Blockchain-gestütztes verteiltes Maschinelles Lernen für autonomes Fahren

Felix Reichel

XAIN AG

Zentrum für Luft- & Raumfahrt III
Schmiedestraße 2A
15745 Wildau
E-Mail: felix.reichel@xain.io

ABSTRACT

Bei der Entwicklung von Modellen für das autonome Fahren gilt als verbreitete Praxis das primäre Machine Learning nicht lokal, sondern an einem zentralen Ort durchzuführen. Das Resultat ist die Speicherung, Verwaltung und Auswertung einer riesigen Menge an personenbezogenen Daten, welche das Fahrverhalten und das Fahrzeug selbst betreffen. Dieser Aspekt ist bezogen auf die Privatsphäre und die Wirtschaftlichkeit kritisch zu betrachten.

Der Ansatz verteiltes Maschinelles Lernen mit der Blockchain Technologie zu kombinieren ist eine Möglichkeit der massiven zentralen Datenverwaltung entgegenzuwirken.
Obwohl globale Anpassungen des generellen Modells erfolgen und lokale Modelle unter den Netzwerkteilnehmern ausgetauscht werden, wird die Gesamtheit der Informationen nicht im Netzwerk geteilt.
Als praktischer Anwendungsfall wird ein Testnetzwerk mit geringerer Skalierung beschrieben, welches mit elektrisch betriebenen Modellfahrzeuge im Maßstab 1:10 ausgestattet wird. Die Fahrzeuge zeichnen sich durch eine vollständig autonome Fahrweise aus, welche durch das Machine Learning auf Basis des lokal im Fahrzeug integrierten Embedded-Moduls ermöglicht wird.

Das Ziel besteht zum einen darin die Privatsphäre des Fahrzeugnutzers durch begrenzte Datenspeicherung zu wahren und zum anderen eine gezieltere lokale Auswertung der Daten zu erreichen. Zusätzlich verfolgt der Ansatz die Einhaltung der EU-Datenschutzgrundverordnung, welche under anderem festlegt, dass personenbezogene Daten nicht um ihrer selbst willen erhoben werden sollten, sondern nur wenn es zur Erbringung eines Dienstes wirklich erforderlich ist.

SCHLÜSSELWÖRTER

Verteiltes Maschinelles Lernen, Deep Learning, Blockchain, Autonomes Fahren, Automobiltechnik

1. EINFÜHRUNG

Die Entwicklung autonomer Fahrzeuge schreitet stetig voran. Gegenwärtig findet ein regelrechtes Rennen zwischen einzelnen Automobilkonzernen und IT-Firmen statt, welches Unternehmen als erstes ein autonomes Fahrzeug der Stufe 5 (fahrerlose Fortbewegung) auf den Markt bringt.
Deep Learning als Untergruppe vom Maschinellen Lernen stellt im Zusammenhang mit der Erhebung und Nutzung jeglicher Formen von Daten als Eingabeparameter einen zentralen Grundbaustein zur Umsetzung des autonomen Fahrens dar. Dieser Aspekt birgt jedoch hinsichtlich der Erfassung und Speicherung von personenbezogenen Daten Risiken in Bezug auf die Privatsphäre der Fahrzeughalter. Die ab Mai 2018 in Kraft tretende EU-Datenschutzgrundverordnung (EU-DSGVO) stellt komplett neue Anforderungen an die Verarbeitung personenbezogener Daten. Demnach dürfen die Daten lediglich für festgelegte Zwecke erhoben werden und nur auf eine nachvollziehbare Weise verarbeitet werden. Im Kontext mit autonomen Fahren bedeutet dies für den Großteil der Systeme eine Selektierung der Daten, welche Informationen tatsächlich für eine autonome Fortbewegung erhoben werden müssen.

In dem Artikel wird durch die Kombination der Blockchain Technologie mit verteiltem Maschinellen Lernen (ML) eine Lösung vorgestellt, welche die EU-DSGVO einhält, die Privatsphäre des Fahrzeughalters schützt, auf ML und Blockchain bezogen skalierbar ist und dennoch die Bildung effektiver ML-Modelle für autonomes Fahren zulässt.

2. DEEP LEARNING

Deep Learning bezeichnet eine Gruppe von Methoden zur Optimierung künstlicher neuronaler Netze mit komplexen Strukturen. Die Komplexität entsteht durch zahlreiche Zwischenlagen zwischen der Ein- und Ausgabeschicht.

Während der Mensch Aufgaben wie Sprach- oder Gesichtserkennung intuitiv löst, wurde die künstliche Intelligenz anfänglich vor eine Herausforderung gestellt, da sich diese Aufgaben nur schwer durch mathematische Regeln abbilden lassen. Die Lösung besteht für computerbasierte Systeme in der Lernfähigkeit von Erfahrungen und die Welt in Bezug auf eine Hierarchie von Konzepten zu verstehen. Die Grundidee: jedes Konzept ist durch seine Beziehung zu einfacheren Konzepten definiert. Auf diese Weise wird durch die Ansammlung von Wissen aus der Erfahrung die Notwendigkeit einer vorherigen formalen Spezifizierung des Wissens vermieden. Durch die Hierarchie der Konzepte sind Computersysteme in der Lage komplizierte Konzepte zu erlernen, indem diese aus einfacheren zusammengesetzt werden. Die Gesamtheit der übereinander aufgebauten Konzepte weist eine tiefe Schichtenstruktur auf. Dementsprechend wird dieser Ansatz in der künstlichen Intelligenz als „Deep Learning bezeichnet [1][2].
Der Automobilbereich bietet für Maschinelles Lernen und Deep Learning eine Vielzahl von Anwendungsfällen.

- **Autonomes Fahren:** Insbesondere die Verarbeitung von riesigen von den Sensoren bereitgestellten Daten (zum Beispiel LiDAR), das Lernen des Fahrverhaltens und das Meistern von bestimmten Situationen resultiert in der Notwendigkeit von Techniken des Maschinellen Lernens. Gegenwärtig werden neuronale Netzwerke in Fahrerassistenzsystemen (ADAS) zur Sicht, Sonar und Radar eingesetzt. Die Sicherheitskomponente wird unter anderem über KI-basierte Software-Lösungen durch visuelle Daten von Kamerasensoren sichergestellt.
- **Intelligente Robotersysteme:** Deep Learning findet in Bezug auf die Erkennung von Merkmalen in Kamerabildern und in Sensordaten zur Maschinensteuerung Anwendung. Weiterhin werden Deep Neural Networks (DNN), insbesondere im Zusammenhang mit Robotersystemen, zur Objekterkennung und als Schlüsselkomponente für eigenständiges Lernen über die Lebensdauer hinweg genutzt.
- **Visuelle Prüfung in Herstellungsprozessen:** Die stetig voranschreitende Weiterentwicklung von Mobilgeräten und IoT-Sensoren führt zu einer Flut von Bild- und Videodaten. Diese werden weitestgehend per Hand ausgewertet. Deep Learning kann helfen bestimmte Muster zu erkennen und die Daten nach erfolgter Analyse organisiert abzulegen [3][4].

3. HERAUSFORDERUNG: Modell-Skalierbarkeit

Die Möglichkeit neuronale Netze zu skalieren, ist eine notwendige Voraussetzung um Netzwerke, basierend auf großen Datenmengen in kurzer Zeit zu trainieren. In diesem Zusammenhang setzt der Begriff „Big Data“, welcher eine Vielzahl an Datenmengen zu unterschiedlichsten Bereichen beschreibt neue Anforderungen an Systeme, welche Maschinelles Lernen unterstützen. DNNs umfassen je nach Anwendung Millionen von Parametern und sind daher rechenintensiver als andere Lernalgorithmen. Der Umfang des Modells variiert je nach Aufbau und steigt mit der Tiefe und der Anzahl der Parameter. Damit die ML Algorithmen in dieser Skalierung betrieben werden können, ist die technische Planung des Aufbaus und der Architektur des Systems von entscheidender Bedeutung.

4. VERTEILTES MASCHINELLES LERNEN

Das Trainieren umfangreicher Datensätze innerhalb tiefer ML Modellstrukturen erfordert eine Form der Verteilung des Trainingsprozesses, wodurch die Zeitspannen für Forschungs- und Produktionszyklen möglichst gering gehalten werden können. Verteiltes Maschinelles Lernen setzt zusätzlich ein sorgfältiges

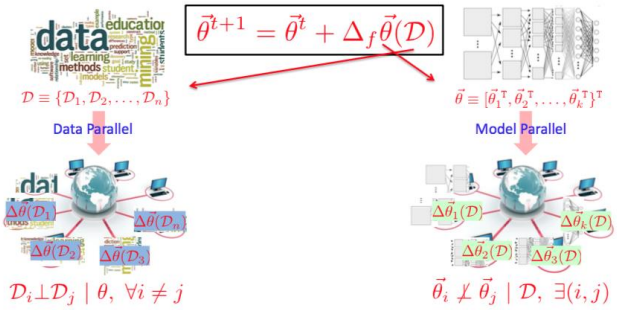

Abbildung 1: Unterschied zwischen Datenparallelität und Modellparallelität [5]

Management der Rechenleistung und der Kommunikationsintervalle, sowie eine verteilte systematische Koordination voraus.
Diesbezüglich existieren im Allgemeinen zwei näher zu betrachtende wesentliche Arten der Parallelisierung: Parallelität der Daten und Parallelität des von Maschinellen Lernen genutzten Modells.
Bei der vielgenutzten und weit verbreiteten Anwendung der datenparallelen Strategie wird der Datensatz mehrfach unterteilt und den verschiedenen parallelen angelegten Arbeitern oder direkt den Rechnern als übergeordnete Instanz zugewiesen. Jeder Arbeiter führt Aktualisierungen der Parameter speziell für seinen Datenabschnitt durch. Nach jeder Trainingsepoche wird das Modell nach vorhergehender globaler Aggregation aller Parameter aktualisiert. Der Ansatz beinhaltet die Sammlung zahlreicher Aufgaben innerhalb eines Arbeiterjobs mit dem Ziel ein und dasselbe ML Modell

auf Grundlage von verschiedenen Datenteilstücken zu trainieren.
Im modellparallelen Training wird das Modell in einzelne Stücke unterteilt und anschließend den Arbeitern zugewiesen. Die Updates des Modells laufen über diverse Aktualisierungsfunktionen ab, welche parallel ausgeführt werden.

Im Gegensatz zu datenparallelem Training wird jeder Aktualisierungsfunktion zusätzlich eine Funktion zur zeitlichen Regelung und Koordinierung zugeteilt, die die Aktualisierungen insofern beeinträchtigt und begrenzt, dass nur eine Teilmenge der Modellparameter erneuert wird.
Hierdurch werden Updates derselben Parameter und somit ein ständiges Überschreiben durch verschiedene Arbeiter verhindert. Generell bestehen zwischen den Modellparametern Abhängigkeiten. Es hat sich jedoch herausgestellt, dass ein modellparalleler Algorithmus ausschließlich effektiv arbeitet, wenn jede Trainingsepoche und die damit einhergehenden parallelen Parameterupdates genau auf eine Teilmenge von gegenseitig unabhängigen Parametern begrenzt werden [5][7].
Beide Strategien werden heutzutage in vielfältigen Arten verwendet. Die Variationen der Systeme treten vor allem zum einen in der Art der Aktualisierung und Speicherung der Modelle und zum anderen in der Koordinierung der Arbeiter zur Abarbeitung der Aufgabenliste auf. Die Modellspeicherung kann zentral über einen Master-Knoten, eine Reihe von einzelnen Netzwerkknoten oder ausdrücklich für diesen Zweck vorgesehene Parameter-Server-Knoten erfolgen. Darüber hinaus existiert die Möglichkeit das Modell zu entlang der Arbeiterknoten aufzuteilen. Die Aktualisierung des Modells kann synchron oder asynchron durchgeführt werden [6].
Die meisten Bibliotheken zum verteilten tiefen Maschinellen Lernen stellen bereits eine verteilte Integrierung des Gradientenabstiegsverfahren speziell für paralleles Lernen zur Verfügung. Im Framework Tensorflow, einer Open-Source-Programmbibliothek für künstliche Intelligenz, werden die Aufgaben je nach Anspruch der Rechenleistung entweder als Parameter Server Job (ps) oder als Arbeiterjob (worker) deklariert. Während der rechenintensive Teil des Modells, beispielsweise die Optimierung der Parameter, im Arbeiterjob kreiert wird, werden geteilte Parameter in einem oder mehreren Aufgaben innerhalb eines Parameter Server Jobs aktualisiert und gespeichert. Da alle Aufgaben gewöhnlicherweise auf verschiedenen Maschinen ausgeführt werden, erlaubt diese Art der Aufgabenzuordnung einen kontinuierlichen Datenaustausch zwischen den Jobs.

5. HERAUSFORDERUNG: Privatsphäre / Schutz der personenbezogenen Daten

An dieser Stelle soll nicht weiter auf die theoretische Funktionalität des verteilten Maschinellen Lernens und der derzeit in diesem Zusammenhang existierenden Programmbibliotheken eingegangen werden. Im Folgenden wird näher betrachtet, welche Anwendungsfälle von datenparallelen und modellparallelen Training sich in Bezug auf eine autonome Fortbewegungsweise ergeben. Dazu wird als aktuelles Beispiel zu Maschinellem Lernen in der Automobilindustrie konkret das Unternehmen Tesla aufgegriffen.
Tesla nutzt im Bereich autonomen Fahren einen Autopiloten, welcher durch den Einsatz von Algorithmen des Maschinellen Lernens kontinuierlich lernt. Die Gesamtheit der Tesla Fahrzeuge fungiert als ein Netzwerk. Die Unternehmen sind bestrebt so viele Informationen wie möglich zur Optimierung der Autopilot-Systeme im Netzwerk zu sammeln. Dies beinhaltet die lokale Anhäufung an Daten in jedem Fahrzeug wie Informationen zum Fahrverhalten über die drahtlose Verbindung des Fahrzeugs, detaillierte GPS- und Kartenkoordinaten, Sensoren und direkte Daten von den unternehmenseigenen Forschungsfahrzeugen. Diese Datensätze bilden die Grundlage für das Training des ML-Modells. Es handelt sich hierbei um „Supervised Learning“ -überwachtes Lernen, bei welchem jeder Datenpunkt in der vorhandenen Datenmenge aus einem Eingabe- und Ausgabewert besteht. Dementsprechend lernt der Autopilot von dem Fahrzeugführer eine korrekte Fahrweise.
Damit der Autopilot nicht nur von dem individuellen Fahrverhalten des Fahrers abhängig ist, wird zuvor ein computerbasiertes Vokabular rund ums Fahren erstellt. Dazu durchlaufen Videos und Daten von bis zu

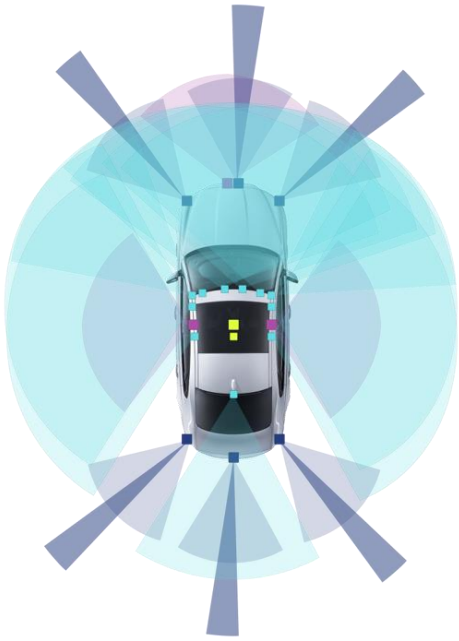

Abbildung 2: Typischer Aufbau eines selbstfahrenden Fahrzeugs bestehend aus einer Reihe von Sensoren: Kameras, Radar, Lidar zur Sicherstellung eines 360° Blickwinkels [9]

Millionen von zurückgelegten Kilometern ein Computer-Datenmodell. Die Algorithmen nutzen visuelle Techniken um die Videos aufzuschlüsseln und den Inhalt verstehen zu können. Der Hintergrundgedanke ist eine zielführende augenblickliche Reaktion des Fahrzeugs in einer unerwarteten Situation.
Das Fahrvokabular wird anschließend auf leistungsfähige Rechenhardware geladen, welche im Fahrzeug verbaut ist. Automobilunternehmen wie Tesla nutzen das Konzept als Grundbaustein um später das Modell über die Zeit hinweg durch weitere Anbindungen von Daten unterschiedlichster Natur aus vielfältigen Quellen zu prägen.
Die Daten des Fahrverhaltens werden über die Wahl der verbauten Hardware greifbar. Ein Fahrzeug, welches zur

Datenanreicherung im Anwendungsfall autonomes Fahren genutzt wird, ist mit zahlreichen Sensoren ausgestattet. Die Fahrzeugausstattung schließt typischerweise Technologien wie Radar, Kameras, Lidar, Ultraschallsensoren und ein breites Spektrum an Fahrzeugsensorik verteilt über die Netzwerk-Steuereinheit des Fahrzeugs, Flexray, Ethernet und andere Netzwerktypen ein.
Der Nutzen der Verwendung von Kameras, Radars und anderen Sensoren liegt darin das Computersystem mit einer Art Sicht und Redundanz zu versehen, welche das Fahrzeug gegen widrige Wetterbedingungen oder Fehlverhalten einzelnen Komponenten schützt. Die Komponenten erfassen und sammeln kontinuierlich Daten zur Verbesserung des selbstfahrenden Assistenzsystems, aber auch zur Datenanreicherung um in Zukunft eventuell Optimierungen an Modellen durch neue Trainingsdurchläufe zu erzielen. Tesla's Fahrzeugflotte ist daher mit dauerhaft eingeschalteten drahtlosen Netzwerkverbindungen ausgerüstet, welche die Datenmengen zum Fahrverhalten sammeln und zur Unternehmens-Cloud senden. Im Anschluss erfolgt mittels Software eine Analyse und Auswertung. Die daraus resultierenden Modelle werden getestet und in die zukünftig erfolgenden Software Updates integriert. Jene werden in regelmäßigen Abständen von den Servern zu den Kundenfahrzeugen kopiert. Diese Praxis wird allgemein als „over-the-air" Update betitelt.

Die zentrale Speicherung der personenbezogenen Daten hinsichtlich der Fahrweise, der genauen Fahrzeugpositionen und der zurückgelegten Strecken ist an diesem Punkt kritisch zu betrachten. Im Gegensatz zu Anwendungen, denen der Nutzer explizit Berechtigungen zur Ausführung bestimmter Funktionen erteilt, weiß der Fahrzeughalter weder welche Daten genau erhoben werden noch wofür jene Informationen Verwendung finden. Dies stellt insbesondere für den Fahrzeughalter eine Beeinträchtigung der Privatsphäre dar.

6. BLOCKCHAIN TECHNOLOGIE

Bei der Blockchain Technologie handelt es sich stark vereinfacht um eine Art dezentraler Datenbank, welche wie eine Liste an Datensätzen kontinuierlich erweitert wird („Blöcke"). Mithilfe kryptographischer Verfahren werden die Blöcke miteinander verbunden. Damit eine Transaktion zum Beispiel in Form einer Information der Blockchain hinzugefügt werden kann, muss diese vorher durch die Netzwerkteilnehmer verifiziert werden. Über ein Konsensverfahren wird sichergestellt, dass alle Netzwerkteilnehmer über die gleiche Datenbasis verfügen.

7. MASCHINE LEARNING & BLOCKCHAIN

Im nachfolgenden Abschnitt wird eine Variante vorgestellt der massiven Datenspeicherung entgegenzuwirken und eine gezieltere lokale Auswertung der Daten zu ermöglichen.
Die Grundlage hierfür bilden die bereits zuvor beschriebenen Strategien des datenparallelen und modellparallelen Trainings des verteilten Maschinellen Lernens. Der Ansatz verfolgt das Ziel Maschine Learning Modelle noch vor einer gesamten Aggregation der Datensätze der einzelnen Fahrzeuge zu trainieren. Im Gegensatz zu Tesla's Ansatz die Daten lokal zu sammeln und nach erfolgter Übertragung an einem zentralen Ort zu analysieren und auszuwerten, findet das Maschinelle Lernen direkt lokal in den Fahrzeugen statt. Dazu werden die erfassten Daten im Speicher des Fahrzeugs für das Training der autonomen Fortbewegungsweise verwendet. Durch gezielt parallel stattfindende Trainingsdurchläufe werden voneinander unterschiedliche Modelle gebildet. Auch hier existieren je nach verfolgtem Ziel zwei voneinander abweichende Grundideen: die Bildung von Modellen mit ein und demselben Ziel auf ähnlicher Datenlage oder die Bildung von Modellen mit unterschiedlichen Zielen mit komplett grundverschiedenen Daten für das Machine Learning Training. Letztere Idee spielt im Zusammenhang mit der Entwicklung mehrerer Modelle zur Kombination miteinander in einem Gesamtsystem eine übergeordnete Rolle. Auf die Praxis bezogene Anwendungsfälle sind das Erlernen eine Fahrspur zu halten, im Detail die Fahrbahnmarkierungen nicht zu überschreiten, und den Sicherheitsabstand zum vorausfahrenden Fahrzeug je nach Geschwindigkeit, Wetter- und Straßenbedingungen einzuhalten. Beide Modelle werden auf Grundlage von verschiedenen Datentypen unabhängig voneinander in zwei unterschiedlichen Fahrzeugen trainiert. Anschließend werden die Modelle in einer Gesamtstruktur integriert.

Unabhängig von der Vorgehensweise ist die Validierung des erlernten Modells, in welcher Qualität die Aufgabe ausgeführt oder ein bestimmtes Ziel erreicht wird von großer Bedeutung. Dies beschreibt in erster Linie, welche Aspekte das ML-Modell zufriedenstellend, mittelmäßig oder nur unzureichend erlernt hat. Dies ist entweder direkt über das Fahrverhalten des Fahrers oder über Fahrsimulationen möglich.

Der Austausch der Modellparameter erfolgt unter den jeweiligen Fahrzeugen über Fahrzeugknoten innerhalb

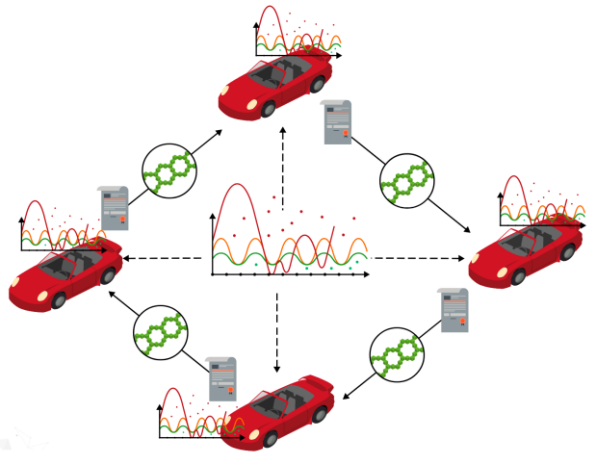

Abbildung 3: Lokales Machine Learning im Blockchain Netzwerk

des Netzwerks. Um einer etwaigen Manipulation des Informationsaustausches durch fremde und externe V2X und IoT Objekten zu blockieren, findet eine Verifizierung der Kommunikation auf Basis der Blockchain Technologie statt. Weiterhin wirkt das Blockchain Netzwerk als vertrauensvolle Plattform zur Einhaltung der Privatsphäre der personenbezogenen Daten, da nicht alle Daten im Netzwerk geteilt werden müssen. Das bedeutet, dass das durch ML lokal erlernte Wissen ohne die Notwendigkeit von Vertrauen unter den Netzwerkteilnehmer geteilt werden kann, welches der DSGVO entspricht die Daten privat zu halten. Des weiteren können bestimmte Netzwerkeffekte erzielt und Innovationen implementiert werden. Besonders für hochgradig lokale Probleme ist die Kombination von lokalen und globalen ML-Modellen erstrebenswert.

8. NETZWERK AUS MODELLFAHRZEUGEN ALS PRAKTISCHER ANWENDUNGSFALL

Um die Kombination der Blockchain Technologie mit verteilten Maschinellen Lernen in einem greifbaren Praxiseinsatz näher zu beleuchten und einzelne Funktionalitäten zu testen, werden innerhalb eines Fahrzeug-Testnetzwerks mit geringer Skalierung elektrisch betriebene Modellfahrzeuge aufgesetzt. Die Grundarchitektur des Netzwerks baut auf der Blockchain Technologie auf, wodurch die Fahrzeuge miteinander vernetzt werden. Die Vorteile, welche sich aus der dezentralen Netzstruktur ergeben, werden später im Zusammenhang mit Maschinellem Lernen näher betrachtet.

Die Fahrzeuggröße richtet sich nach der Anzahl und der zur Erfüllung der Anforderungen vorausgesetzten Leistungen der zu verbauenden Hardwarekomponenten. Der Maßstab wird somit auf 1:10 festgelegt. Von der Hardwareseite werden Kernkomponenten wie Prozessorsteuereinheit, Sensorik (Kameramodul & Ultraschall) zur Erfassung der Straßenführung und der statischen und dynamischen Objekte im Straßenverkehr, Motorantrieb, sowie Energieversorgung (Lithium-Ionen-Akkumulatoren) in den Fahrzeugen montiert. Die Bauteile werden zwecks gegenseitiger Abschirmung auf mindestens zwei übereinander montierten Plateauplatten angeordnet. Die räumliche Trennung erfolgt auf Grundlage von elektromagnetischer Verträglichkeit (EMV). Während im unteren Chassis die Grundausstattung zur Umsetzung der Fortbewegung platziert wird, erfolgt die Anordnung der für zusätzliche Anwendungen erforderten technischen Bestandteile (Sensorik, Steuereinheit, Kommunikationsmodul) auf den oberen Plateaus. Die Nachstellung von für die Technologien relevanter Testfälle setzt eine vollständig autonome Fortbewegungsweise der Modellfahrzeuge voraus. Dazu wird die Straßenführung in den von der Kamera erfassten Bildern vorgeprägt, welche anschließend als Eingabeparameter für das Training eines neuronalen Netzwerks für die autonome Steuerung dienen.

Zum allgemein besseren Verständnis ist anzumerken, dass die Algorithmen des Maschinellen Lernens lokal im Fahrzeug selbst durch den Einsatz von Embedded-Modulen ausgeführt werden. Es wird ein globales Modell für die grundautonome Fahrweise trainiert, welches auf der Erkennung und der Orientierung der Straßenlinien beruht. Zu Beginn wird das Modell auf allen Fahrzeugen gleichermaßen ausgerollt. Über das Blockchain Netzwerk können die Modellparameter im Laufe der Zeit modifiziert und an alle Fahrzeugteilnehmer, welche Knotenpunkte darstellen, über „over-the-air“ Updates verteilt werden. Neben den globalen Aktualisierungen lernen die Fahrzeuge komplett unabhängig voneinander eigene Modelle je nach Infrastrukturbedingungen wie in etwa die Seitenabstände ohne Straßenführung einzuhalten oder gar mit Feldwegen umzugehen. Zusätzlich ist durch verstärkendes Lernen (Reinforcement Learning) eine auf den Energieverbrauch optimierte Fahrweise zur Senkung von Energieverschwendungen und Materialverschleiß und zur Erhöhung der Verkehrssicherheit denkbar.

Der Grundgedanke ist, Modelle mit dem Ziel des autonomen Fahrens zu bilden, welche nicht in optimalen Umgebungen entstanden sind und somit einen realitätsnahen Fall widerspiegeln. Auf das lokale Lernen von den Umgebungsbedingungen zu verzichten ist ausgeschlossen, da kein globales Modell entwickelt werden kann, welches alle existierenden Infrastrukturzustände abdecken kann. Dabei handelt es sich um hochgradig lokale Probleme, welche von Land zu Land und Kontinent zu Kontinent sehr differierend voneinander sind. Um dennoch autonomes Fahren in möglichst jeder denkbaren Situation sicherzustellen, fungiert die Blockchain Technologie als zentrale Rolle, indem das verifizierbare „Wissen“ der Fahrzeuge untereinander geteilt wird. Dazu fragen die Netzwerkteilnehmer je nach Situation und Position nach einem im Netzwerk befindlichen Modell an. Die Freigabe eines Modells kann für einen bestimmten Zeitraum erfolgen ohne alle Daten transferieren oder tauschen zu müssen. Hierdurch werden Netzwerkeffekte

des Lernens erzielt ohne der Notwendigkeit von zentralen Datenbanken.
Ferner findet auf Basis der Blockchain eine Verifizierung der Kommunikation statt um eine etwaige Manipulation des Informationsaustausches durch fremde externe V2X und IoT Objekte zu blockieren.

Folgende Testfälle gilt es in Bezug auf die autonome Fahrweise an sich und das Netzwerk selbst zu untersuchen:

- Einhaltung der Straßenmarkierungen
- Beachtung von Verkehrsschildern, Ampeln und anderen Verkehrsteilnehmern
- Reaktion auf veränderte Infrastruktur
- Simulation einer Cyberattacke von außen auf die Blockchain zur Beeinflussung der Netzwerkstabilität
- Flexibilität der Integrierung weiterer IoT-Funktionen

Da die Testfälle und Rahmenbedingungen mit der Zeit variieren, ist eine dynamisch anpassbare Fahrzeuginfrastruktur erstrebenswert.

Zusätzlich zu den bereits erwähnten Faktoren erlaubt das Netzwerk das Testen alternativer Machine Learning Modelle, wodurch neue Anwendungsfälle erschlossen werden können. Dabei sind die Grenzen und demzufolge die Menge der möglichen Ereignisse durch die kleinere Skalierung begrenzt.
Über diesem Netzwerk gestellt sind weitere intelligente Systeme wie ein automatisiert geregelter Verkehrsfluss durch die Übermittlung von Parametern zum Verkehrsstand und zur Straßenbeschaffenheit zwischen den Fahrzeugen und registrierten IoT-Objekten denkbar.

9. ZUSAMMENFASSUNG

Der Artikel präsentierte eine Möglichkeit das heutzutage zentral angewandte Machine Learning zum Zweck einer autonomen Fahrweise auf eine Vielzahl an Endgeräten lokal zu verteilen. Zugleich wurden die Herausforderungen des Datenschutzes in Bezug auf personenbezogene Daten im digitalen Wandel der Automobilindustrie unter Einhaltung der EU-DSGVO näher betrachtet. Als Kommunikationsverifizierung zur Sicherung des Modellaustausches eignete sich ein eigens optimiertes Blockchain Framework. Der Nutzen der Kombination der beiden Technologien wurde anhand eines praktischen Anwendungsfalls herausgestellt. In jenem werden verschiedenen Szenarien zur testweisen Untersuchung in einem dynamischen Umfeld angesprochen und analysiert.

10. LITERATUR

[1] Ian Goodfellow, Yoshua Bengio und Aaron Courville: Deep Learning (Adaptive Computation and Machine Learning), MIT Press, Cambridge (USA), 2016

[2] Michael A. Nielsen: Neural Networks and Deep Learning, Determination Press, 2015

[3] A. Luckow, M. Cook, N. Ashcraft, E. Weill, E. Djerekarov und B. Vorster: Deep learning in the automotive industry: Applications and tools, in Big Data, IEEE International Conference on. IEEE, 2016

[4] F. Falcini, G. Lami, A. M. Costanza: Deep Learning in Automotive Software, IEEE Computer Society, 2017

[5] Eric P. Xing, Qirong Ho, Pengtao Xie, and Dai Wei: Strategies and principles of distributed machine learning on big data. Engineering (179-195), 2016

[6] F. Niu, B. Recht, C. Re, and S. J. Wright. HOGWILD!: A Lock-Free Approach to Parallelizing Stochastic Gradient Descent, ArXiv e-prints, 2011

[7]S Lee,JK Kim, X Zheng, Q Ho, GA Gibson, EP Xin g: On model parallelization and scheduling strategies for distributed machine learning, in Proceedings of the Neural Information Processing Systems, 2014

[8] Iandola, F.N., Moskewicz, M.W., Ashraf, K., Han, S., Dally, W.J., Keutzer, K.: Squeezenet: alexnet-level accuracy with 50x fewer parameters and 1mb model size, ArXiv preprint (arXiv:1602.07360), 2016

[9] Adam Grzywaczewski: Training AI for Self-Driving Vehicles: the Challenge of Scale, NVIDIA Developer Blog, verfügbar unter: https://devblogs.nvidia.com/training-self-driving-vehicles-challenge-scale, 2017

[10] L.-N. Lundbæk, D. J. Beutel, M. Huth, S. Jackson, L. Kirk, S. Schwerin: Practical Proof of Kernel Work & Distributed Adaptiveness, verfügbar unter: https://www.xain.io/pdf/XAIN_Yellow_Paper.pdf, 2017

[11] Codie Thompson: Why driverless cars will be safer than human drivers, verfügbar unter: http://www.businessinsider.de/why-driverless-cars-will-be-safer-than-human-drivers-2016-11?r=US&IR=T, 2016

KONTAKT

Felix Reichel
XAIN AG
Schmiedestraße 2A, 15745 Wildau
Email: felix.reichel@xain.io