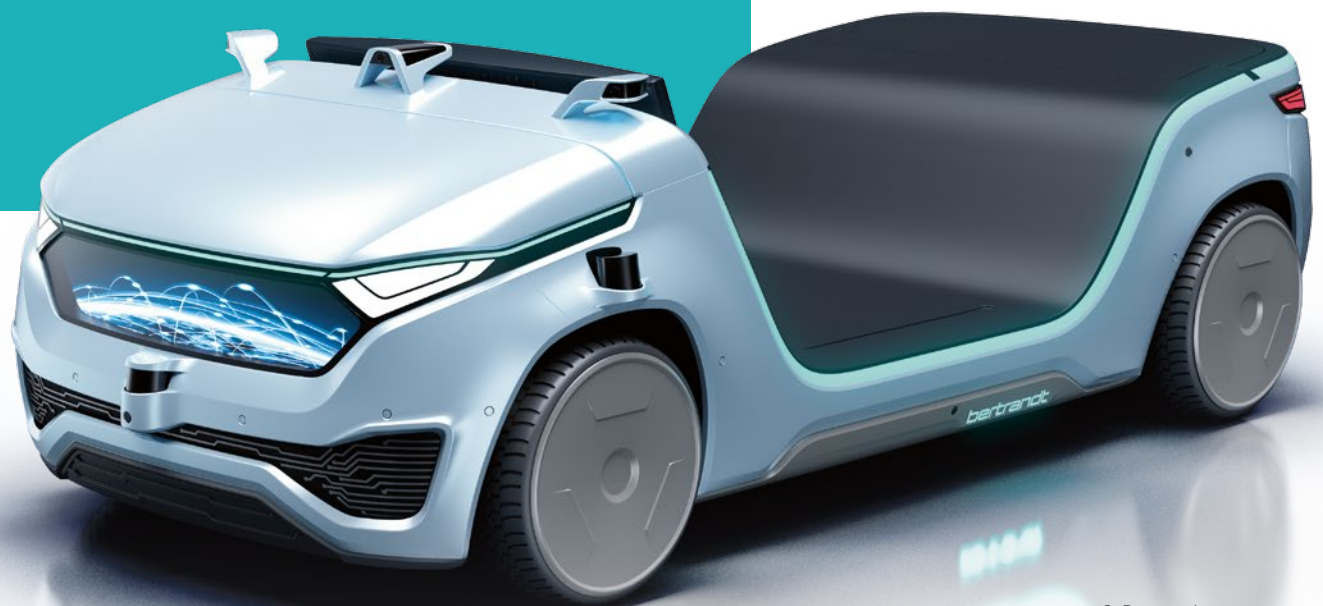


Maschinelles Lernen für das automatisierte Fahren



© Bertrand

Zu den Kernthemen der Mobilität der Zukunft zählt das automatisierte und das autonome Fahren in den SAE-Levels 3 bis 5. Bertrandt zeigt in seinem Innovationsprojekt „Park and Charge“, wie wichtig die Umfelderkennung und die präzise Trajektorienplanung mithilfe künstlicher Intelligenz ist. Das maschinelle Lernen wird dabei genutzt, um Lokalisierung, Vernetzung und Cloud-Anwendungen zu verbessern.

AKTUELLER STATUS DER OBJEKTERKENNUNG

Bereits heute werden autonome Fahr-funktionen in Serie zum Einsatz gebracht. Dabei darf das Fahrzeug den Fahrer in verschiedenen Situationen ersetzen, muss dafür aber selbst in sehr kurzen Zeiträumen verlässliche Entscheidungen treffen. Dieser Umstand macht den Einsatz von künstlicher Intelligenz (KI) und insbesondere des maschinellen Lernens (ML) unverzichtbar. Die Entscheidungsfindung beruht auf trainierten Deep-Learning-Algorithmen. Als Eingangsgrößen dienen Messungen von zahlreichen Sen-

soren wie Kameras, Radare und Lidare. Die Hauptaufgabe der Algorithmen ist, durch Objekterkennung die Umgebung in der Fahrsituation zu rekonstruieren und diese Information für die Fahrzeug-systeme zur Verfügung zu stellen.

GRUNDLAGEN VON MASCHINELLEM LERNEN UND NEURONALEN NETZEN

Im allgemeinen Sprachgebrauch haben sich für den Begriff KI eine Vielzahl von Interpretationen etabliert. Dieses Teilgebiet der Informatik befasst sich mit dem selbstständigen, automatischen Lösen

von Problemen [1]. Es wird dabei zwischen der starken und schwachen KI unterschieden. Während sich die starke KI damit beschäftigt, den Menschen und seine Vorgehensweise beziehungsweise sein Denken (zum Beispiel Bewusstsein) zu imitieren, beschränkt sich die schwache KI darauf, gezielt Algorithmen zu erlernen, um eine klar umrissene Problemstellung zu lösen [2].

Der heutzutage prominenteste Vertreter aus dem Bereich der KI ist das ML, was der schwachen KI zuzuordnen ist. Viele führende Experten auf dem Gebiet der KI handeln ML als eine der wichtigsten Technologien des aktuellen Zeitalters

AUTOREN



Dipl.-Ing. Peter Schiekofer
ist Director Advanced Engineering and Autonomous Driving des Bertrandt-Konzerns in Ehningen.



Dr. Yusuf Erdogan
ist Leiter Data Science und Entwicklung bei Bertrandt in Frankfurt am Main.



Dipl.-Phys. Stefan Schindler
ist Data Scientist im Bereich Machine Learning bei Bertrandt in Frankfurt am Main.



Markus Wendl, M. Eng.
ist Lead Engineer im ADAS-Bereich und Teilprojektleiter Softwareentwicklung für die autonomen Fahrfunktionen im Projekt „Park and Charge“ bei Bertrandt in Köln.

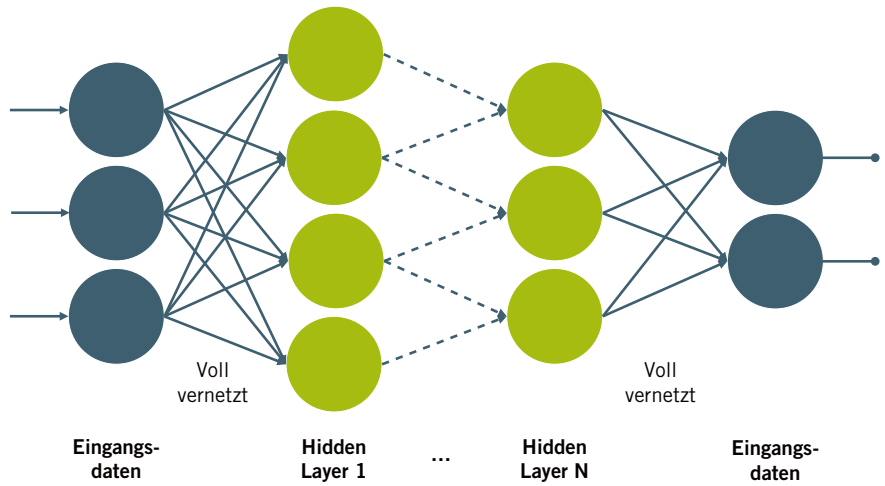


BILD 1 Illustration eines neuronalen Netzes mit verschiedenen Lagen (© Bertrandt)

gie, ein Problem zu lösen, indem man mit Belohnung und Bestrafung entsprechende Anreize schafft. Innerhalb der genannten Arten findet das Supervised Learning am häufigsten Anwendung. In der Welt der Open-Source-Werkzeuge stehen hier zahlreiche Umgebungen mit vielen verschiedenen Architekturen zur Verfügung. Eine spezielle Gruppe dieser Architekturen bilden die neuronalen Netze, die von ihrem Aufbau den Neuronen des menschlichen Gehirns nachempfunden sind. Die Neuronen sind dabei in Schichten angeordnet, die unterschiedliche Funktionen erfüllen. Jede Schicht besteht aus sehr vielen einzelnen Neuronen, wobei die Anzahl je nach Architektur unterschiedlich ist. Jedes Neuron einer Schicht besitzt eine der Architektur zugrundeliegende Vernetzung mit den Neuronen der darauffolgenden Schicht. Ist jedes Neuron einer Schicht mit allen Neuronen der Folgeschicht verbunden, spricht man von einer Fully Connected Layer (FCL). Da typischerweise nur die erste und letzte Schicht für den In- beziehungsweise Output angesprochen werden, bezeichnet man die Zwischenschichten als versteckte Schichten (Hidden Layers), **BILD 1**.

Zur Objekterkennung auf Bildern haben sich Faltungsnetzwerke (Convolution Neural Net, CNN) bewährt. Hier wird das Bild der Eingabeschicht zugeführt und in seine Pixel-Information zerlegt. Durch Faltung wird nun die Dimension der Informationen stufenweise reduziert und gleichzeitig Merkmale (Features) erlernt, die in den Neuronen des Netzes gespeichert werden. Je nach Architektur, zum Beispiel bei einem Faster-R-CNN [4], durchläuft das

Bild Hunderte dieser Schichten im neuronalen Netz, was zur Begrifflichkeit des tiefen Lernens (Deep Learning) führt. In der Ausgabeschicht wird dem Netz nun gezeigt, welches Objekt sich wo auf dem Bild befindet. Diese Information wird mit dem Ergebnis des Netzwerks verglichen und die Abweichungen anschließend rückwärts in das Netz propagiert, um die Neuronen entsprechend einzustellen. Hat man dem Netzwerk genügend Objekte auf den Bildern gezeigt und so die Abweichungen reduziert, erlernt das Netz diese Objekte eigenständig zu lokalisieren und zu erkennen.

UMSETZUNG DER TRAJEKTORIENPLANUNG FÜR AUTONOME FAHRSYSTEME

Für das „Park and Charge“-Projekt wurde ein Trajektorienplanungs-Algorithmus entwickelt, der einen validen Pfad zu einem beliebigen Punkt mit Zielausrichtung berechnet. Dabei werden Hindernisse aus der lokalen Umfeldkarte, die durch Lidar- und Ultraschallsensoren erzeugt wird, berücksichtigt und umfahren. Zusätzlich können hochgenau vermessene Routen hinterlegt werden, die von dem Algorithmus anschließend für die Berechnung der Trajektorie zum Zielpunkt bevorzugt werden.

Die Basis des Algorithmus bildet ein modifizierter, geschachtelter A*-Algorithmus mit einem mathematischen Vergleichsverfahren, das die Rechenzeit stark verringert und die Endausrichtung des berechneten Pfads definiert.

Für die Berechnung eines befahrbaren Pfads werden fahrzeugspezifische Daten benötigt, die unter anderem die fahrdynamischen Grenzen des Fahrzeugs

[3]. Grundsätzlich unterscheidet man zwischen drei Arten des ML:

- Supervised Learning: Dem Algorithmus werden Daten und das erwartete Ergebnis gezeigt. Dabei erlernt der Algorithmus, diese beiden Größen miteinander zu verknüpfen. Ist das Lernen abgeschlossen, kann der Algorithmus selbstständig ein Ergebnis zu unbekanntem Daten finden.
- Unsupervised Learning: Hier lernt der Algorithmus nur anhand der Daten, Muster zu erkennen und diese zu klassifizieren.
- Reinforcement Learning: Der Algorithmus lernt hierbei die optimale Strategie.

beschreiben. Die Berechnung der möglichen Pfade wird nicht durch den Algorithmus begrenzt, sondern unter anderem durch die Agilität des Fahrzeugs limitiert. Es besteht somit zum Beispiel auch die Option, eine agile Zieltrajektorie für einen Gabelstapler zu berechnen.

DAS FAHRSCENARIO FÜR „PARK AND CHARGE“

In dem angestrebten Fahrscenario wird dem Fahrzeug eine empfohlene Route zur Parkmöglichkeit über die Bertrandt-Cloud übermittelt. Mit dem entwickelten Algorithmus besteht die Möglichkeit, die empfohlene Route in Echtzeit anhand der Umfelddaten zu optimieren und anschließend zur Zielvorgabe zu folgen, **BILD 2**.

Sobald das Fahrzeug die Zielvorgabe erreicht hat, wird die vorgegebene Parklücke anhand der Umfeldsensorik validiert und der Einparkvorgang in ein bis drei Zügen berechnet. Die Art und Ausrichtung der Parklücke spielt dabei keine Rolle. Ein exemplarischer Einparkvorgang ist in **BILD 2** anhand simulierter Daten dargestellt. Die Vorgabe der Route wird bei der Kartendarstellung dabei als Vertiefung und die Wände und Fahrzeuge als Erhöhung dargestellt.

Daraus geht hervor, in welchen Bereichen sich das Fahrzeug selbstständig, ohne externe Vorgaben, bewegen kann. Des Weiteren ist es möglich, Zonen zu definieren, die nicht befahren werden dürfen. Hierzu zählen zum Beispiel die Gegenfahrbahn, Fußgängerüberwege oder Bürgersteige.

Zurzeit wird eine Methodik entwickelt, die es schafft, mithilfe eines kontinuierlichen Abgleichs der Umfeldsensorik und der erstellten lokalen Karte mögliche Kollisionen vorherzusehen. Dabei wird der Abgleich zwischen der extrapolierten Fahrzeugbewegung und den dynamischen sowie statischen Umfelddaten ausgeführt.

ANTRAINIEREN DER OBJEKTERKENNUNG

Für die kamerabasierte Objekterkennung wurde ein Ansatz verfolgt, mit dem der Aufwand zur Generierung von Trainingsbildern deutlich reduziert werden kann [5]. Diese Methodik kann für die Erkennung von Objekten verwendet werden, bei denen das Erscheinungsbild nicht allzu stark

variiert. Gut geeignete Objekte sind zum Beispiel Verkehrsschilder und Ampeln, schlecht geeignete sind Fahrzeuge, Tiere, Fußgänger oder Fahrradfahrer.

Die Objekte werden zunächst aus verschiedenen Perspektiven, aus denen sie im Anwendungsfall erkannt werden sollen, vor einem homogenen Hintergrund aufgenommen. Anschließend werden die Objekte aus einer repräsentativen Menge der aufgenommenen Bilder mit entsprechenden Bildverarbeitungsalgorithmen automatisiert ausgeschnitten. Die ausgeschnittenen Objekte werden abschließend in verschiedene Hintergrundbilder der Zielumgebung (zum Beispiel Straßenszenen) eingefügt.

Hierbei wurde mit verschiedenen Blending-Verfahren zur Verbesserung der Ergebnisse experimentiert.

Um verschiedene Umwelteinflüsse zu simulieren und generell eine Überanpassung (Overfitting) beim Trainieren zu vermeiden, werden das Aussehen der Objekte sowie der Hintergrund zusätzlich mit Data Augmentation variiert. In **BILD 3** werden die zu erkennenden Objekte dargestellt, die im Fahrbetrieb vom Fahrzeug erkannt werden müssen. Die Objekte dienen zum Beispiel als Landmarken-Referenzpunkte, sodass die Positionsbestimmung des Fahrzeugs verbessert werden kann. Wird die verwendete IoT-Ampel durch das Backend auf Grün geschaltet, soll das Fahrzeug den Wechsel der Ampelphase erkennen können und die Fahrt fortsetzen, sofern die Umgebungsfreigabe erfolgt ist.

In **BILD 3** (unten) ist ersichtlich, dass die Relationen und Positionen der eingefügten Objekte zu denen im Hintergrund nicht realistisch dargestellt werden. Aufgrund der Struktur von Faltnetzwerken sind für das Training beziehungsweise eine hohe Trefferquote lediglich lokale Merkmale relevant, nicht jedoch globale Bildzusammenhänge wie beispielsweise die Position im Gesamtbild.

Das Training des neuronalen Netzes wurde zunächst nur mit künstlich erzeugten Daten durchgeführt. Anschließend wurde es mit einer geringen Menge an realen Daten aus eigenen Fahrzeugaufnahmen validiert und optimiert.

ZEIT- UND KOSTENERSPARNIS DURCH SYNTHETISCH ERZEUGTE DATEN

Somit wurde ein Ansatz zur Erstellung von synthetisch erzeugten Trainingsdaten angewandt, mit dem zeit- und kostenaufwendiges Einfahren von realen Daten und anschließendes Annotieren stark minimiert werden können. Durch den beschriebenen Ansatz können unter anderem zeit- und kostengünstig annotierte Trainingsbilder für neue Schilder im Straßenverkehr generiert werden.

Nach dem Training des neuronalen Netzes wurde die Inferenzzeit (Laufzeit) durch eine niedrige Quantisierung der Gewichte (von 32-bit-Float nach 8-bit-Integer) für die Ziel-Hardware (Nvidia Drive PX2 Autochauffeur) und die Ein-

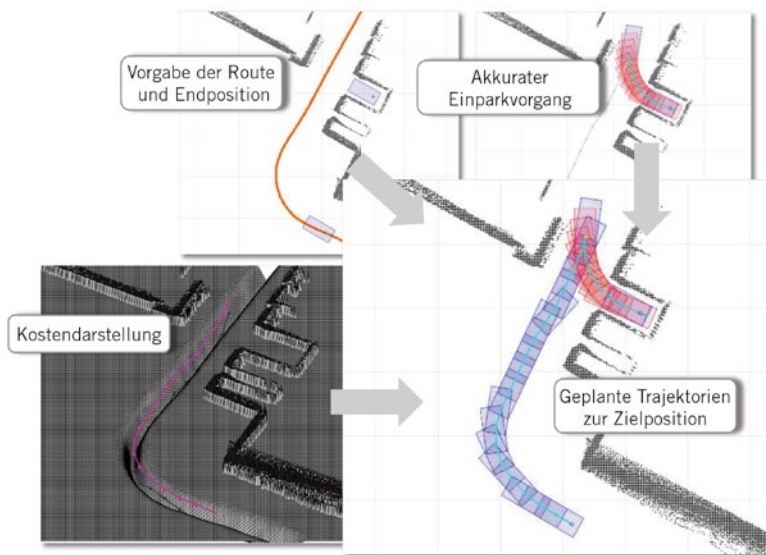


BILD 2 Algorithmus für die Trajektorienplanung (© Bertrandt)

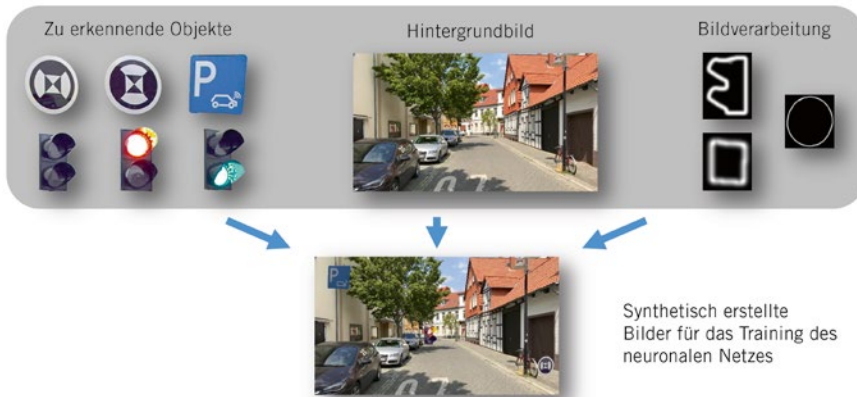


BILD 3 Grafische Darstellung des Ansatzes zur Generierung von künstlichen Trainingsbildern (© Bertrandt)

haltung der notwendigen Laufzeitkriterien optimiert.

AUTOMATISCHE GENERIERUNG VON GROUND-TRUTH-DATEN

Aktuell ist bei Bertrandt ein Produkt in der Fertigstellung, das zukünftig für die automatisierte Generierung von Ground-Truth-Informationen aus Sensordaten verwendet werden kann. Diese Software basiert auf Computer-Vision-Algorithmen und vortrainierten neuronalen Netzen. Die neuronalen Netze erkennen selbstständig auf Kamera-Frames unterschiedliche Objekte und markieren diese mit hoher Genauigkeit.

Die Computer-Vision-Algorithmen werden genutzt, um die markierten Objekte über eine Abfolge von mehreren Frames zu verfolgen. In Kombination können diese beiden Algorithmen entweder den Nutzer beim manuellen Annotieren der Kameradaten unterstützen oder den kompletten Annotationsprozess vollständig automatisieren.

Das Bertrandt-Labeling-Werkzeug wird in diesem Jahr als Software-as-a-Service (SaaS) zahlreichen Kunden zur Verfügung gestellt. Darüber hinaus wird die Software genutzt, um die Trainingsdaten für die in diesem Artikel vorgestellten neuronalen Netze zu generieren. In weiteren Entwicklungsphasen wird die Funktionsweise auf die Objekterkennung aus Lidardaten erweitert. Das für die Entwicklung dieses Produkts benötigte Know-how, insbesondere für die intelligenten Algorithmen, baut auf den Erkenntnissen aus dem genannten Innovationsprojekt „Park and Charge“ auf.

OBJEKTERKENNUNG MIT KAMERA- UND LIDARDATEN

Im weiteren Verlauf dieses Projekts soll mittels ML eine kamera- und lidarbasierte Objekterkennung entwickelt und für die Ziel-Hardware optimiert werden. Hierfür wird auf aktuelle Ansätze aus der Forschung zurückgegriffen, bei denen ein neuronales Netz sowohl die Merkmale der zu erkennenden Objekte in Kamera- und Lidardaten als auch die Fusion dieser Daten lernt. Eine kamera- und lidarbasierte Objekterkennung soll die Vorteile der einzelnen Sensoren vereinen und die jeweiligen Nachteile der Sensoren kompensieren. Das Projekt wird von einem interdisziplinären Team verschiedener Bertrandt-Standorte bearbeitet.

LITERATURHINWEISE

- [1] Michalski, R. S.; Carbonell, J. G.; Mitchell, T. M.: Machine Learning. Berlin, Heidelberg: Springer-Verlag, 1983
- [2] Goertzel, B.: Toward a Formal Characterization of Real-World General Intelligence. Tagungsband, 3rd Conference on Artificial General Intelligence, Lugano, 2010, Nr. 10.2991/agi.2010.17
- [3] Brynjolfsson, E.; McAfee, A.: The business of artificial intelligence. In: Harvard Business Review, Nr. 7, 2017
- [4] Ren, S.; He, K.; Girshick, R.; Sun, J.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In: IEEE Trans. Pattern Anal. Mach. Intell., Bd. 39, Nr. 6 (Juni 2017), S. 1137-1149
- [5] Dwivedi, D.; Misra, I.; Hebert, M.: Cut, Paste and Learn: Surprisingly Easy Synthesis for Instance Detection. In: arXiv: 1708.01642 (2017)



READ THE ENGLISH E-MAGAZINE

Test now for 30 days free of charge:
www.atz-worldwide.com