe, die Auswirkungen von Erweiterungsgruppen und Zuordnungsgruppen auf Genauigkeit, Rückruf und Trainingszeit sind in Abbildung 7 dargestellt. Für Genauigkeit und Rückruf ist das Modell schwankt stärker mit der Änderung der Parameter, da das binäre Klassifizierungsaufgabenmodell empfindlicher auf diese beiden Parameter. Mit der Zunahme der Abbildungsmatrix Z **und der Verbesserungsmatrix H, ihre Eingangsmatrix** A **steigt und die Leistung der Modellerkennung ist stark abhängig von der Matrix A. Die Trainingszeit steigt mit der Zunahme von Erweiterungsgruppen und Abbildungsgruppen, da die Eingangsmatrix** A **mit der Erhöhung der Parameter. Aus den experimentellen Ergebnissen ist bekannt, dass bei der Verstärkung Gruppen und Zuordnungsgruppen werden als 1 ausgewählt, dies kann sowohl eine kürzere Trainingszeit als auch höhere Genauigkeit und Rückruf.**



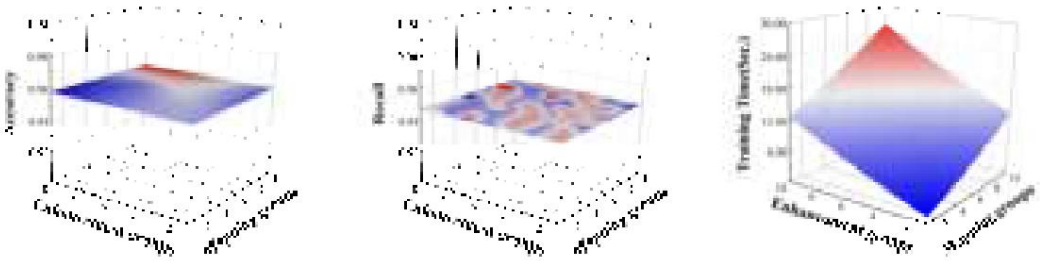
(A) **(B)** **(C)**

Abbildung 7. Die Wirkung von Erweiterungsgruppen und Zuordnungsgruppen auf die binäre Klassifizierung. Die Auswirkungen auf die Leistung: (a) die Auswirkung auf die Genauigkeit; (b) die Auswirkung auf das Erinnerungsvermögen; (c) die Auswirkung auf die Trainingszeit.

4.4.4. Auswirkungen von Mapping- und Erweiterungsgruppen auf die Mehrfachklassifizierung

Die Auswirkungen auf Genauigkeit und Rückruf bei der Mehrfachklassifizierungsaufgabe werden in Abbildung 8 dargestellt. Mit der Zunahme von Erweiterungsgruppen und Zuordnungsgruppen werden Genauigkeit und Rückruf tendieren fast zur Stabilisierung, was darauf hindeutet, dass für die Multiklassifizierungsaufgabe diese beiden Metriken von Das DBELS-Modell ist nicht sensitiv gegenüber den Mapping- und Enhancement-Gruppen. Zeit zeigt einen Aufwärtstrend im Zeitverbrauch sowohl mit der Zunahme der Kartierung und Erweiterungsgruppen. Der Grund dafür ist, dass die Zunahme an Mappinggruppen und Erweiterungsgruppen vergrößern die Abbildungsmatrix Z **und die Erweiterungsgruppen Matrix H, was zu einer Erhöhung der Matrix A führt, was die Belastung und den Verbrauch erhöht des Modells. Es kann gefolgert werden, dass seit der Zunahme der Enhancement-Gruppen und Mapping-Gruppen haben weniger Einfluss auf die Erkennungsleistung des Modells und erhöhen der Zeitaufwand bei der Auswahl von Erweiterungsgruppen und Mappinggruppen als 1, die Genauigkeit und Rückruf halten eine höhere Leistung und den Zeitverbrauch wird minimiert.**

Machine Translated by Google



(A) **(B)** **(C)**

Abbildung 8. Die Auswirkung von Erweiterungsgruppen und Zuordnungsgruppen auf die Mehrfachklassifizierung. Die Auswirkungen auf die Leistung: (a) die Auswirkung auf die Genauigkeit; (b) die Auswirkung auf das Erinnerungsvermögen; (c) die Auswirkung auf die Trainingszeit.

4.5. Ablationsstudien

Um die Wirksamkeit der vorgeschlagenen Methode zu bewerten, haben wir Ablationsexperimente auf der Grundlage dreidimensionaler Reduktionsmodelle und zweier Klassifizierungs Die dreidimensionalen Reduktionsmodelle sind traditionelle PCA, traditionelle DBN und DeBN mit der Einführung von EQL v2. Die Klassifizierungsmethoden sind traditionelle BLS und unsere verbesserten BELS. Wir evaluierten die Detektionsleistung von insgesamt acht Methoden unter einer binären Klassifizierungsaufgabe und einer Multiklassifizierungsaufgabe basierend auf der CICIDS2017-Datensatz. Alle Experimente wurden in derselben Umgebung durchgeführt mit identische Formen der Datenverarbeitung, und die Erkennungsleistungen verschiedener Modelle sind in den Tabellen **3 und** 4 **dargestellt . Die Erkennungsrate jeder Probe unter verschiedenen Klassifizierungen** Aufgaben sind in den Tabellen **5 und 6 dargestellt .**

Tabelle 3. Die Leistung verschiedener Methoden bei der binären Klassifizierung.

Methode	Algorithmus					Bewertungsmetrik		
	BLS	BELS	PCA	DBN	DeBN	Genauigkeit	Abrufen	Mal)
BLS	ÿ					0,88105	0,63062	1,83126
PBLS	ÿ		ÿ			0,86780	0,59044	0,75575
DBLS	ÿ			ÿ		0,91531	0,73711	0,60949
DeBLS	ÿ				ÿ	0,29939	0,58103	0,60531
BELS		ÿ				0,96670	0,92483	4,86849
PBELS		ÿ	ÿ			0,92279	0,85250	1,64884
DBELS		ÿ		ÿ		0,99240	0,98579	1,63914
DeBELS		ÿ			ÿ	0,98896	0,97516	1,57186

Das ÿ in der Tabelle zeigt an, dass die Methode den Algorithmus enthält.

Tabelle 4. Die Leistung verschiedener Methoden bei der Mehrfachklassifizierung.

Methode	Algorithmus					Bewertungsmetrik		
	BLS	BELS	PCA	DBN	DeBN	Genauigkeit	Abrufen	Mal)
BLS	ÿ					0,97450	0,47382	1,74388
PBLS	ÿ		ÿ			0,93755	0,29756	0,51006
DBLS	ÿ			ÿ		0,99286	0,65797	0,51341
DeBLS	ÿ				ÿ	0,98936	0,67712	0,50913
BELS		ÿ				0,97060	0,52527	5,00825
PBELS		ÿ	ÿ			0,93848	0,47024	1,77999
DBELS		ÿ		ÿ		0,98719	0,80346	1,49983
DeBELS		ÿ			ÿ	0,95790	0,94741	1,63619

Das ÿ in der Tabelle zeigt an, dass die Methode den Algorithmus enthält.

Machine Translated by Google

Tabelle 5. Der Rückruf jeder Probe bei binärer Klassifizierung.

Kategorie	Methode							
	BLS	PBLS	DBLS	DeBLS	BELS	PBELS	DBELS	DeBELS
Gutartig	0,99983	0,99934	0,99983	0,16582	0,98655	0,95613	0,99554	0,99551
Attacke	0,26141	0,18154	0,47439	0,99625	0,86312	0,74888	0,97604	0,95480

Tabelle 6. Der Rückruf jeder Probe bei Mehrfachklassifizierung.

Kategorie	Methode							
	BLS	PBLS	DBLS	DeBLS	BELS	PBELS	DBELS	DeBELS
Gutartig	0,99094	0,99252	0,99787	0,99665	0,99057	0,95042	0,98886	0,95903
Botnetz	0	0	0	0,13660	0	0	0	0,94845
ARES	0	0	0,98480	0,98655	1	0	0,99766	0,98830
Rohe Gewalt	0,91515	0,79286	0,97416	0,95507	0,87920	0,89234	0,98244	0,94544
DoS/DDoS	0,93683	0	0,99101	0,98787	0,85313	0,97871	0,99093	0,98709
Port-Scan	0	0	0	0	0	0	0,86085	0,85613
Web-Angriff								

4.5.1. Leistungsanalyse bei binärer Klassifizierung

Bei der binären Klassifizierungsaufgabe ist die traditionelle BLS-Struktur bei der Erkennung des CICIDS2017-Datensatzes weniger effektiv, hauptsächlich aufgrund der hochdimensionalen redundanten Informat im Datensatz, was die Klassifizierungsleistung beeinträchtigt, was zu einer geringeren Genauigkeit führt und Rückruf, insbesondere die mangelnde Aufmerksamkeit für die Minderheitenklassen-Stichproben in der unausgewogenen Datensatz. Nach der Einführung von EQL v2 verbesserten sich die Genauigkeit und der Rückruf des BELS-Modells erheblich , was darauf hindeutet, dass EQL v2 die Fähigkeit verbessert, Minderheitenklassen zu er Proben im Umgang mit dem Ungleichgewichtsproblem. Das Modell Deep Belief Network Broad Learning System (DBLS) reduziert die Datenredundanz durch die DBN-Dimensionalität Reduktion, die die Genauigkeit und den Rückruf verbessert, die um 0,034 verbessert wurden und 0,106 bzw. Das DeBLS-Modell konzentriert sich zu sehr auf Stichproben aus Minderheitenklassen während Reduzierung der DeBN-Dimensionalität, was zu einer Verringerung der Erkennungsleistung für Stichproben der Mehrheitsklasse.

DBELS hat die höchste Erkennungsleistung bei der binären Klassifizierungsaufgabe und Die Einführung von EQL v2 auf DBLS verbessert die Genauigkeit und den Rückruf um 0,077 bzw. 0,249, was darauf hinweist, dass die Kombination von DBN-Dimensionalität Reduktion und EQL v2 verbessern die Modellleistung erheblich. Obwohl die DeBELS Modell führt EQL v2 in den DeBN-Datensatz zur Dimensionsreduktion ein und verbessert die Modellleistung durch Anpassen der Parameter des positiven und negativen Gradienten Faktoren, die Genauigkeit und der Rückruf sanken leicht um 0,004 bzw. 0,01 im Vergleich mit DBELS, was die negativen Auswirkungen des Überlernens abnormaler Proben auf Klassifizierungsleistung. Für die Hauptkomponentenanalyse breites Lernsystem (PBLS)-Modell, nach der PCA-Dimensionsreduktion verringerten sich Genauigkeit und Rückruf um 0,014 bzw. 0,040 im Vergleich zum BLS, hauptsächlich aufgrund zusätzlicher Informationen Verlust bei der PCA-Dimensionsreduktion und die Unfähigkeit, mit unausgeglichenen Datensätze effizient. Das Hauptkomponentenanalyse-Lernsystem mit breiter Entzerrung (PBELS) Modell hatte eine Genauigkeit von 0,923 und einen Rückruf von 0,853, verglichen mit PBLS, mit eine Verbesserung von 0,055 und 0,262, was darauf hindeutet, dass EQL v2 noch deutlich verbessern kann die Erkennungsleistung von Minderheitenklassenproben selbst bei PCA-Dimensionalität Reduktion ist unwirksam. Allerdings reduzierte die PBELS Genauigkeit und Rückruf um 0,070 und 0,106 im Vergleich zu DBELS, was weiter belegt, dass die DBN-Dimensionalität Die Reduzierung übertrifft die PCA-Dimensionsreduzierung bei binären Klassifizierungsaufgaben.

Machine Translated by Google

4.5.2. Leistungsanalyse bei Mehrfachklassifizierung Bei

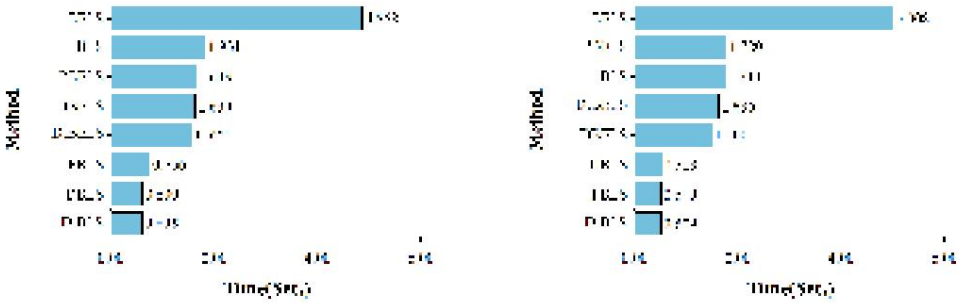
der Mehrfachklassifizierungsaufgabe schnitt das herkömmliche BLS mit einer Genauigkeit von 0,975 und einem Rückruf von nur 0,474 schlecht ab, hauptsächlich aufgrund der starken Unausgewogenheit des Datensatzes, die die Klassifizierungsleistung beeinträchtigt. Das BELS-Modell hatte nach Einführung von EQL v2 einen verbesserten Rückruf von 0,051, hauptsächlich weil EQL v2 das Problem der Datenunausgewogenheit verbessert, wodurch das Modell die Gewichtung der Angriffsproben für Minderheitsklassen erhöhen und somit die Erkennungsrate verbessern kann. Das DBLS-Modell hatte nach der Dimensionsreduzierung durch DBN eine verbesserte Genauigkeit von 0,018 und einen verbesserten Rückruf von 0,184, was darauf zurückzuführen ist, dass die Dimensionsreduzierung durch DBN die Datenredundanz verringert, die negativen Auswirkungen der Rohdaten mildert und die Erkennungsleistung des Modells verbessert. Das DeBLS-Modell verbesserte die Trefferquote im Vergleich zu DBLS um 0,02, was darauf zurückzuführen ist, dass DeBN mehr über die Proben der Minderheitsklasse lernt, Redundanz zwischen verschiedenen Proben eliminiert und gleichzeitig die wichtigsten Merkmale der Proben der Minderheitsklasse beibehält, wodurch die Klassifizierungs- und Erkennungsleistung verbessert wird. Das DBELS-Modell verbesserte sich bei der DBN-Dimensionsreduzierung um 0,018, und mit der Einführung von EQL v2 verbesserten sich die Genauigkeit um 0,017 und die Trefferquote um 0,278; im Vergleich zu DBLS verbesserte sich die Trefferquote von 0,658 auf 0,803, was da

Die Genauigkeit des DeBELS-Modells von 0,957 und der Rückruf von 0,947 stellten die beste Leistung bei der Mehrfachklassifizierungsaufgabe dar und der Rückruf war im Vergleich zu DeBLS signifikant um 0,27 verbessert, was daran liegt, dass das verbesserte BLS-Modell besser mit unausgewogenen Datensätzen umgehen kann und die Fähigkeit zum Erkennen von Angriffen von Minderheitsklassen verbessert. Der Rückruf von DeBELS verbesserte sich im Vergleich zu DBELS um 0,144 und die Gesamtklassifizierungsleistung verbesserte sich signifikant. Die Genauigkeit des PBLS-Modells verringerte sich um 0,037 und der Rückruf verringerte sich um 0,176 nach der PCA-Dimensionalitätsreduzierung, da die PCA-Dimensionalitätsreduzierung zu einem größeren Informationsverlust führt, der den Mehrfachklassifizierungseffekt beeinflusst . Die Genauigkeit des PBELS-Modells betrug 0,938 und der Rückruf 0,470; Im Vergleich zu PBLS verbesserten sich Genauigkeit und Rückruf um 0,001 bzw. 0,173, was darauf hinweist, dass die Einführung von EQL v2 die Erkennungsrate von Minderheitenklassenproben verbessert. Im Vergleich zu DeBELS verringerten sich Genauigkeit und Rückruf von PBELS jedoch um 0,019 bzw. 0,477, was darauf hinweist, dass die Dimensionsreduktion von DBN in Bezug auf die Erkennungsleistung viel besser ist als das Dimensionsr

4.5.3. Zeit-Kosten-Analyse

Der Vergleich der Trainingszeiten für die verschiedenen Modelle ist in Abbildung 9 dargestellt. Bei der binären Klassifizierung und Mehrfachklassifizierung betrug die Trainingszeit des herkömmlichen BLS etwa 1,8 s, was den Vorteil flacher Netzwerke mit geringerem Zeitaufwand bei der Modellklassifizierung widerspiegelt. Die Trainingszeit von BELS betrug etwa 5 s, und die Zeiterhöhung war darauf zurückzuführen, dass die Einführung von EQL v2 die zusätzliche Durchführung mehrerer Matrixoperationen erfordert . Die Trainingszeit des DBLS-Modells betrug etwa 0,5 s, was eine erhebliche Verkürzung im Vergleich zum BLS-Modell darstellt und darauf hindeutet, dass hochdimensionale Daten einen erheblichen negativen Effekt auf die Modelltrainingszeit haben und die verringerte Dimensionalität des Datensatzes die Trainingszeit auf 25 % der ursprünglichen Zeit reduziert. Die Trainingszeit von DeBLS betrug ebenfalls etwa 0,5 s, was darauf hindeutet, dass die Einführung von EQL v2 nach der Dimensionsreduzierung von DBN fast keine Auswirkungen auf die Trainingszeit hat. Die Trainingszeit für DBELS betrug etwa 1,5 s, was im Vergleich zu BELS erheblich reduziert ist, was **der Dimensionsreduzierung des DBN-Modells für hochdimensionale Daten zu verdanken ist .** Im Vergleich zu DBLS **gab es jedoch einen Anstieg aufgrund einiger zusätzlicher Berechnungen, die das verbesserte BELS-Modell erforderte. Die Trainingszeit von DeBELS betrug etwa 1,6 s, was fast der von DBELS entspricht, was darauf hindeutet, dass die Einführung des niedrigdimensionalen Datensatzes mit EQL v2 keinen signifikanten Einfluss auf die Trainingszeit hat und der Zeitaufwand aufgrund einer geringen Anzahl zusätzlicher Berechnungen der positiven und negativen Gradientenmatrizen im Verglei**

Machine Translated by Google



(A) **(B)**

Abbildung 9. Vergleich der Trainingszeit für verschiedene Methoden: (a) binäre Klassifizierung; (b) Mehrfachklassifizierung.

Die Ergebnisse zeigen, dass die hybriden DBN- und BLS-basierten Modelle DBELS und DeBELS im Vergleich zu den traditionellen BLS-Modellen einen geringen Zeitaufwand aufweisen, was den Vorteilen der flachen BLS-Netzwerke entspricht, die rechnerisch schnell und einfach zu trainieren sind. Obwohl das PCA-Modell zur Dimensionsreduzierung ebenfalls einen Teil der Trainingszeit verkürzt , verschlechtert sich seine Erkennungsleistung erheblich, was wiederum die Vorteile des DBN-basierten Modells zur Dimensionsreduzierung bestätigt. Das vorgeschlagene Modell ist mit dem traditionellen Deep-Learning-Modell vergleichbar, das einen erheblichen Vorteil beim Zeitaufwand aufweist , was dazu beiträgt, in **praktischen Anwendungen eine gewisse Menge an Rechenressourcen einzusparen .**

4.5.4. Ergebnisanalyse für den Rückruf jeder Probe bei binärer Klassifizierung

Die Erkennungsrate von gutartigen und Angriffsproben bei der binären Klassifizierungsaufgabe ist in Tabelle 5 dargestellt. Das traditionelle BLS zeigte bei der Erkennung von Angriffsproben eine schlechte Leistung mit einem Rückruf von nur 0,261, was auf das Überlernen der Mehrheitsklassenproben bei der Verarbeitung unausgewogener Datensätze zurückzuführen ist, was zu einer niedrigen Erkennungsrate für Minderheitsklassenproben führt . Mit der Einführung des BELS-Modells mit EQL v2 verbesserte sich die Erkennungsrate von Angriffen deutlich auf 0,863, was die Identifizierung von Minderheitsklassenproben erheblich verbessert. Das DBLS-Modell mit DBN-Dimensionsreduzierung verbesserte die Erkennungsrate von Angriffen um 0,213, was darauf hindeutet, dass die DBN-Dimensionsreduzierung die Datenredundanz verringert und sich positiv auf die nachfolgende Modellklassifizierung auswirkt. Im DeBLS- Modell betrug die Erkennungsrate von Angriffen bis zu 0,996, die Erkennungsrate von gutartigen Proben jedoch nur 0,165, was zeigt, dass sich die DeBN-Dimensionsreduzierung übermäßig auf Minderheitsklassen von Proben konzentriert, was zu einer Verringerung der Erkennungsleistung der Mehrheitsklassen von Proben führt.

Das DBELS-Modell hatte nach der Dimensionsreduzierung durch die Einführung von EQL v2 eine Erkennungsrate von 0,996 für „Gutartig“ und 0,976 für „Angriff“, was die Erkennungsrate der beiden Probenklassen im Vergleich zum BELS-Modell um 0,009 bzw. 0,1

Da die Dimensionsreduzierung von DBN redundante Merkmale verringert und die Klassifizierungsleistung des Modells verbessert, verbesserte DBELS die Erkennungsrate von Attack im Vergleich zu DBLS erheblich von 0,474 auf 0,955, was darauf zurückzuführen ist, dass EQL v2 die Probengewichte ausgleicht, sodass das Modell sich **ausgewogener auf jede Probenklasse konzentriert .** Die Erkennungsraten von Benign und Attack für das **DeBELS-Modell betrugen 0,995 bzw. 0,955. Im Vergleich mit DeBLS war die Erkennungsrate von Benign wesentlich höher und die Erkennungsrate von Attack etwas niedriger, aber EQL v2 verbesserte die Erkennungsleistung, indem das Modell sich ausgewogener auf jede Probenklasse konzentrierte. Im Vergleich zu DBELS blieb die Erkennungsrate von Benign für DeBELS nahezu unverändert und die Erkennungsrate von Attack ging leicht zurück, was auf das Überlernen von Proben der Minderheitsklasse durch DeBN zurückzuführen ist, aber durch die Anpassung der positiven und negativen Gradientenfaktoren funktionierte die Klassifizierung weiterhin gut.**

Machine Translated by Google

Die Angriffserkennungsrate für das PBLS-Modell betrug nach der PCA-Dimensionalitätsreduzierung nur 0,182, was einen Rückgang gegenüber dem BLS-Modell von 0,080 darstellt. Die PCA-Dimensionalitätsreduzierung führt zu Informationsverlust, der sich negativ auf die Klassifizierungserkennungswirkung von Minderheitsklassenproben auswirkt. Die Angriffserkennungsrate des PBELS- Modells betrug 0,749, was eine Verbesserung um 0,567 gegenüber PBLS darstellt. Durch die Einführung von EQL v2 konzentriert sich das Modell ausgewogener auf jede Probenklasse, was die Erkennungsleistung der Minderheitsklassenproben verbessert. Trotz der Verbesserung des PBELS-Modells war seine Angriffserkennungsrate noch immer um 0,227 niedriger als die von DBELS, was die Überlegenheit des DBN-Modells mit Dimensionsreduzierung hinsichtlich der Erkennungsleistung weiter beweist.

4.5.5. Ergebnisanalyse für den Rückruf jeder Probe bei Mehrfachklassifizierung

Die Erkennungsraten der verschiedenen Methoden für jeden Probentyp bei der Mehrfachklassifizierungsaufgabe werden in Tabelle 6 dargestellt. Das herkömmliche BLS funktionierte beim Umgang mit dem unausgewogenen CICIDS2017-Datensatz nicht gut, und obwohl es bei einer größeren Anzahl von Proben (z. B. gutartig, DoS/DDoS, Port-Scan) eine bessere Erkennungsleistung zeigte, war es beim Erkennen von Proben einiger Klassen (z. B. Botnet ARES, Brute Force, Web-Angriff) schlecht. Das liegt daran, dass das BLS-Modell beim Umgang mit unausgewogenen Datensätzen die Beiträge der verschiedenen Probenklassen nicht ausgleichen kann, was zu einer sehr niedrigen Erkennungsrate für Proben von Minderheitsklassen führt. Die Einführung des BELS-Modells mit EQL v2 verbesserte die Erkennungsrate von Brute-Force-Angriffen erheblich, was wiederum die Angriffsidentifizierung verb Botnet ARES und Web Attack waren jedoch immer noch nicht erkennbar, da die Merkmale dieser Minderheitsklassenproben anderen Proben sehr ähnlich sind, die das Modell nicht unterscheiden kann, was zu falschen Positivergebnissen führt. Das DBLS-Modell eliminiert bestimmte Datenredundanzen nach der Dimensionsreduzierung durch DBN, was die Erkennungsleistung verbessert . Die Erkennungsrate von Brute Force wurde auf 0,985 verbessert, Botnet ARES und Web Attack wurden jedoch immer noch nicht erkannt. Dies zeigt, dass die Dimensionsreduzierung allein das Problem unausgewogener Datensätze nicht lösen kann, insbesondere bei Klassen mit einer sehr geringen Anzahl von Proben.

Das DeBLS-Modell verbesserte die Erkennungsleistung weiter und die meisten Angriffsarten wurden erkannt, einschließlich Botnet ARES. Dies liegt daran, dass die Dimensionsreduktion von DeBN mehr über die Minderheitsklassenproben lernt und ihre Hauptmerkmale bewahrt, aber Web Attack war immer noch nicht erkennbar, was eine weitere Konzentration auf die Minderheitsklassenproben erfordert . Das DBELS-Modell zeigte eine signifikante Verbesserung der Erkennungsrate der meisten Angriffsarten im Vergleich zu BELS und DBLS, insbesondere Web Attack, von 0 auf 0,856.
Dies liegt daran, dass die DBN-Dimensionsreduktion die Datenredundanz verringert, während EQL v2 dafür sorgt, dass sich das Modell besser auf die Stichproben der Minderheitsklasse konzentriert. Botnet ARES wurde jedoch weiterhin nicht erkannt, was darauf hindeutet, dass die DBN-Dimensionsreduktion nicht ausreicht, um das Problem zu lösen. Das DeBELS-Modell behielt eine hohe Erkennungsrate für alle Angriffsarten bei . Im Vergleich zu DeBELS verbesserte sich die Erkennungsrate von Botnet ARES von 0,137 auf 0,948 und die von Web Attack von 0 auf 0,856. Dies liegt daran, dass die Einführung von EQL v2 die Fähigkeit verbessert, die Stichproben der Minderheitsklasse besser zu erkennen und zu erlernen, was wiederum die Erkennungsrate deutlich verbessert. DeBELS verbesserte die Erkennungsrate von Botnet ARES im Vergleich zu DBELS weiter, da die DeBN-Dimensionsreduktion die wichtigsten Merkmale der Stichproben der Minderheitsklasse in einer feineren Granularität bewahrt und das Problem des Datenungleichgewichts mildert.

Das PBLS-Modell mit PCA-Dimensionsreduzierung erkannte nur die höchste Anzahl von DoS/DDoS-Angriffen, während Port Scan, Botnet ARES, Brute Force, Web Attack und andere Angriffe nicht erkannt wurden. Im Vergleich zum BLS-Modell konnte Port Scan aufgrund des durch die PCA-Dimensionsreduzierung verursachten Informationsverlusts nicht erkannt werden, was den Multiklassifizierungseffekt beeinträchtigt. Nach der Einführung von EQL v2 in das PBELS- Modell verbesserte sich die Erkennungsrate von Port Scan auf 0,978, aber die anderen wenigen Angriffsklassen (z. B. Botnet ARES, Brute Force und Web Attack) blieben weiterhin unentdeckt. Dies deutet darauf hin, dass die PCA-Dimensionsreduzierung zum Verlust wichtiger Funktionen führt. Im Vergleich zu

Machine Translated by Google

DeBELS, die Erkennungsleistung von PBELS war schlechter, was darauf hindeutet, dass die DBN-basierte Dimensionsreduktion die traditionelle PCA-Dimensionsreduktion in Bezug auf die Erkennungsleistung übertrifft.

4.6. Analyse von ROC-AUC

Abbildung 10 zeigt die Receiver-Operating-Characteristic-Kurven (ROC) für die binäre DBELS-Klassifizierung. Die durchschnittliche ROC-Kurve insgesamt weist eine Fläche unter der Kurve (AUC) von 0,9949 auf, was darauf hinweist, dass der DBELS-Klassifikator im gesamten Datensatz eine hervorragende Leistung erbracht hat und nahezu ein perfekter Klassifikator ist. Zweitens sind die ROC-Kurven für die beiden Kategorien (Angriff und Gutartig) zu sehen, jeweils mit einer AUC von 0,9916, was darauf hinweist, dass der Klassifikator auch bei individueller Betrachtung jeder Kategorie eine hohe Leistung erbracht hat, insbesondere bei der Verarbeitung einer Minderheit von Proben. Die AUC nahe 1 bedeutet, dass der Klassifikator bei seinen Vorhersagen für die Kategorie eine sehr hohe True-Positive-Rate (TPR) aufweist und gleichzeitig eine relativ niedrige FPR beibehält. Dies bedeutet, so viele True-Positive -Proben wie möglich zu identifizieren und gleichzeitig die Fehldiagnoserate so gering wie möglich zu halten. Insgesamt zeigte der DBELS-Klassifikator im gesamten Ungleichgewichtsdatensatz sowohl für Stichproben der Minderheits- als

Die ROC-Kurve der DeBELS-Mehrfachklassifizierung ist in Abbildung 11 dargestellt. Die durchschnittliche ROC-Kurve insgesamt weist einen AUC von 0,9760 auf, was darauf hinweist, dass der Klassifikator im gesamten Datensatz gut abgeschnitten hat. Die Erkennungen Benign, Botnet ARES, Brute Force, DoS/DDoS und Port Scan zeigten gute Ergebnisse; ihre AUC-Bereiche lagen alle über 0,9700, obwohl Web Attack den niedrigsten AUC-Bereich von 0,9100 aufwies, aber dieser ist bereits besser Insgesamt kann das Modell nach der Einführung der positiven und negativen Gradientenfaktoren von EQL v2 durch Anpassen der Parametergewichte den Minderheitsklassen von Proben angesichts unausgewogener Datensätze mehr Aufmerksamkeit schenken und so die Erkennungsleistung

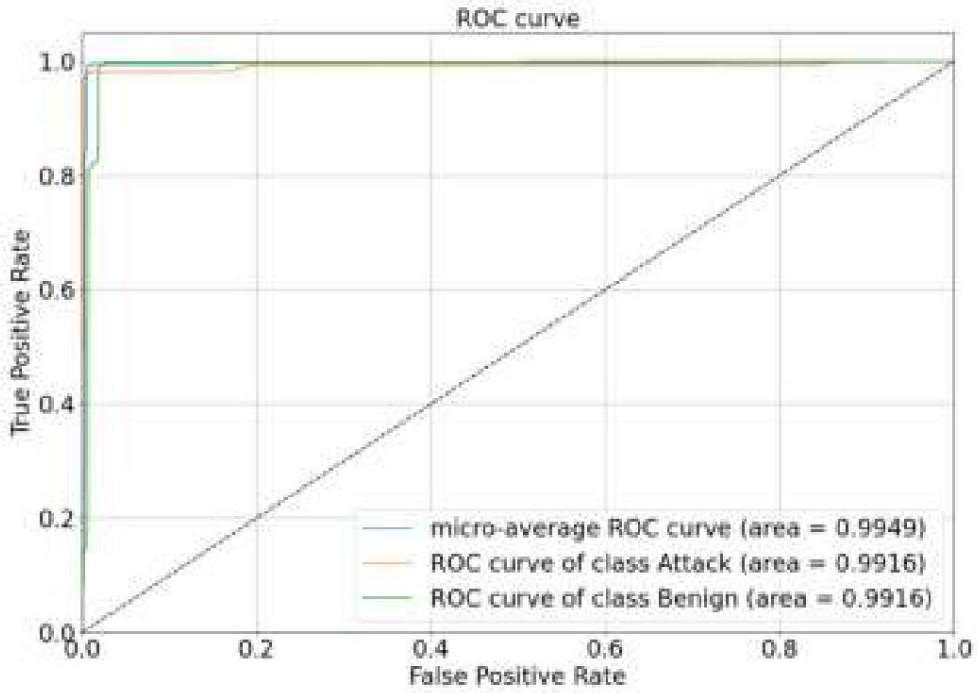


Abbildung 10. Der ROC von DBELS bei binärer Klassifizierung.

Machine Translated by Google

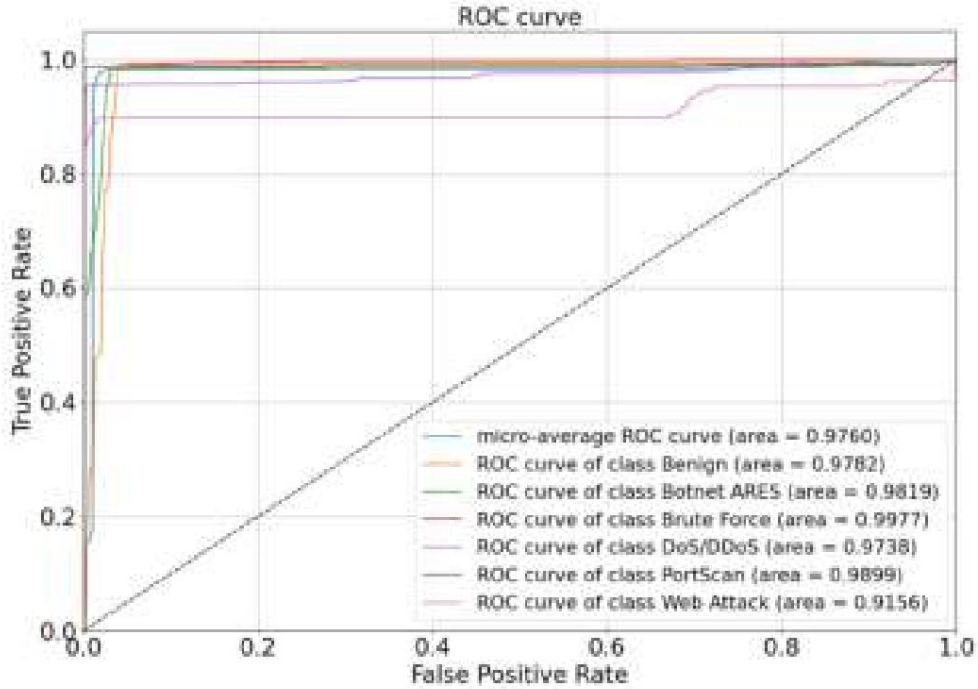


Abbildung 11. Der ROC von DeBELS bei Mehrfachklassifizierung.

4.7 Vergleich mit modernsten Methoden **4.7.1.**

Binäre Klassifizierung Die

Falschalarmrate und der Zeitaufwand aller Algorithmen wurden bei der binären Klassifizierungsaufgabe verglichen, wie in Abbildung 12 dargestellt . Bei FPR hatte das vorgeschlagene Modell eine niedrigere Falschalarmrate von 0,014, was besser war als die anderen Modelle. Dies deutet darauf hin, dass das vorgeschlagene Modell im Vergleich zu anderen Methoden normale Situationen seltener fälschlicherweise als abnormal kennzeichnet und normale Situationen häufiger richtig identifiziert und klassifiziert. Was die Modelltrainingszeit betrifft, benötigte das vorgeschlagene Modell bei der binären Klassifizierungsaufgabe weniger Trainingszeit und Erkennungszeit, mit einem Zeitaufwand von etwa 1 s. Der Zeitaufwand des Modells ist viel niedriger als der des Deep-Learning-Modells und hat einen Vorteil gegenüber anderen maschinellen Lernalgorithmen, da es den Zeitaufwand erheblich reduzieren und ein schnelleres

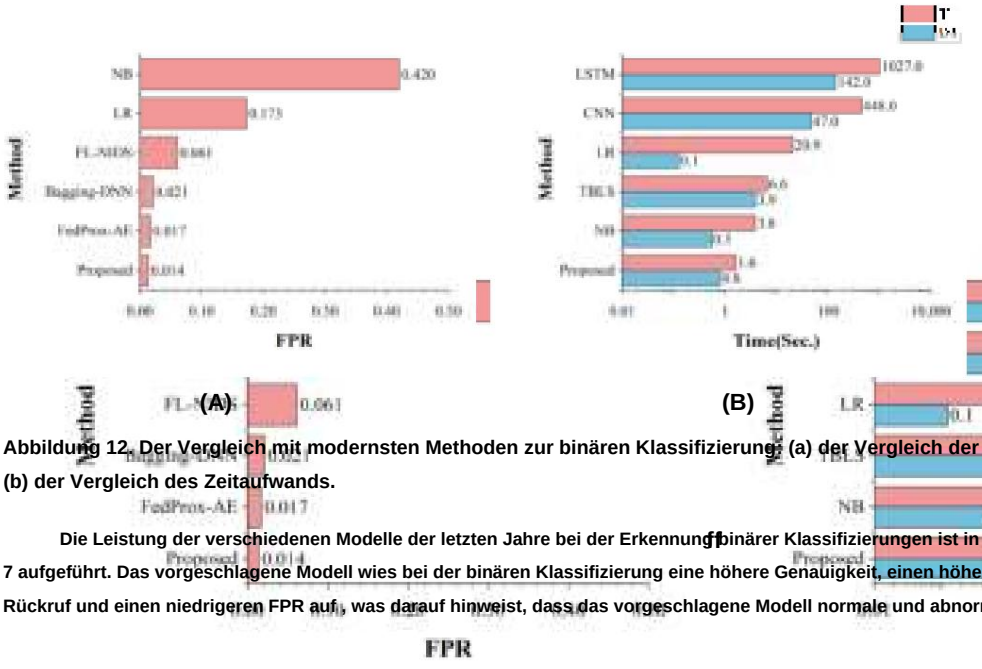


Abbildung 12. Der Vergleich mit modernsten Methoden zur binären Klassifizierung: (a) der Vergleich der FPR; (b) der Vergleich des Zeitaufwands.

Die Leistung der verschiedenen Modelle der letzten Jahre bei der Erkennung binärer Klassifizierungen ist in Tabelle 7 aufgeführt. Das vorgeschlagene Modell wies bei der binären Klassifizierung eine höhere Genauigkeit, einen höheren Rückruf und einen niedrigeren FPR auf, was darauf hinweist, dass das vorgeschlagene Modell normale und abnormale Klassifizierung

Machine Translated by Google

Verhaltensweisen gut. Im Vergleich zu traditionellen maschinellen Lernmodellen ist das vorgeschlagene Modell übertrifft die logistische Regression (LR) und Naive Bayes (NB) in Genauigkeit, Rückruf und FPR. Im Vergleich zu anderen BLS-basierten Modellen sind Genauigkeit und Rückruf besser als bei TBLS. Im Vergleich zu Deep-Learning-Modellen übertrifft das vorgeschlagene Modell die meisten Modelle. und die Leistungen der Fusion von statistischem Deep Neural Network (FS-DNNs) und Multi- Objective Evolutionary Convolutional Neural Network (MECNN) sind etwas höher als das vorgeschlagene Modell. Da jedoch beide Modelle auf Deep-Learning-Modellen basieren, Modellgröße und Trainingszeitaufwand sind viel höher als beim BLS-basierten Modell. Das Bagging Ensemble Learning Deep Neural Network (Bagging-DNN) Modell, seine Genauigkeit ist niedriger als das vorgeschlagene Modell, und sein Rückruf ist etwas höher als das vorgeschlagene Modell, aber zusätzlich zu der höheren räumlichen und zeitlichen Komplexität, die das Deep-Learning-Modell mit sich bringt, Die Falschalarmrate ist ebenfalls höher als beim vorgeschlagenen Modell. Zusammenfassend lässt sich sagen, dass das vorgeschlagene Das Modell weist eine bessere Erkennungsleistung und eine niedrigere Falschalarmrate bei der binären Klassifizierung auf und hat klare Vorteile gegenüber den anderen Modellen.

Tabelle 7. Der Vergleich mit modernsten Methoden der binären Klassifizierung.

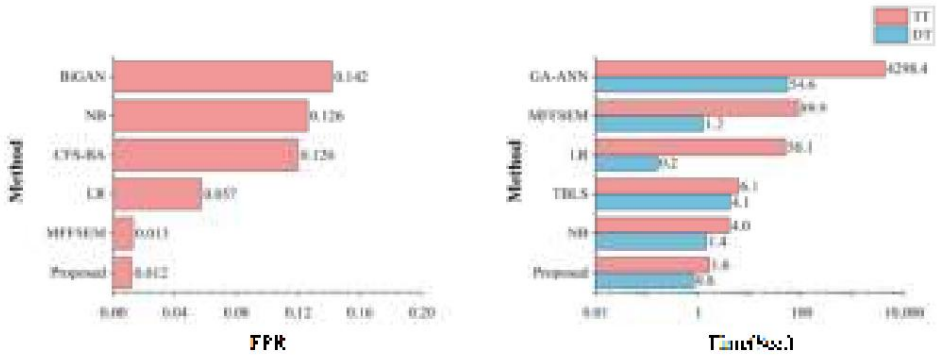
Methode	Genauigkeit	Abrufen	FPR
Unsere Studie	0,992	0,986	0,014
LR	0,934	0,827	0,173
Achtung	0,307	0,580	0,420
DCAE	0,925	0,925	-
TBLS(W) [18]	0,982	0,975	-
DNN [18]	0,868	-	-
CNN [18]	0,844	-	-
LSTM [18]	0,365	-	-
2D-CNN [26]	0,980	-	-
CNN-LSTM [27]	0,930	0,768	-
FS-DNN [28]	0,998	0,999	0,012
MECNN [29]	0,998	0,998	-
Bagging-DNN [30]	0,987	0,999	0,021
FL-NIDS [31]	0,943	0,947	0,061
FedProx-AE [32]	0,935	-	0,017

4.7.2. Mehrfachklassifizierung

Die Falsch-Positiv-Rate und der Zeitaufwand aller anderen Algorithmen bei der Multi-Klassifizierung Aufgabe sind in Abbildung 13 dargestellt. Die FPR-Grafiken zeigen, dass das vorgeschlagene Modell nur eine 0,012 Falsch-Positiv-Rate, was niedriger ist als bei den anderen Modellen. Das bedeutet, dass das Modell können unterschiedliche Angriffskategorien besser erkennen und seltener falsch klassifizieren während der Multi-Klassifizierungsaufgabe und Durchführung der Aufgabe der Erkennung und Identifizierung von Angriffsproben besser. Das Modell benötigte weniger Trainingszeit und Erkennungszeit im Vergleich zu anderen Modelle, die Vorteile bei der Handhabung großer Datenmengen bieten und ihre Eignung für reale Datensatzumgebungen.

Die Leistung der Multiklassifizierungserkennung verschiedener Modelle in den letzten Jahren ist in Tabelle 8 angegeben. Das vorgeschlagene Modell hat eine höhere Genauigkeit und einen höheren Rückruf und eine niedrigere FPR unter Multi-Klassifikationsaufgabe, was darauf hindeutet, dass das vorgeschlagene Modell sowohl normale und abnormales Verhalten gut und erkennen jede spezifische Klasse von Angriffen besser. Vergleichen den traditionellen Algorithmen des maschinellen Lernens sehen wir, dass das vorgeschlagene Modell LR und NB in Bezug auf Genauigkeit, Rückruf und FPR. Im Vergleich zu anderen BLS-Modellen ist das Das TBLS(W)-Modell weist eine etwas höhere Genauigkeit und Trefferquote auf als das vorgeschlagene Modell, da Für die TBLS wählten die Autoren nur den Datensatz eines einzigen Tages, Mittwoch, für das Modell Training und Tests. Es macht nur 30 % des gesamten Datensatzes aus, und die Angriffsarten sind alle Angriffe vom Typ DoS/DDoS, die im Vergleich zum Gesamtdatensatz besser ausgeglichen sind, und Die Erkennungsleistung des Modells wird bis zu einem gewissen Grad beeinträchtigt.

Machine Translated by Google



(A) **(B)**

Abbildung 13. Der Vergleich mit den modernsten Methoden zur Mehrfachklassifizierung: (a) der Vergleich von FPR; b) der Vergleich der Zeitkosten.

Tabelle 8. Der Vergleich mit modernsten Methoden zur Mehrfachklassifizierung.

Methode	Genauigkeit	Abrufen	FPR
Unsere Studie	0,958	0,947	0,012
LR	0,934	0,301	0,057
Achtung	0,286	0,746	0,126
GA-ANN	0,815	-	-
GSPSO-ANN	0,840	-	-
TBLS(W) [18]	0,978	0,977	-
BiGAN [33]	0,823	0,763	0,142
MFFSEM(W) [2]	0,999	0,999	0,013
DNN [34]	0,946	0,846	-
MECNN [29]	0,997	0,791	-
AdaBoost [35]	0,889	0,234	-
SDAE-SVM [36]	0,954	0,444	-
Die Genauigkeitsrate von Angreifern (DR) [37]	0,985	0,742	-
LSTM [38]	0,989	0,748	-
CFS-BA(W) [39]	0,999	0,999	0,120

Im Vergleich zu den anderen Algorithmen des maschinellen Lernens für die mehrdimensionale Feature Fusion and Stacking Ensemble Mechanism (MFFSEM) und korrelationsbasierte Feature Auswahl Fledermaus-Algorithmus (CFS-BA) Modelle, die Erkennungsleistungen sind besser als die des vorgeschlagenen Modells, aber ihre FPRs sind höher als die des vorgeschlagenen Modells, was dazu führen wird, zu einer Fehleinschätzung normalen Verhaltens als Angriff. Die beiden Modelle werden auch mit dem gleichen Datensatz von nur einem Tag, Mittwoch. Im Vergleich zu den langfristigen kurzfristigen Memory (LSTM)-Modell, seine Genauigkeit ist etwas höher als die des vorgeschlagenen Modells, aber seine Der Rückruf ist viel niedriger als beim vorgeschlagenen Modell, was bedeutet, dass das LSTM-Modell nicht erkennen die Angriffe gut. Zusammenfassend lässt sich sagen, dass das vorgeschlagene Modell im Vergleich zu anderen Modellen hat Vorteile wie eine bessere Erkennungsleistung und eine geringere Falschalarmrate bei Durchführen der Multiklassifizierungsaufgabe.

5. Modellvalidierung

Der Netzwerkverkehr in realen Umgebungen ist groß und komplex. Zur Bewertung des DBELS Um die Auswirkungen realer Umgebungen besser zu modellieren und zu simulieren, haben wir uns für das CICIDS2018 entschieden. Datensatz, um die Leistung des Modells in verschiedenen Datenumgebungen zu testen und zu validieren. Der Der CICIDS2018-Datensatz ist ein weit verbreiteter Datensatz in der Cybersicherheitsforschung und wurde vom kanadischen Institut für Cybersicherheitsforschung, das reale Netzwerksimulationen simuliert Verkehr, einschließlich normalem Verkehr und vielen Arten von Angriffen. Im Vergleich zum CICIDS2017 Der Datensatz hat sich in seiner Größe etwa vervierfacht und enthält mehr als acht Millionen Verkehrsaufzeichnungen; es gibt jedoch auch ein ernstes Datenungleichgewicht. Der Datensatz enthält über

Machine Translated by Google

acht Millionen Verkehrsdatensätze mit bis zu 80 Merkmalen pro Stichprobe. Diese Daten umfassen sowohl normaler Verkehr und ein breites Spektrum an Angriffsverkehr, wobei die Beispieltypen hauptsächlich Folgendes umfassen: Gutartig, DoS/DDoS, Brute Force, Botnet und Infiltration. Die Vielfalt und Komplexität von Die Stichprobentypen im CICIDS 2018-Datensatz sorgen dafür, dass der Datensatz die realen Cyberbedrohungen besser widerspiegelt, was eine umfassende Bewertung der Leistung ermöglicht und Generalisierungsfähigkeiten von Intrusion Detection Systemen.

Die Modellvalidierungsexperimente mit dem CICIDS2018-Datensatz umfassten eine Vorverarbeitung der Daten und eine Reduzierung der Datendimensionalität, und die Form der Operationen bl im Einklang mit CICIDS2017. Die Modelle für ihre Vergleichsexperimente umfassen BLS, DT, NB, MLP und CNN. Um die Gültigkeit der Experimente sicherzustellen, wird die Datenvorverarbeitung gleich, und alle Experimente werden in einer Umgebung der Zustimmung durchgeführt. Die Auswertung Zu den Messgrößen für das Vergleichsexperiment gehörten Genauigkeit, Rückruf, Falsch-Positiv-Rate und Trainingszeit. Die Leistung verschiedener Modelle bei der binären Klassifizierungsaufgabe ist in Tabelle 9 dargestellt, und die Leistung bei der Mehrfachklassifizierungsaufgabe ist dargestellt in Tabelle 10.

Tabelle 9. Die Leistung verschiedener Methoden zur binären Klassifizierung für den CICIDS2018-Datensatz.

Methode	Bewertungsmetrik			
	Genauigkeit	Abrufen	FPR	Mal)
Unsere Studie	0,981	0,956	0,044	5.295
BLS	0,958	0,931	0,069	18.826
DT	0,917	0,862	0,138	16.770
Achtung	0,791	0,742	0,258	2.556
MLP	0,963	0,933	0,067	861.788
CNN	0,957	0,926	0,074	949.476

Tabelle 10. Die Leistung verschiedener Methoden bei der Mehrfachklassifizierung für den CICIDS2018-Datensatz.

Methode	Bewertungsmetrik			
	Genauigkeit	Abrufen	FPR	Mal)
Unsere Studie	0,946	0,923	0,016	4.731
BLS	0,953	0,702	0,031	12.729
DT	0,891	0,499	0,060	17.266
Achtung	0,828	0,860	0,047	2.694
MLP	0,934	0,485	0,046	1025.025
CNN	0,966	0,785	0,026	916.550

Die Ergebnisse zeigen, dass unser vorgeschlagenes Modell die meisten anderen Modelle in der Genauigkeit übertrifft, Rückruf, FPR und Zeitaufwand bei verschiedenen Klassifizierungsaufgaben. Für die binäre Klassifizierung Aufgabe, im Vergleich zur traditionellen BLS-Struktur, Genauigkeit und Rückruf werden um 0,023 verbessert und 0,025, und die Falschalarmrate wird um 0,025 verringert. Im Vergleich zum Im Vergleich zu den herkömmlichen maschinellen Lernalgorithmen DT und NB weisen beide eine deutlich verbesserte Erkennungsleistung auf, während die Rate falscher Alarme deutlich sinkt. Für die Deep-Learning-Algorithmen MLP und CNN wird die Erkennungsleistung verbessert und die Trainingszeit des Modells wird erheblich reduziert. Für die Multiklassifizierungsaufgabe ist die Das vorgeschlagene Modell weist außerdem eine hohe Erkennungsleistung und eine geringe Modelltrainingszeit auf. und im Vergleich zum NB-Modell hat es einen erheblichen Vorteil bei der Erkennungsleistung, obwohl die Modelltrainingszeit leicht verbessert ist. Im Vergleich zum CNN-Modell, Obwohl die Genauigkeit leicht um 0,02 abnimmt, ist es besser als das CNN-Modell in Bezug auf von Rückruf, FPR und Modelltrainingszeit.

Machine Translated by Google

Aus den obigen experimentellen Ergebnissen lässt sich schlussfolgern, dass das vorgeschlagene Modell auch bei einem größeren **und komplexeren CICIDS2018-Datensatz eine bessere Erkennungsleistung zeigen** kann, was weiter darauf hinweist, dass das Modell über ein gewisses Maß an **Skalierbarkeit und Anpassungsfähigkeit verfügt. Der Maßstab dieses Datensatzes ist im Vergleich zum CICIDS2017-Datensatz stark erhöht; dennoch kann das vorgeschlagene Modell weiterhin einen geringen Zeitaufwand aufrechterhalten , was weiter bestätigt, dass das Modell über ein gewisses Maß an Anpassungsfähigkeit und Leichtigkeit für reale komplexe Umgebungen verfügt. Unser vorgeschlagenes Modell zeigt durchgängig eine starke Erkennungsleistung und einen geringen Zeitaufwand über verschiedene Datensätze hinweg, sei es für die binäre oder die Mehrklassenklassifizierung. Diese Effizienz ist entscheidend für reale Anwendungen, bei denen die Aufrechterhaltung einer hohen Genauigkeit und Geschwindigkeit unerlässlich ist. Darüber hinaus steigern die Anpassungsfähigkeit des Modells an unterschiedliche Datenumgebungen und seine Robustheit gegenüber unterschiedlichen Bedrohungen seinen praktischen** Wert. Seine Skalierbarkeit stellt sicher, dass es zunehmende Datenmengen verarbeiten kann, was es z

6. Schlussfolgerungen

In diesem Artikel schlagen wir auf der Grundlage des BLS-Modells und unter Einführung der EQL v2-Neugewichtungsidee ein DBELS-Modell vor, das das Problem der niedrigen Erkennungsraten des herkömmlichen BLS bei hochdimensionalen und unausgewogenen Datensätzen im Bereich der NIDS sowie der geringen Erkennung von Angriffen der Minderheitsklasse lösen soll. Zunächst lernt das Modell den groß angelegten hochdimensionalen Datensatz vollständig über DBN, trainiert jede RBM-Struktur vollständig, erfasst die optimale Struktur und die Schlüsselinformationen des hochdimensionalen Datensatzes und stellt ihn als optimalen niedrigdimensionalen Datensatz dar. Anschließend wird das BELS-Modell angewendet, um die Klassifizierung des niedrigdimensionalen Datensatzes zu erlernen, und der Lerngrad des Modells für die Stichproben der Minderheitsklasse wird verbessert, indem die positiven und negativen Gradientenfaktoren angepasst werden, um das Klassifizierungsgewicht des Modells zu erhöhen und die Erkennung der Angriffe der Minderheitsklasse zu verbessern. Das DBELS wird mithilfe des CICIDS2017-Datensatzes vollständig experimentell getestet, und unser vorgeschlagenes Modell übertrifft andere Algorithmen in Bezug auf Genauigkeit, Rückruf, FPR und Zeit. Abschließend wird die Nützlichkeit des vorgeschlagenen Modells mit dem CICIDS2018-Datensatz weiter getestet und validiert. Die Ergebnisse zeigen, dass unser vorgeschlagenes DBELS-Modell erhebliche Vorteile bei der Lösung des Problems hochdimensionaler und unausgewogener Netzwerk-Intrusion-Detection-Daten bietet und die Probleme der langen Trainingszeit von Intrusion-Detection-Systemen und der geringen Erkennung von Minority-Class-Angriffen lindert, was es zu einer praktischen Intrusion-Detection-Methode macht. Wir planen , **in Zukunft Initialisierungsstrategien und Flow-Learning-Techniken einzuführen,** um die Stabilität und **Erkennungsleistung des BLS zu verbesserr**

Beiträge der Autoren: Konzeptualisierung, MD; Methodik, MD und CS; Software, MD, CS und YK; Validierung, MD; Untersuchung, CS, MD und SF; Ressourcen, XZ; Datenkuratierung, CS und SF; Schreiben – Vorbereitung des Originalentwurfs, MD und HX; Schreiben – Überprüfung und Bearbeitung, MD und HX; Visualisierung, MD und YK; Überwachung, XZ und HX. Alle Autoren haben die **veröffentlichte Version des Manuskripts gelesen und stimmen ihr zu .**

Finanzierung: Diese Forschung wurde von der National Natural Science Foundation of China (62276091) finanziert.

Erklärung zur Datenverfügbarkeit: Daten werden auf Anfrage zur Verfügung gestellt.

Interessenkonflikte: Die Autoren erklären, dass keine Interessenkonflikte bestehen.

Verweise

1. **Anderson, JP** Computer Security Threat Monitoring and Surveillance; **Technischer Bericht; James P. Anderson Company: Washington, DC, USA, 1980.**
2. **Zhang, H.; Li, J.-L.; Liu, X.-M.; Dong, C. Mehrdimensionaler Merkmalsfusions- und Stapelensemblemechanismus für Netzwerk Einbruchserkennung.** Future Gen. Comput. Syst. **2021,** 122, **130–143.** [CrossRef]
3. **Sain, H.; Purnami, SW Combine Sampling Support Vector Machine für die Klassifizierung unausgewogener Daten.** Procedia Comput. Sci. **2015,** 72, **59–66.** [CrossRef]

Machine Translated by Google

4. Thapa, N.; Liu, Z.; Kc, DB; Gokaraju, B.; Roy, K. Vergleich von maschinellen Lern- und Deep-Learning-Modellen für Netzwerk Intrusion Detection Systems. Future Internet **2020,** 12, **167.** [CrossRef]

5. Namakshenas, D.; Yazdinejad, A.; Dehghantanha, A.; Srivastava, G. Föderiertes, quantenbasiertes, datenschutzerhaltendes Bedrohungserkennungsmodell für das Internet der Dinge für Verbraucher. IEEE Trans. Consum. Electron. **2024.** [CrossRef]

6. Almiani, M.; AbuGhazleh, A.; Al-Rahayfeh, A.; Atiewi, S.; Razaque, A. Tiefes rekurrierendes neuronales Netzwerk zur Erkennung von IoT-Eindringlingen System. Simul. Modell. Praxis. Theorie **2020,** 101, **102031.** [CrossRef]

7. Wu, P. Deep Learning zur Erkennung von Netzwerkangriffen: Angriffserkennung mit Computational Intelligence; **UNSW: Sydney, Australien, 2020.**

8. Yazdinejad, A.; Dehghantanha, A.; Karimipour, H.; Srivastava, G.; Parizi, RM Ein robustes, die Privatsphäre wahrendes Federated-Learning- Modell gegen Model-Poisoning-Angriffe. IEEE Trans. Inf. Forensics Secur. **2024,** 19, **6693–6708.** [CrossRef]

9. Yazdinejad, A.; Dehghantanha, A.; Srivastava, G.; Karimipour, H.; Parizi, RM Hybrides, die Privatsphäre bewahrendes föderiertes Lernen gegen unregelmäßige Benutzer im Internet der Dinge der nächsten Generation. J. Syst. Archit. **2024,** 148, **103088.** [CrossRef]

10. Chen, CP; Liu, Z. Breites Lernsystem: Ein effektives und effizientes inkrementelles Lernsystem ohne die Notwendigkeit tiefer Architektur. IEEE Trans. Neural Netw. Learn. Syst. **2017,** 29, **10–24.** [CrossRef] [PubMed]

11. Gong, X.; Zhang, T.; Chen, CP; Liu, Z. Forschungsbericht für breites Lernsystem: Algorithmen, Theorie und Anwendungen. IEEE Trans. Cybern. **2021,** 52, **8922–8950.** [CrossRef]

12. Li, Z.; Rios, ALG; Xu, G.; Trajkovi´c, L. Maschinelle Lerntechniken zur Klassifizierung von Netzwerkanomalien und -einbrüchen. In Proceedings des 2019 IEEE International Symposium on Circuits and Systems (ISCAS), Sapporo, Japan, 26.–29. Mai 2019; S. 1–5.

13. Wu, T.; Fan, H.; Zhu, H.; You, C.; Zhou, H.; Huang, X. Intrusion Detection System kombiniert Enhanced Random Forest mit SMOTE Algorithmus. EURASIP J. Adv. Signal Process. **2022,** 2022, **39.** [CrossRef]

14. Zhang, B.; Liu, Z.; Jia, Y.; Ren, J.; Zhao, X. Methode zur Netzwerk-Intrusion-Detection basierend auf PCA und Bayes-Algorithmus. Secur. Commun. Netw. **2018,** 2018, **1914980.** [CrossRef]

15. Li, Z.; Batta, P.; Trajkovic, L. Vergleich von Machine-Learning-Algorithmen zur Erkennung von Netzwerkangriffen. In Proceedings der 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 7.–10. Oktober 2018; S. 4248–4253.

16. Rios, ALG; Li, Z.; Xu, G.; Alonso, AD; Trajkovi´c, L. Erkennen von Netzwerkanomalien und Eindringlingen in Kommunikationsnetzwerken. In Proceedings der 2019 IEEE 23rd International Conference on Intelligent Engineering Systems (INES), Gödöllÿo, Ungarn, 25.–27. April 2019; S. 000029–000034.

17. Rios, ALG; Li, Z.; Bekshentayeva, K.; Trajkovi´c, L. Erkennung von Denial-of-Service-Angriffen in Kommunikationsnetzwerken. In Proceedings des 2020 IEEE International Symposium on Circuits and Systems (ISCAS), Sevilla, Spanien, 12.–14. Oktober 2020; S. 1–5.

18. Li, J.; Zhang, H.; Liu, Z.; Liu, Y. Netzwerk-Einbruchserkennung über ein dreifaches Lernsystem basierend auf räumlich-zeitlicher Granularität. J. Supercomput. **2023,** 79, **9180–9205.** [CrossRef]

19. Ahmad, T.; Aziz, MN Datenvorverarbeitung und Merkmalsauswahl für maschinelle Lernsysteme zur Erkennung von Angriffen. ICIC Express Lett. **2019,** 13, **93–101.**

20. Hao, X.; Jiang, Z.; Xiao, Q.; Wang, Q.; Yao, Y.; Liu, B.; Liu, J. Mehr produzieren mit weniger: Ein GAN-basierter Ansatz zur Erkennung von Netzwerkangriffen für unausgewogene Daten. In Proceedings der 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Dalian, China, 5.–7. Mai 2021; S. 384–390.

21. Tan, J.; Lu, X.; Zhang, G.; Yin, C.; Li, Q. Equalization Loss v2: Ein neuer Gradientenausgleichsansatz für die Erkennung langschwänziger Objekte. In Proceedings der IEEE/CVF-Konferenz zu Computer Vision und Mustererkennung, Nashville, TN, USA, 20.–25. Juni 2021; S. 1685–1694.

22. Shen, Z.; Zhang, Y.; Chen, W. Eine bayesianische Klassifizierungsmethode zur Intrusion Detection basierend auf der Fusion von PCA und LDA. Secur. Commun. Netw. **2019,** 2019, **6346708.** [CrossRef]

23. Salo, F.; Nassif, AB; Essex, A. Dimensionsreduktion mit IG-PCA und Ensemble-Klassifikator zur Netzwerk-Einbruchserkennung. Comput. Netw. **2019,** 148, **164–175.** [CrossRef]

24. Hinton, GE; Osindero, S.; Teh, Y.-W. Ein schneller Lernalgorithmus für tiefe Glaubensnetze. Neural Comput. **2006,** 18, **1527–1554.** [CrossRef] [PubMed]

25. Belarbi, O.; Khan, A.; Carnelli, P.; Spyridopoulos, T. Ein Intrusion Detection System basierend auf Deep Belief Networks. In Proceedings der International Conference on Science of Cyber Security, Shimane, Japan, 10.–12. August 2022; S. 377–392.

26. Andresini, G.; Appice, A.; Malerba, D. Nearest-Cluster-basierte Einbruchserkennung durch Convolutional Neural Networks. Knowl. Based Syst. **2021,** 216, **106798.** [CrossRef]

27. Kim, A.; Park, M.; Lee, DH AI-IDS: Anwendung von Deep Learning zur Echtzeit-Erkennung von Web-Angriffen. IEEE Access **2020, 8, 70245–70261.** [CrossRef]

28. Thakkar, A.; Lohiya, R. Fusion statistischer Bedeutung für die Merkmalsauswahl in Deep Neural Network-basierter Intrusion Detection System. Inf. Fusion **2023,** 90, **353–363.** [CrossRef]

29. Chen, Y.; Lin, Q.; Wei, W.; Ji, J.; Wong, K.-C.; Coello, CAC Intrusion Detection mithilfe eines multi-objektiven evolutionären Convolutional Neural Network für das Internet der Dinge im Fog Computing. Knowl. Based Syst. **2022,** 244, **108505.** [CrossRef]

30. Thakkar, A.; Lohiya, R. Angriffsklassifizierung von unausgewogenen Intrusionsdaten für IoT-Netzwerke unter Verwendung von auf Ensemble-Lernen basierenden Deep neuronales Netzwerk. IEEE Internet Things J. **2023,** 10, **11888–11895.** [CrossRef]

Machine Translated by Google

31. Mulyanto, M.; Faisal, M.; Prakosa, SW; Leu, J.-S. Wirksamkeit des Fokusverlusts für die Klassifizierung von Minderheiten bei Netzwerkintrusion Detektionssysteme. Symmetry **2020,** 13, **4.** [CrossRef]

32. Idrissi, MJ; Alami, H.; El Mahdaouy, A.; El Mekki, A.; Oualil, S.; Yartaoui, Z.; Berrada, I. Fed-anids: Föderiertes Lernen für Anomalie-basierte Netzwerk-Intrusion-Detection-Systeme. Expert Syst. Appl. **2023,** 234, **121000.** [CrossRef]

33. Yao, W.; Shi, H.; Zhao, H. Skalierbare anomaliebasierte Intrusion Detection für sicheres Internet der Dinge unter Verwendung generativer kontradiktorischer Netzwerke in einer Fog-Umgebung. J. Netw. Comput. Appl. **2023,** 214, **103622.** [CrossRef]

34. Jose, J.; Jose, DV Deep-Learning-Algorithmen für Intrusion Detection Systeme im Internet der Dinge unter Verwendung des CIC-IDS 2017-Datensatzes. Int. J. Electr. Comput. Eng. (IJECE) **2023,** 13, **1134–1141.** [CrossRef]

35. Rojas, R. AdaBoost und der Super Bowl der Klassifikatoren – eine Einführung in Adaptives Boosting. Freie Univ. Berl. Tech. Rep. **2009, 1, 1–6.**

36. Lv, Z.; Qiao, L.; Li, J.; Song, H. Sicherheitsprobleme durch Deep Learning im Internet der Dinge. IEEE Internet Things J. **2020, 8, 9531–9538.** [CrossRef]

37. Assis, MV; Carvalho, LF; Lloret, J.; Proença Jr, ML Ein GRU-Deep-Learning-System gegen Angriffe in softwaredefinierten Netzwerken. J. Netw. Comput. Appl. **2021,** 177, **102942.** [CrossRef]

38. Imrana, Y.; Xiang, Y.; Ali, L.; Abdul-Rauf, Z. Ein bidirektionaler LSTM-Deep-Learning-Ansatz zur Einbruchserkennung. Expert Syst. Appl. **2021,** 185, **115524.** [CrossRef]

39. Zhou, Y.; Cheng, G.; Jiang, S.; Dai, M. Aufbau eines effizienten Intrusion Detection Systems basierend auf Merkmalsauswahl und Ensemble-Klassifikator. Comput. Netw. **2020,** 174, **107247.** [CrossRef]

Haftungsausschluss/Anmerkung des Herausgebers: Die in allen Veröffentlichungen enthaltenen Aussagen, Meinungen und Daten sind ausschließlich die der einzelnen Autoren und Mitwirkenden und nicht die von MDPI und/oder den Herausgebern. MDPI und/oder die Herausgeber lehnen jegliche Verantwortung für Personen- oder Sachschäden ab, die aus den im Inhalt erwähnten Ideen, Methoden, Anweisungen oder Produkten resultiere