



HOCHSCHULE RUHR WEST  
UNIVERSITY OF APPLIED SCIENCES

# **Bachelorarbeit**

## **Aspekte der Datensammlung beim Einsatz von Convolutional Neural Network für das autonome Fahren**

### **Bachelorarbeit**

Studiengang Betriebswirtschaftslehre – Industrielles  
Dienstleistungsmanagement (Bachelor of Science B.Sc.)  
der Hochschule Ruhr West

**Arda Kaya**

**Matrikelnr. 10010776**

Erstprüfer: Prof. Dr. Christian Weiß

Zweitprüferin: Prof. Dr.-Ing Karla Ohler Martins

Mülheim an der Ruhr, 28. Juni 2022

## **Abstract**

Die vorliegende Bachelorarbeit thematisiert, ob ein Convolutional Neural Network für das autonome Fahren geeignet ist und wie der gesamte Datenprozess für das Convolutional Neural Network Modell aussehen könnte. Dazu wurden mehrere Forschungsfragen gestellt.

Um die Forschungsfragen zu beantworten, wurde eine quantitative Literaturarbeit mit aktuellen wissenschaftlichen Kenntnissen durchgeführt. Die Bachelorarbeit ist in der deutschen Sprache geschrieben, aus diesem Grund orientiert sich die Bachelorarbeit an die deutschsprachige Bevölkerung.

Die quantitative Literaturarbeit zeigte, dass die größten Herausforderungen für das Convolutional Neural Network Training in der Datenspeicherung und der Datenverarbeitung liegen. Ebenfalls zeigte die quantitative Literaturarbeit, dass das Convolutional Neural Network durch die qualifizierte Wahl von Methoden und Daten eine bessere Erfolgsquote als das menschliche Verhalten für das autonome Fahren nachweist.

Dies zeigt, dass der Einsatz von Convolutional Neural Network Modellen mit einem optimalen Datenprozess und der Dateninfrastruktur für das autonome Fahren empfehlenswert sein kann.

## Vorwort

Die vorliegende Bachelorarbeit behandelt mehrere Forschungsfragen, wie zum Beispiel „warum sind Convolutional Neural Networks geeignet für das autonome Fahren?“, oder „woher werden die Daten für das autonome Fahren gewonnen?“, die intensiv beantwortet werden sollen.

Dabei wurde die Forschung für diese Bachelorarbeit in Form einer quantitativen Literaturarbeit durchgeführt.

Diese Bachelorarbeit verfasste ich als Abschlussarbeit meines Studiums der Betriebswirtschaftslehre – Industrielles Dienstleistungsmanagement in Mülheim an der Ruhr an der Hochschule Ruhr West.

Zusammen mit meinem Betreuer, Herrn Christian Weiß, entwickelte ich die Forschungsfragen für diese Bachelorarbeit. Durch seine fachlichen Kenntnisse im Bereich des maschinellen Lernens, half er mir den Einstieg in die Thematik zu finden und Ideen für mögliche Forschungsfragen zu entwickeln. Gewisse Themenkomplexe gestalteten sich anfänglich als schwierig, dennoch konnten alle wesentlichen Fragestellungen erläutert und beantwortet werden.

Aus diesem Grund möchte ich meinem Betreuer für seine intensive Unterstützung während der Durchführung meiner Bachelorarbeit danken.

Zusätzlich möchte ich mich bei Sascha Müller und Alina Witte für das Korrekturlesen dieser Abschlussarbeit bedanken.

Ich wünsche Ihnen viel Freude beim Lesen dieser Abschlussarbeit.

Arda Kaya

# Inhaltsverzeichnis

<b>Abstract</b> .....	<b>I</b>
<b>Vorwort</b> .....	<b>II</b>
<b>Inhaltsverzeichnis</b> .....	<b>III</b>
<b>Abbildungsverzeichnis</b> .....	<b>V</b>
<b>Tabellenverzeichnis</b> .....	<b>VIII</b>
<b>Abkürzungsverzeichnis</b> .....	<b>IX</b>
<b>Symbolverzeichnis</b> .....	<b>X</b>
<b>Glossar</b> .....	<b>XI</b>
<b>1 Einleitung</b> .....	<b>1</b>
<b>2 Maschinelles Lernen</b> .....	<b>5</b>
2.1 Einführung maschinelles Lernen .....	5
2.2 Künstliche Neuronale Netzwerke.....	11
2.3 Convolutional Neural Network und seine Komponenten .....	16
2.4 Convolutional Neural Network und seine Wichtigkeit im Bereich des autonomen Fahrens .....	21
2.5 Zielbild bei der Klassifikation Muster- und Bilderkennung im öffentlichen Straßenverkehr.....	23
<b>3 Daten für den Einsatz von Convolutional Neural Network und für das autonome Fahren</b> .....	<b>26</b>
3.1 Bedeutung von Daten und die Datenstruktur für das Convolutional Neural Network.....	26
3.2 Daten und die Dateninfrastruktur für das autonome Fahren .....	27
3.2.1 Datengewinnung .....	30
3.2.2 Datenspeicherung und die Herausforderungen .....	36
3.2.3 Datenverarbeitung und die Herausforderungen .....	39
3.3 Potenzial für ein Convolutional Neural Network Modell mit den Datensätzen von GTSRB .....	42
3.4 Anwendungsbeispiel: Der gesamte Datenprozess für das Convolutional Neural Network.....	45
<b>4 Forschungsstand, Problematik, weitere Herausforderungen und Ausblick</b> .....	<b>48</b>
<b>5 Fazit</b> .....	<b>52</b>

<b>Literaturverzeichnis .....</b>	<b>54</b>
<b>Anhang A: .....</b>	<b>66</b>
<b>Stichwortverzeichnis .....</b>	<b>67</b>

## Abbildungsverzeichnis

Abbildung 1: Entwicklung der Anzahl der angemeldeten Patente im Bereich autonomes Fahren weltweit (Analog zu: <a href="https://de.statista.com/statistik/daten/studie/1062558/umfrage/anzahl-der-patente-im-bereich-autonomen-fahren-weltweit/">https://de.statista.com/statistik/daten/studie/1062558/umfrage/anzahl-der-patente-im-bereich-autonomen-fahren-weltweit/</a> , eigene Darstellung, online abgerufen am 30. April 2022, 16:30 Uhr).....	2
Abbildung 2: Akzeptanzprobleme für das autonome Fahren: März 2019; 500 Befragte Personen; ab 18 Jahren (eigene Darstellung) (Analog zu: <a href="https://de.statista.com/statistik/daten/studie/270612/umfrage/nachteile-von-autonomen-fahrzeugen/">https://de.statista.com/statistik/daten/studie/270612/umfrage/nachteile-von-autonomen-fahrzeugen/</a> , online abgerufen am 30. April 2022, 13:50 Uhr) .....	3
Abbildung 3: Programmierung auf klassische Weise (links) und durch maschinelles Lernen (rechts) (Sigurd Schacht, Blockchain und maschinelles Lernen - Wie das maschinelle Lernen und die Distributed-Ledger-Technologie voneinander profitieren, 2019, S. 90).....	5
Abbildung 4: Hierbei handelt es sich um eine grafische Darstellung des Ansatzes zur Generierung von künstlichen Trainingsbildern ( Schiefkofer, Erdogan, Schindler, & Wendl, Maschinelles Lernen für das automatisierte Fahren, 2019, S. 51).....	6
Abbildung 5: Einteilung und Verfahren des maschinellen Lernens (Breyer-Mayländer, Industrie 4.0 bei Hidden Champions, 2022, S.173) .....	7
Abbildung 6: Eigene Darstellung: Überwachtes Lernen: Modell trainiert anhand von bekannten Daten und Beispielen. Es gibt eine klare Zielvariable, die vorhergesagt wird. (Quelle: Analog zu <a href="https://datasolut.com/wiki/supervised-learning/">https://datasolut.com/wiki/supervised-learning/</a> , online abgerufen am 10.05.202) .....	8
Abbildung 7: Bestandteile eines künstlichen neuronalen Netzes (Eigene Darstellung, analog zu: Wennker, Künstliche Intelligenz in der Praxis - Anwendung in Unternehmen und Branchen: KI wettbewerbs- und zukunftsorientiertes einsetzen 2020).....	11
Abbildung 8: Bestandteile eines Neurons (Wennker, Künstliche Intelligenz in der Praxis - Anwendung in Unternehmen und Branchen: KI wettbewerbs- und zukunftsorientiertes einsetzen 2020).....	12
Abbildung 9: Formales Neuron mit den Eingaben der Aktivierungsfunktion (Ertel, Grundkurs Künstliche Intelligenz, 2008) .....	13
Abbildung 10: Funktionsverlauf der Aktivierungsfunktion im Intervall -10 bis 10 (Eigene Darstellung, analog zu: Neiden, Analyse hochdimensionaler MALDI-Daten mittels Convolutional Neural Networks, 2017) .....	14
Abbildung 11: Grafische Abbildung vom lokalen und globalen Fehlerminimum beim Gradientenabstiegsverfahren (Sauer, Wiesche, & Krimmling, Management digitaler Plattformen – Datengetriebene Geschäftsmodelle für Mobilität in der Smart City 2018) .....	15
Abbildung 12: Convolutional Layer (Online: <a href="https://databasecamp.de/ki/convolutional-neural-network">https://databasecamp.de/ki/convolutional-neural-network</a> ; Abgerufen am 25.04.2022; 19:29 Uhr) .....	17
Abbildung 13: Max Pooling mit 2x2-Filter und Stride=2 (Selle, Künstliche Neuronale Netzwerke und Deep Learning, 2018) .....	18

Abbildung 14: Pooling Layer übernimmt die 3x3 Matrix der Convolutional Layer als Input – Dimensionsreduktion (Quelle: <a href="https://databasecamp.de/ki/convolutional-neural-network">https://databasecamp.de/ki/convolutional-neural-network</a> ; online abgerufen am 26. April, 22:00 Uhr).....	19
Abbildung 15: Allgemeiner Aufbau eines CNN in Anlehnung an MathWorks (Neiden, Analyse hochdimensionaler MALDI-Daten mittels Convolutional Neural Networks, 2017).....	20
Abbildung 16: Aufbau - vereinfacht dargestellt (Online abgerufen am 28.05.22, 21:10 Uhr : <a href="https://de.mathworks.com/discovery/convolutional-neural-network-matlab.html">https://de.mathworks.com/discovery/convolutional-neural-network-matlab.html</a> ; MathWorks, 2021) .....	21
Abbildung 17: Die fünf Stufen des automatisierten Fahrens (Online abgerufen: <a href="https://blog.doubleslash.de/vision-zero-durch-autonomes-fahren/">https://blog.doubleslash.de/vision-zero-durch-autonomes-fahren/</a> ; am 19.06.2022; 18:40 Uhr).....	23
Abbildung 18: Eigenes bearbeitetes Foto – Zielbild CNN im öffentlichen Straßenverkehr Mülheim (1) [Analog zu: <a href="https://artemoppermann.com/de/ki-use-case-computer-vision/">https://artemoppermann.com/de/ki-use-case-computer-vision/</a> ].....	24
Abbildung 19: Eigenes bearbeitetes Foto – Zielbild CNN im öffentlichen Straßenverkehr Mülheim (2) [Analog zu: <a href="https://artemoppermann.com/de/ki-use-case-computer-vision/">https://artemoppermann.com/de/ki-use-case-computer-vision/</a> ].....	25
Abbildung 20: Professionelles Rechenzentrum [Online abgerufen: <a href="https://www.it-daily.net/it-management/data-center/wie-sieht-ein-modernes-monitoring-der-rechenzentrums-infrastruktur-aus">https://www.it-daily.net/it-management/data-center/wie-sieht-ein-modernes-monitoring-der-rechenzentrums-infrastruktur-aus</a> , am 10.05.22; 13:30 Uhr].....	29
Abbildung 21: Der Weg der Daten von der Straße ins Rechenzentrum und in die Cloud (eigene Darstellung) (Analog zu: Nguyen, Autonomes Fahren beginnt im Rechenzentrum, 2020).....	30
Abbildung 22: Autonomes Fahren im Urbanen Raum - Innovative Verarbeitung von Sensordaten für komplexe Verkehrssituationen (Volkswagen, <a href="https://www.volkswagenag.com/de/news/stories/2019/12/the-champions-league-of-autonomous-driving.html#">https://www.volkswagenag.com/de/news/stories/2019/12/the-champions-league-of-autonomous-driving.html#</a> , online abgerufen am 08. Mai.....	34
Abbildung 23: Nur wenige Teile der Daten können real sein [Online abgerufen: <a href="https://www.bmwk.de/Redaktion/DE/Downloads/W/ws5-praesentation-ki-und-daten.pdf?__blob=publicationFile&amp;v=4">https://www.bmwk.de/Redaktion/DE/Downloads/W/ws5-praesentation-ki-und-daten.pdf?__blob=publicationFile&amp;v=4</a> ; am 17.05.2022, 20:55 Uhr].....	35
Abbildung 24: Energiebedarf pro Jahr [Mrd. kWh/a] (online abgerufen: <a href="https://www.bundestag.de/resource/blob/863850/423c11968fcb5c9995e9ef9090edf9e6/WD-8-070-21-pdf-data.pdf">https://www.bundestag.de/resource/blob/863850/423c11968fcb5c9995e9ef9090edf9e6/WD-8-070-21-pdf-data.pdf</a> ; am 20.06.2022; 23:10 Uhr) .....	37
Abbildung 25: Speicheroptionen für die KI-Datenpipeline (Ziegler, Krämer, & Haustein, KI und Big-Data-Management für autonomes Fahren, 2020, S.43).....	38
Abbildung 26: Datensätze mit mehreren Simulationen (Berghoff, Erkennung und Klassifizierung von Verkehrsschildern unter schwierigen Verhältnissen mit Convolutional Neural Network (CNN), 2020, S.5) .....	40
Abbildung 27: Funktionsweise CNN anhand von deutschen Verkehrsschildern (Online abgerufen: <a href="http://www.durak-bolat.de/RSRwithCNN.html">http://www.durak-bolat.de/RSRwithCNN.html</a> ; am 10.Juni 2022, 14:50 Uhr) .....	43
Abbildung 28: GTSRB Results 2011 ( <a href="https://benchmark.ini.rub.de/gtsrb_results.html">https://benchmark.ini.rub.de/gtsrb_results.html</a> , online abgerufen am 29.05.2022, 18:20 Uhr).....	43

- Abbildung 29: Der Weg zum Modell [Online abgerufen:  
[https://www.bmwk.de/Redaktion/DE/Downloads/W/ws5-  
praesentation-ki-und-daten.pdf?\\_\\_blob=publicationFile&v=4;](https://www.bmwk.de/Redaktion/DE/Downloads/W/ws5-<br/>praesentation-ki-und-daten.pdf?__blob=publicationFile&v=4;)  
Bundesministerium für Wirtschaft und Energie, 2020, S. 2) ..... 45
- Abbildung 30: Anteil der Fahrzeuge mit Automatisierungsfunktionen an den  
Neuzulassungen (Prognos-Forschungsinstitut. Einführung von  
Automatisierungsfunktion in der Pkw-Flotte, 2018, S.15)..... 50



## Tabellenverzeichnis

Tabelle 1: Strukturierte Daten (Betriebswirtschaftliche KI-Anwendungen, Herrmann, 2021, S.33) .....	27
Tabelle 2: Vor- und Nachteile der verschiedenen Datenspeichermöglichkeiten (eigene Darstellung) [online abgerufen: <a href="https://www.cio.de/a/aktuelle-datensicherungsmethoden-im-ueberblick,3564528">https://www.cio.de/a/aktuelle-datensicherungsmethoden-im-ueberblick,3564528</a> ;Püschel, 2017).....	39

## Abkürzungsverzeichnis

ML	Machine Learning
KI	Künstliche Intelligenz
CNN	Convolutional Neural Network
LIST	List Processing (Listen-Verarbeitung)
ReLu	Rectified Linear Unit Function
SQL	Structured Query Language
IoT	Internet of Things
GPU	Graphic Processing Unit
HDD	Hard-Disk-Drive
SSD	Solid-State-Drives
SDS	Special Direct System
ETL	Extract-Transform-Load
NVMe	Non-Volatile Memory Express
DL	Data language
RTK	Real Time Kinematic
GNSS	Globales Navigationssatellitensystem
GTSRB	German Traffic Sign Recognition Benchmark

## Symbolverzeichnis

$\Sigma$	Summe
$x$	gewichtete Ausgabewerte
$f(x)$	Funktionswert
$\sum_{j=1}^n$	Summe von $i = 1$ bis $n$
$<$	kleiner als
$\geq$	größer gleich
$e$	Eulersche Zahl
$\sigma$	Sigma
$K$	Körper (Algebra)

## Glossar

<i>Bias</i>	Es ist die Verzerrung einer Schätzfunktion in der Schätztheorie. Es ist eine Kennzahl, welche die systematische Über- oder Unterschätzung der Schätzfunktion qualifiziert. (Statista, 2019)
<i>Datenbank</i>	Eine Datenbank ist eine Sammlung von Daten oder Informationen, die dann zu einer logischen Einheit verknüpft wird. Die Datenbank wird elektronisch erfasst. (Preuss, 2017, S. 35)
<i>Künstliche Intelligenz</i>	Die KI ist der Versuch, das menschliche Denken und Lernen auf Computer zu übertragen. Somit soll ihm damit Intelligenz verliehen werden, wodurch die KI eigenständig lernt, Probleme zu lösen. Unter der KI fällt das maschinelle Lernen und Deep Learning an. (Witte, 2020, S. 243)
<i>Internet of Things</i>	Das Internet of Things (IoT) ist die Definition für das Netzwerk physischer Objekte. Die Objekte werden mit die mit Software, Sensoren und anderer Technik ausgestattet sind. Diese werden dann mit anderen Geräten über das Internet vernetzt, somit ist ein Datenaustausch zwischen den Objekten möglich. (Oracle University, 2020)
<i>Trainingsdaten</i>	Trainingsdaten im Bereich maschinelles Lernen sind Datensätze, die für das Lernen der erkannten Muster und Zusammenhänge in den Daten verwendet werden. (Freedden & Rummel, 2017, S. 557)

# 1 Einleitung

Maschinelles Lernen ist für autonomes Fahren zwingend erforderlich. Jedes Fahrzeug, welches sich im öffentlichen Straßenverkehr autonom bewegt, ist eine potenzielle Gefahr für andere Teilnehmer. Zu diesen gehören zum Beispiel weitere Autofahrer, Fußgänger und Fahrradfahrer. Es ist daher wichtig, dass der vollständige Datenprozess und die notwendige Technik für die Objekterkennung im Fahrzeug verstanden werden. (Krol, 2021, S. 9-13)

Aus diesem Grund ergibt sich die Besonderheit für das Convolutional Neural Network Modell und welchen Einfluss es nach ausführlicher Ausarbeitung für das autonome Fahren hat. Insbesondere die Prozesse während der Datensammlung sowie die Verarbeitung der Daten sind in diesem Modell von hoher Bedeutung. Hierbei ist zu unterscheiden, dass nur gewisse Daten benötigt werden und woher diese gewonnen werden.

Um die steigende Bedeutung des autonomen Fahrens zu verdeutlichen, wird in der ersten Abbildung eine Statistik über die Entwicklung der angemeldeten Patente im Bereich des autonomen Fahrens weltweit aufgezeigt. Im Jahr 2019 gab es insgesamt 18.260 Patente, was eine Steigerung von 1.344,62% im Vergleich zum Jahr 2009 entspricht. In den letzten Jahren wurde der Fokus in der Automobilindustrie vermehrt auf das autonome Fahren gelegt. Zukünftig soll es möglich sein, dass sich Fahrzeuge im öffentlichen Straßenverkehr auch selbstständig fortbewegen können, ohne dass eine Person aktiv das Fahrzeug betätigen muss. Dafür sind verschiedene lernfähige Systeme notwendig. (Süllwold, 2020, S. 7)

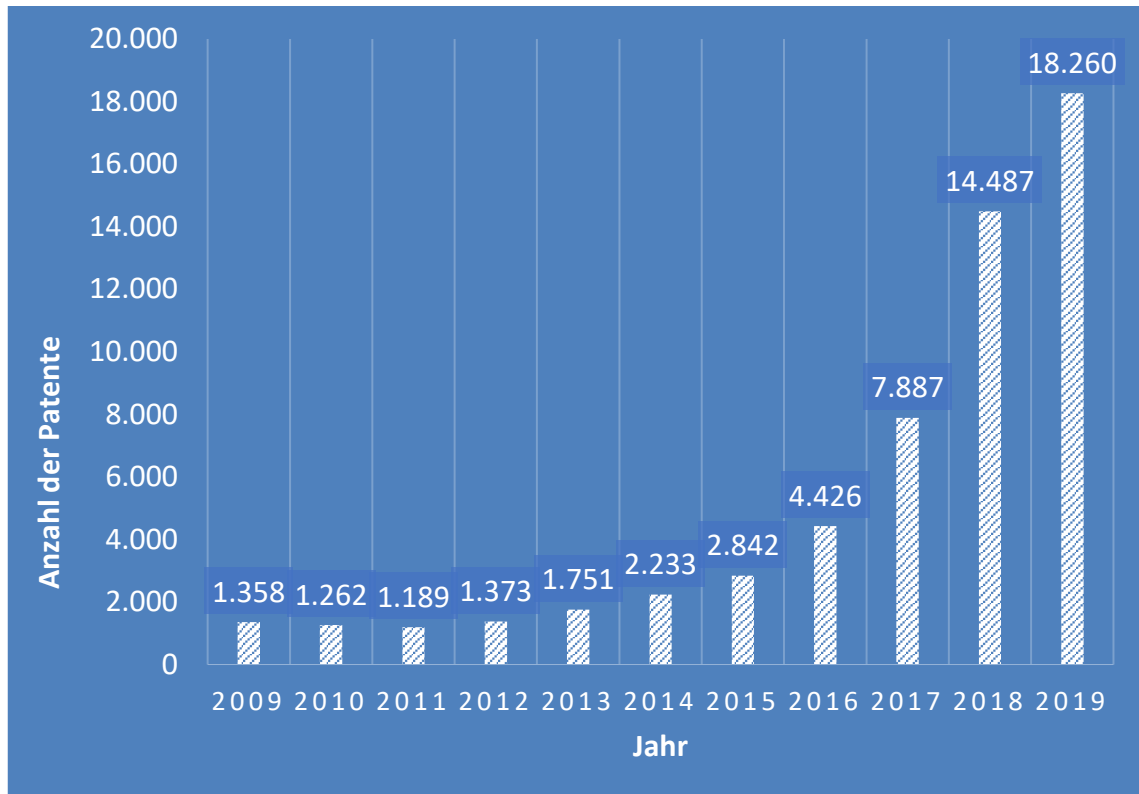


Abbildung 1: Entwicklung der Anzahl der angemeldeten Patente im Bereich autonomes Fahren weltweit (Analog zu: <https://de.statista.com/statistik/daten/studie/1062558/umfrage/anzahl-der-patente-im-bereich-autonomen-fahren-weltweit/>, eigene Darstellung, online abgerufen am 30. April 2022, 16:30 Uhr)

Durch die stetige Steigerung der Anmeldungen der Patente ist deutlich zu erkennen, dass das autonome Fahren in naher Zukunft eine größere Rolle spielen wird als heute. Das klingt in der jetzigen Situation, da noch nicht viele autonome Fahrzeuge im öffentlichen Straßenverkehr unterwegs sind, sehr revolutionär, allerdings ist es in der Luftfahrt oder bei U-Bahnen schon längst Alltag geworden und nicht mehr wegzudenken, denn kein menschlicher Pilot kann einen Autopiloten in den Kategorien Präzision und Fehlerfreiheit übertreffen. (Koch, 2018, S. 4-5)

Wie schnell das autonome Fahren flächendeckend präsent sein wird, hängt von der gesellschaftlichen Akzeptanz ab. Verantwortung, Sicherheit und Vertrauen in die Technik sind die entscheidenden Themenfelder für die Gesellschaft. In der Gesellschaft herrschen allerdings in vielen Bereichen Akzeptanzprobleme für das autonome Fahren im Straßenverkehr. (Audi, 2021, S. 38)

## WELCHE AKZEPTANZPROBLEME SEHEN SIE BEIM AUTONOMEN FAHREN?

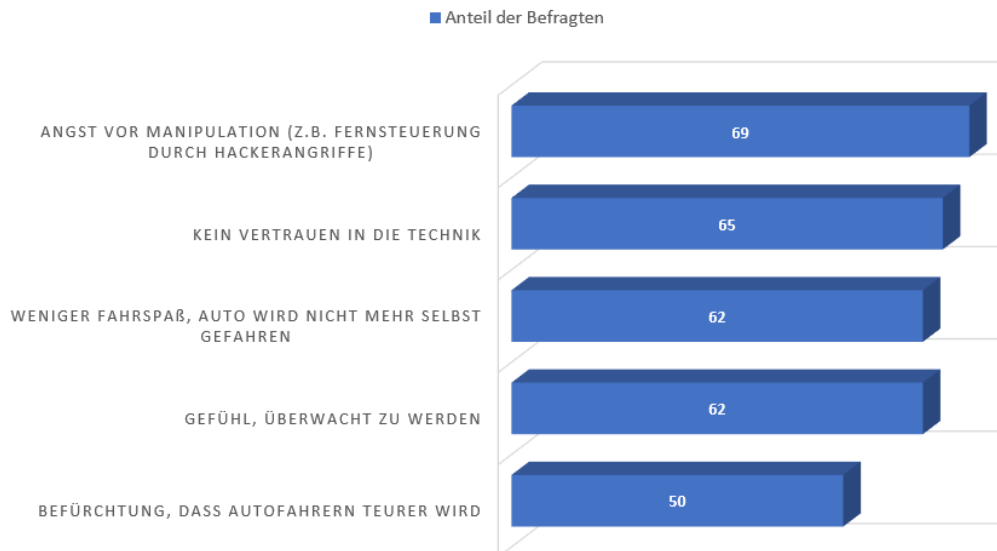


Abbildung 2: Akzeptanzprobleme für das autonome Fahren: März 2019; 500 Befragte Personen; ab 18 Jahren (eigene Darstellung) (Analog zu : <https://de.statista.com/statistik/daten/studie/270612/umfrage/nachteile-von-autonomen-fahrzeugen/>, online abgerufen am 30. April 2022, 13:50 Uhr)

In Abbildung 2 ist zu erkennen, dass von 500 befragten Teilnehmern in Deutschland 69% Angst vor Cyberangriffen oder Fernsteuerung an ihrem Fahrzeug und dadurch mögliche Manipulationen haben. Daraufhin folgen mit 65% das fehlende Vertrauen in die Technologie.

Wie in der ersten Abbildung zu erkennen ist, gewinnt das autonome Fahren an Bedeutung, allerdings sind noch viele Schwierigkeiten insbesondere in der Akzeptanz von der Gesellschaft vorhanden. Die Bachelorarbeit hat das Ziel, das mangelnde Vertrauen in die Technik durch Convolutional Neural Network für das autonome Fahren umzukehren. Ein großes Einsatzgebiet von Convolutional Neural Network ist in der Bild- und Mustererkennung. Allerdings kann die Erkennung von Verkehrsschildern und Ampeln nur dann funktionieren, wenn ein Convolutional Neural Network Modell die Fähigkeit besitzt, Bilder und Muster zu erkennen, um letztlich auf ein Verhaltensmuster zu schließen. Im Bereich des autonomen Fahrens spielt aus diesem Grund das Convolutional Neural Network eine wichtige Rolle. (Dimitri, 2016)

Daher wird in dieser Bachelorarbeit unter anderem der Forschungsstand von Convolutional Neural Network dargestellt, sowie die Funktionsweise und dessen Komponenten, die erst den funktionalen Zusammenhang zwischen dem Convolutional Neural Network und des autonomen Fahrens ermöglichen. Der Schwerpunkt dieser Bachelorarbeit liegt

darin, den gesamten Datenprozess von der Datensammlung bis zur Datenverarbeitung für das Convolutional Neural Network Modell aufzuzeigen. Dabei zeigt die Arbeit die Herausforderungen der Datenerfassung, wie zum Beispiel die Tatsache, dass aufgrund der großen Datenmengen, die gesammelt werden müssen, ein großes Rechenzentrum erforderlich ist. Aus den Problemen werden Lösungsansätze abgeleitet, welche eine Darstellung von Prozessen geben, um das autonome Fahren ohne menschlichen Fahrer zu ermöglichen.

Der Leser soll verstehen, wie die Daten für das Convolutional Neural Network Modell gewonnen, gespeichert und verarbeitet werden und welche Herausforderungen im gesamten Prozess zu bewältigen sind, damit das Convolutional Neural Network Modell sein Potenzial entfalten kann. Durch das Aufzeigen des gesamten Datenprozesses für das Convolutional Neural Network Modell, kann Vertrauen in die Technik gewonnen und der Schwierigkeit bezüglich der Akzeptanz entgegengewirkt werden. Die Ergebnisse dieser Bachelorarbeit eignen sich für die weiterführende Analyse von autonomem Fahren und leisten einen akademischen Beitrag zur Diskussion über den gesamten Datenprozess für das Trainieren des Convolutional Neural Network Modells bei.

Eingeleitet wird die Bachelorarbeit durch eine Einführung in die Thematik des maschinellen Lernens, welches die Grundlage für das autonome Fahren bildet. Im nachfolgenden Kapitel werden künstliche neuronale Netzwerke beschrieben, deren Bestandteile ebenfalls wie das Convolutional Neural Network auf mathematischen Modellen aufgebaut werden. Um die Basis der Abschlussarbeit festzulegen, wird sowohl auf die Geschichte als auch auf die für Convolutional Neural Network benötigten Komponenten Bezug genommen. Das Convolutional Neural Network hat einen hohen Stellenwert in dieser Bachelorarbeit, da gezeigt wird, ob ein Convolutional Neural Network Modell für das autonome Fahren geeignet ist. Zusätzlich werden in diesem Zusammenhang wichtige Begriffe definiert. Am schließlich erfolgt die Auseinandersetzung mit der eigentlichen Thematik bezüglich der Daten für das autonome Fahren. Speziell im Bereich der Datensammlung wird aufgezeigt, welcher große Aufwand von Rechenzentren betrieben werden muss, um die hohe Quantität an gesammelten Daten zu speichern und zu verarbeiten. Weiter wird beschrieben, wie Daten im Kontext zum autonomen Fahren gewonnen werden und welche Daten unumgänglich für ein gutes trainiertes Convolutional Neural Network Modell sind. Abschließend werden die Ergebnisse der Bachelorarbeit zusammenfassend dargestellt, um eine präzise Beantwortung der Forschungsfragen zu ermöglichen, welche im Fazit sowohl ausgewertet als auch interpretiert wird.



## 2 Maschinelles Lernen

Im Folgenden Kapitel werden nun zunächst die theoretischen Grundlagen des maschinellen Lernens betrachtet, die zum Verständnis für das Convolutional Neural Network notwendig sind. Nach der Erläuterung des maschinellen Lernens werden die Künstlichen Neuronalen Netzwerke sowie das Convolutional Neural Network und dessen Komponenten erklärt. Das Kapitel wird mit der Erklärung der Wichtigkeit des Convolutional Neural Networks für das autonome Fahren und dem Zielbild bei der Klassifikation im öffentlichen Straßenverkehr abgeschlossen.

### 2.1 Einführung maschinelles Lernen

Das maschinelle Lernen ist ein wichtiger Bestandteil der Künstlichen Intelligenz und gehört in die Kategorie der Informatik. Durch das maschinelle Lernen werden IT-Systeme, in eine Lage versetzt, wodurch die anhand der vorhandenen Datenbestände und Algorithmen Muster und Gesetzmäßigkeiten erkennen. Dadurch entwickelt das maschinelle Lernen Lösungen. (Sigurd Schacht, 2019, S. 90-91)



Abbildung 3: Programmierung auf klassische Weise (links) und durch maschinelles Lernen (rechts) (Sigurd Schacht, Blockchain und maschinelles Lernen - Wie das maschinelle Lernen und die Distributed-Ledger-Technologie voneinander profitieren, 2019, S. 90)

Für die Intuition wie maschinelles Lernen funktioniert, kann eine Analogie-Bildung angeschaut werden, wie ein Mensch lernt. Um neues Wissen zu erlernen, oder sich neue Fähigkeiten anzueignen, besteht für einen Menschen die Möglichkeit, Daten und Informationen aus bspw. Lexika zu lesen oder Wissen aus alten und neuen Erfahrungen zu generieren, um eigenständige Lösungen für neue und unbekannte Probleme zu finden. Die dafür aufgewendete zielgerichtete Anstrengung zeigt einen Prozess, durch den eine Anreicherung des Wissensschatzes erreicht wird. Durch das wiederholte lesen der Kernaussage in weiteren Lexika oder durch neue Erfahrungen, lassen sich Verhaltensänderungen, welche auf Beobachtung, Übung und zurückzuführen sind, feststellen. Im Vergleich zu Maschinen sind Menschen in der Lage, ihr Verhalten anzupassen sowie Zusammenhänge zu erkennen, Rückschlüsse zu ziehen und Vorhersagen zu treffen. (Aichele & Jörg, 2021, S. 31-32)

Nachdem verdeutlicht wurde, wie das maschinelle Lernen interpretiert werden kann, schauen wir uns ein Praxisbeispiel für das maschinelle Lernen im Bereich autonomes Fahren für das Antrainieren der Objekterkennung an. Für die kamerabasierte Objekterkennung wurde ein Ansatz verfolgt, mit dem der Aufwand zur Generierung von Trainingsbildern deutlich reduziert werden kann. Diese Methodik des maschinellen Lernens kann für die Erkennung von Objekten verwendet werden. Als gutes Beispiel können hier Ampeln und Verkehrsschilder aufgeführt werden. Um die Funktionsweise besser darzustellen, wird Abbildung 4 erläutert. In Abbildung 4 ist zu sehen, wie aus unterschiedlichen Perspektiven und vor einem homogenen Hintergrund Bilder aufgenommen wurden, um dann im späteren Anwendungsfall erkannt zu werden. Der nächste Schritt besteht darin, dass die Objekte (synthetisch erstellte Bilder) aus einer Menge der aufgenommenen Bilder mit entsprechenden Bildverarbeitungsalgorithmen automatisiert ausgeschnitten werden. Die Objekte, die ausgeschnitten wurden, werden abschließend für das Training in ein Hintergrundbild (unterer Bereich der Abbildung) eingefügt. Das Zusammensetzen in ein Hintergrundbild dient als Training für das autonome Fahrzeug. Ein Trainingsbeispiel wäre die Ampelschaltung. Wird die verwendete Ampel im Bild auf Grün geschaltet, soll das autonome Fahrzeug lernen, die Fahrt fortsetzen. Für das maschinelle Lernen mit der Convolutional Neural Network Technologie ist es nur relevant, eine hohe Trefferquote für lokale Merkmale zu erzielen, damit das autonome Auto dementsprechend reagieren kann. (Schiefkofer, Erdogan, Schindler, & Wendl, 2019, S. 50-51)

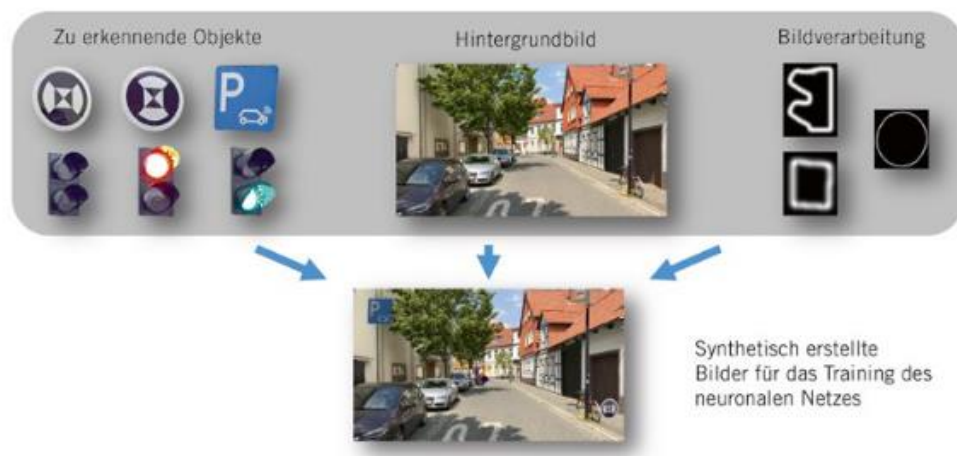


Abbildung 4: Hierbei handelt es sich um eine grafische Darstellung des Ansatzes zur Generierung von künstlichen Trainingsbildern (Schiefkofer, Erdogan, Schindler, & Wendl, Maschinelles Lernen für das automatisierte Fahren, 2019, S. 51)

Im Bereich des maschinellen Lernens gibt es drei Vorgehensweisen, die sich im Lernprozess unterscheiden.

- **Überwachtes Lernen (Supervised Learning):** Im gesamten Lernprozess ist der korrekte Ausgabewert bekannt.
- **Unüberwachtes Lernen (Unsupervised Learning):** Im gesamten Lernprozess ist der Ausgabewert nicht bekannt oder wird nicht genutzt. (Matzka, 2021, S. 11-12)
- **Bestärkendes Lernen (Reinforcement Learning):** Das Lernsystem funktioniert durch die Interaktionen und Erfahrungen direkt von der Umwelt. (Ege & Paschke, 2020, S. 11-12)

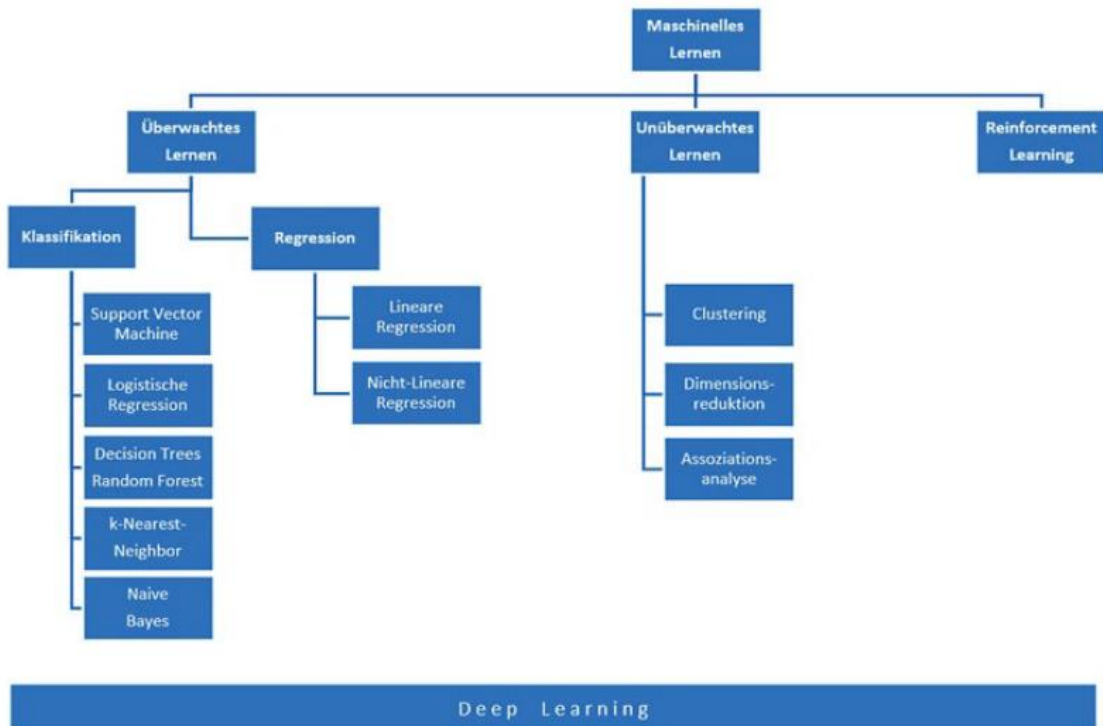


Abbildung 5: Einteilung und Verfahren des maschinellen Lernens (Breyer-Mayländer, Industrie 4.0 bei Hidden Champions, 2022, S.173)

Die erste Variante, die besprochen wird, ist das überwachte Lernen. Das sicherste Lernen stellt das überwachte Lernen dar. Aus diesem Grund ist es am effektivsten, bei sicherheitsrelevanten Lernaufgaben das überwachte Lernen zu bevorzugen. Dadurch können aber keine innovativen Erkenntnisse aus den Daten gewonnen werden, weswegen es nicht für kreative Aufgaben geeignet ist. Das Hauptziel des überwachten Lernens ist die Optimierung. (Mockenhaupt, 2021, S. 139)

Beim überwachten Lernen wird dem Algorithmus des Maschine Learning ein Datensatz vorgelegt, dessen Zielvariable bereits bekannt ist. Dadurch erlernt der Algorithmus bestimmte Abhängigkeiten und Zusammenhänge innerhalb der Daten, mit denen sich die Variablen erklären lassen. Nach der Durchführung wird die Qualität der Vorhersage bewertet. Die erlernten Muster werden dann auf neue, hinzugewonnene und unbekannte Daten angewendet, um Prognosen sowie Vorhersagen zu erstellen. (Wuttke, Was ist Supervised Learning (Überwachtes Lernen)?, 2020)

Beim überwachten Lernen werden zwei Mechanismen eingesetzt:

- **Klassifikation:** Immer bei diskreten Ausgabewerten. Bei einer Klassifikation wird ein Input aus einer Auswahl von vorher festgelegten Eigenschaften (Klassen) zugeordnet. Bei einer Bilderkennung für das autonome Fahren könnten es zum Beispiel Verkehrsschilder sein wie zum Beispiel „Stop“ und „Vorfahrt“, welche vorher in Klassen festgelegt wurden. (DIVISIO, 2019)
- **Regression:** Bei kontinuierlichen Ausgabewerten. Die Regression zeigt Zusammenhänge zwischen zwei oder mehreren Variablen an. (Matzka, 2021, S. 10)

In einer Klassifikation sind in einem Datensatz Eingabeeigenschaften eines beliebigen Typs sowie eine diskrete Eigenschaft, die als Ausgabewert definiert wird, bekannt. Das Ziel hier besteht darin, dass das maschinelle Lernverfahren anhand von Mustern in den Eingabedaten deren konkrete Klasse bestimmen kann. (Matzka, 2021, S. 11)

Nach der Erläuterung der Klassifikation wird die Regression erklärt. Bei der Regression ist der Algorithmus darauf ausgelegt, eine Ausgabe aus einem kontinuierlichen Bereich möglicher Werte zu prophezeien. Dafür muss der Algorithmus eine funktionale Beziehung zwischen den Eingabe- und Ausgabeparametern identifizieren. (TIBCO Software Inc., 2020)

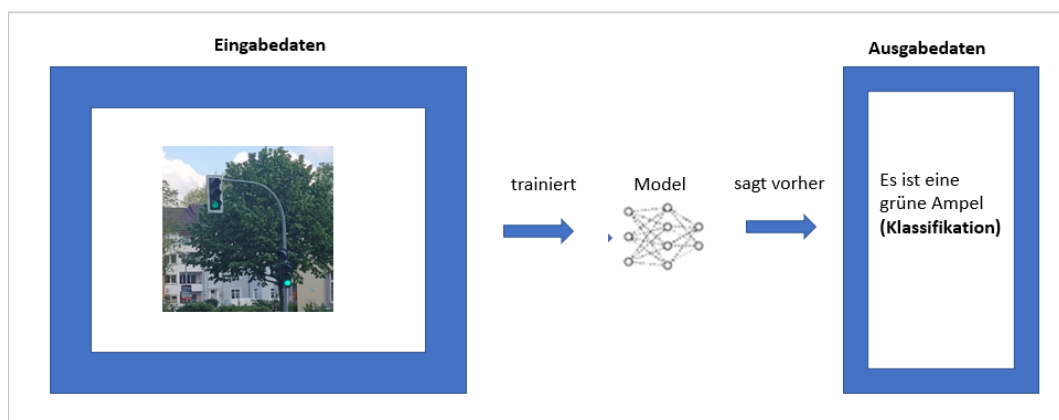


Abbildung 6: Eigene Darstellung: Überwachtes Lernen: Model trainiert anhand von bekannten Daten und Beispielen. Es gibt eine klare Zielvariable, die vorhergesagt wird. (Quelle: Analog zu <https://dataso-lut.com/wiki/supervised-learning/>, online abgerufen am 10.05.202)

Als nächstes wird ein Praxisbeispiel im Bereich des überwachten Lernen aufgezeigt. In Abbildung 6 sind die Eingabedaten vom System zu erkennen. Das überwachte Lernen, worauf das Convolutional Neural Network Modell beim autonomen Fahren beruhen wird, trainiert anhand von bekannten Daten und Beispielen. Das überwachte Lernen unterteilt bei der Ausgabe die Datensätze durch eindeutige Zielvariablen, die vorhergesagt werden. Im Praxisbeispiel ist zu erkennen, dass die Eingabedaten eine grüne Ampel ist. Das Modell wurde darauf trainiert, dass bei einer grünen Ampel die Fahrt fortgesetzt werden kann, wenn die Umgebungsbedingungen stimmen. (Müller, 2021)

Die Gegenseite vom überwachten Lernen ist das unüberwachte Lernen. Das unüberwachte Lernen ist eine Methode, bei der der Algorithmus selbständig und ohne Überwachung Muster und Zusammenhänge in Daten explorativ erkennt. Die Eingangsdaten haben dabei weder eine bestimmte Beschriftung noch eine vorgegebene Zielvariable, weswegen die gewünschten Ausgabedaten nicht vorgegeben sind. (Wuttke, Was ist Unsupervised Learning (Unüberwachtes Lernen), 2020)

Das unüberwachte Lernen hat wie das überwachte Lernen verschiedene Mechanismen:

- **Dimensionsreduktion:** Das Ziel der Dimensionsreduktion besteht darin, die Anzahl der Merkmale zu verringern, ohne bestimmte Charakteristiken – wie zum Beispiel den paarweisen Abstand der Datenpunkte voneinander wesentlich zu verändern.
- **Clusteranalyse:** Die Clusteranalyse und das damit verbundene Verfahren hat das Ziel, eingehende Datensätze in Gruppen ähnlicher Informationsobjekte einzuteilen. Analog zum überwachten Lernen wäre eine Klassifikationsaufgabe – nur werden die Clusterlabels nicht durch einen Trainingsdatensatz vorgegeben, sondern vom Algorithmus bestimmt. Hier wird aus der Struktur der Eingangsdaten erlernt. (Plaue, 2021, S. 255)

Die Clusteranalyse im Bereich autonomes Fahren kann für die Festlegung der Unfallentwicklung auf lokaler Ebene genutzt werden. Das ermöglicht die Erfassung von den notwendigen Ereignisdaten, um gezielte Maßnahmen für lokale Verbesserungen des Verkehrskonzepts abzuleiten. Dadurch ist es möglich, das Verkehrskonzept zu verbessern und damit das Unfallrisiko an den jeweiligen Orten zu verringern. Das führt zu einer Reduzierung des Unfallpotentials. (Breitfuß, 2020, S. 21)

- **Assoziationsanalyse:** Das Verfahren bei der Assoziationsanalyse beschäftigt sich mit der Suche nach interessanten Beziehungen zwischen Variablen in großen Datenmengen. (Töbel & Kropp, 2018, S. 17)

Nachdem die beiden Varianten erläutert wurden, wird die letzte mögliche Vorgehensweise im Bereich des maschinellen Lernen dargestellt, nämlich das bestärkende Lernen (Reinforcement Learning). Im Gegensatz zum überwachten Lernen benötigt das beschränkte Lernen im Vorfeld keine Daten. Die Bildung der Daten erfolgt stattdessen durch ausführliche Trial-and-Error-Abläufe innerhalb eines dafür simulierten Szenarios. Mitten im Trainingsdurchlauf werden beim Reinforcement die erforderlichen Daten markiert und generiert. Im Reinforcement Learning ist eine unabdingbare Komponente vorhanden, nämlich der Agent. Der Agent repräsentiert die Künstliche Intelligenz im Reinforcement Learning. Der Agent ist darauf programmiert, eine Strategie zu erlernen, um die erhaltenen Belohnungen anhand einer Belohnungs-Funktion (Nutzwertfunktion) zu maximieren. Dadurch hat der Agent eigenständig erlernt, welche Aktion in der jeweiligen Situation die beste ist. (Laurenz, 2020)

Die drei möglichen Varianten des maschinellen Lernens basieren auf der Grundlage des Deep Learning, wie in der Abbildung 5 zu erkennen ist. In der Forschung im Bereich der Künstlichen Intelligenz werden immer Systeme genutzt, die vergleichbar wie das neuronale Netzwerk des Menschlichen Gehirns funktioniert. Diese Art von Technologie wird auch als Deep Learning bezeichnet. Deep Learning bedeutet, dass der Lernprozess in künstlichen neuronalen Netzen stattfindet, die mindestens aus mehreren, teilweise aber auch aus sehr vielen inneren Schichten bestehen. (Franken, 2020, S. 39-40)

Das Deep Learning hat in den letzten Jahren einen großen Teil dazu beigetragen, die Interpretation von Sprache, Bildern oder Schrift zu ermöglichen. Es ist ein wesentliches Teilgebiet des maschinellen Lernens. Beim Deep Learning werden parametrisierte Modelle genutzt, welche aus sukzessiv, sogar tief hintereinander aufgebauten Schichten bestehen. Es basiert auf der Grundlage der Künstlichen Neuronalen Netzwerke, die im nächsten Unterkapitel erfolgt. Es sorgt für eine reibungslose Identifikation der passenden Parametern zu einem bestehenden Problem, das gelöst werden muss. Dieser Prozess erfolgt iterativ und wird auch Trainingsprozess genannt. Dabei wird vom „Lernen der Aufgabe“ gesprochen. (Folkers, 2019, S. 3)

## 2.2 Künstliche Neuronale Netzwerke

Künstliche Neuronale Netzwerke sind im Kontext zum Menschen Nervenzellen mittels Synapsen, die vernetzt sind und verschiedene Neuronen mit Verbindungen verknüpfen und ein Netzwerk bilden. Die verknüpften künstlichen Neuronen werden dann zur Berechnungen genutzt. Die Funktionsweise basiert darauf, dass jedes künstliche Neuron ein Signal empfängt und verarbeitet. Das verarbeitete Signal wird vom Neuron ggf. an das nächste künstliche Neuron weitergeleitet. Im Bereich der Bilderkennung kann ein neuronales Netz selbstständig lernen, ob auf einem Bild ein Auto zu sehen ist, sofern es vorher mit Beispielbildern gefüttert wurde, die mit den Labels Auto/kein Auto versehen sind. Das neuronale Netz muss hierfür nicht darauf programmiert werden, dass ein Auto vier Räder hat und auf der Straße fährt. Es lernt selbstständig Attribute, die ein Auto in Bildern ausmacht. Schematisch dargestellt besteht ein künstliches neuronales Netz von künstlichen Neuronen, die in Schichten angeordnet sind. Zwischen der Eingabeschicht (Input Layer) und der Ausgabeschicht (Output Layer) finden sich weitere verdeckte Schichten (Hidden Layers). (Wennker, 2020, S. 18-20)

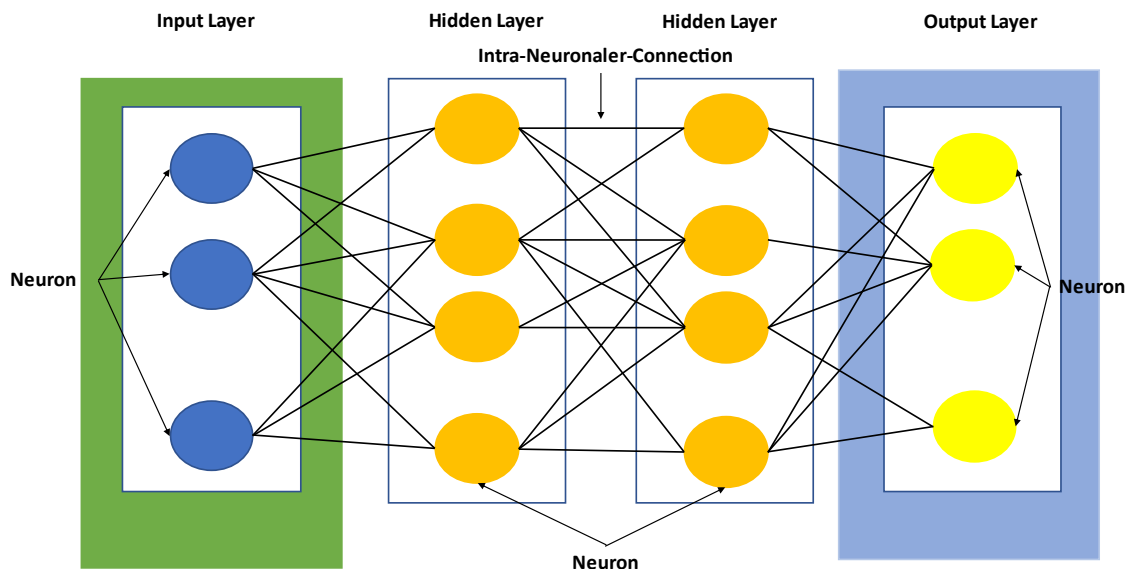


Abbildung 7: Bestandteile eines künstlichen neuronalen Netzes (Eigene Darstellung, analog zu: Wennker, Künstliche Intelligenz in der Praxis - Anwendung in Unternehmen und Branchen: KI wettbewerbs- und zukunftsorientiertes einsetzen 2020)

Nachdem erläutert wurde, wie künstliche neuronale Netzwerke funktionieren, werden die Bestandteile eines KNN erläutert. In Abbildung 7 sind die Bestandteile eines künstlichen neuronalen Netzes zu sehen. Das Input Layer im neuronalen Netz, das sich in der Eingabeschicht befindet, hat die Aufgabe, Informationen aufzunehmen und diese zu gewichten. Dabei ist jedes Neuron individuell. Diese Daten werden als Reiz aufgenommen

und an alle Neuronen im inneren, welche auch Hidden Layer genannt werden, weitergegeben. Diese Daten werden vor jeder erneuten Weitergabe in die nächste Schicht wiederum individuell gewichtet. Die Neuronen im Hidden Layer geben die gewichteten Reize weiter, und diese landen dann im Output Layer. Das Output Layer befindet sich in der Ausgabeschicht und dort wird das Gesamtergebnis abgebildet. (Grünwald, 2020)

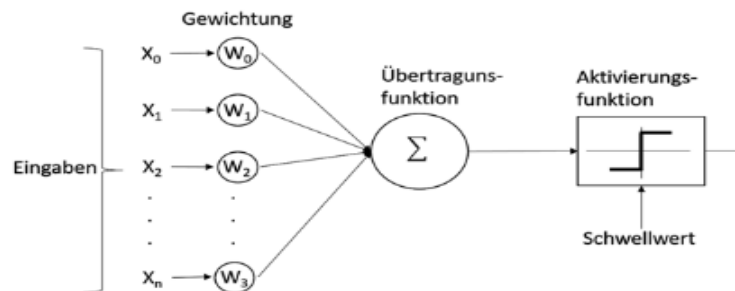


Abbildung 8: Bestandteile eines Neurons (Wennker, Künstliche Intelligenz in der Praxis - Anwendung in Unternehmen und Branchen: KI wettbewerbs- und zukunftsorientiertes einsetzen 2020)

Nach der Erläuterung der Bestandteile wird das mathematische Modell beschrieben. Die künstlichen neuronalen Netzwerke basieren auf mathematischen Modellen. Es wird eine Berechnung mit dem Neuron  $i$  in einem Zeitschritt durchgeführt. Das Aufladen des Aktivierungspotentials, erfolgt durch Summation der gewichteten Ausgabewerte  $x_1 \dots x_n$  aller eingehenden Verbindungen über die Formel:

$$\sum_{j=1}^n w_{ij}x_j$$

Bei den neuronalen Modellen wird diese gewichtete Summe berechnet. Die Aktivierungsfunktion erfolgt nach der gewichteten Summe und des jeweiligen Ergebnisses.

$$x_i = f\left(\sum_{j=1}^n w_{ij}x_j\right)$$

Die Aktivierungsfunktion wird im nächsten Schritt als Ausgabe über die synaptischen Gewichte an die Nachbarneuronen weitergegeben. Es muss beachtet werden, dass für die Aktivierungsfunktion eine Reihe von Möglichkeiten existieren. Um es zu vereinfachen, wird hier die Identität ( $f(x) = x$ ) genommen. (Ertel, Grundkurs Künstliche Intelligenz, 2008, S. 244-245)



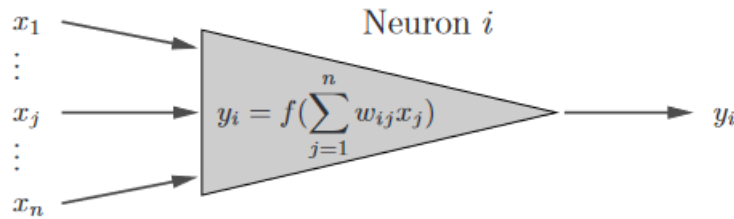


Abbildung 9: Formales Neuron mit den Eingaben der Aktivierungsfunktion (Ertel, Grundkurs Künstliche Intelligenz, 2008)

Neben der Aktivierungsfunktion existiert die Schwellwertfunktion. Die Schwellwertfunktion gibt den Wert 1 zurück, wenn der Bias  $\geq 0$  ist und -1 für alle anderen vorhandenen Werte. (Hölscher, 2021, S. 116)

$$\varphi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{if } v < 0 \end{cases}$$

Die Schwellwertfunktion ist eine von mehreren Aktivierungsfunktionen, weitere Funktionen sind folgende:

- der Tangens Hyperbolicus

$$\varphi(v) = \tanh(v) = \frac{1 - \exp(-v)}{1 + \exp(-v)}$$

- die Sigmoid Funktion mit Steigungsparameter a

$$\varphi(v) = \frac{1}{1 + \exp(-av)}$$

- die Rectified Linear Unit Funktion

$$\varphi(v) = \max(0, v)$$

Der Funktionsverlauf der Aktivierungsfunktionen wird in einem Diagramm in Intervallen dargestellt. (Neiden, 2017, S. 22-23)

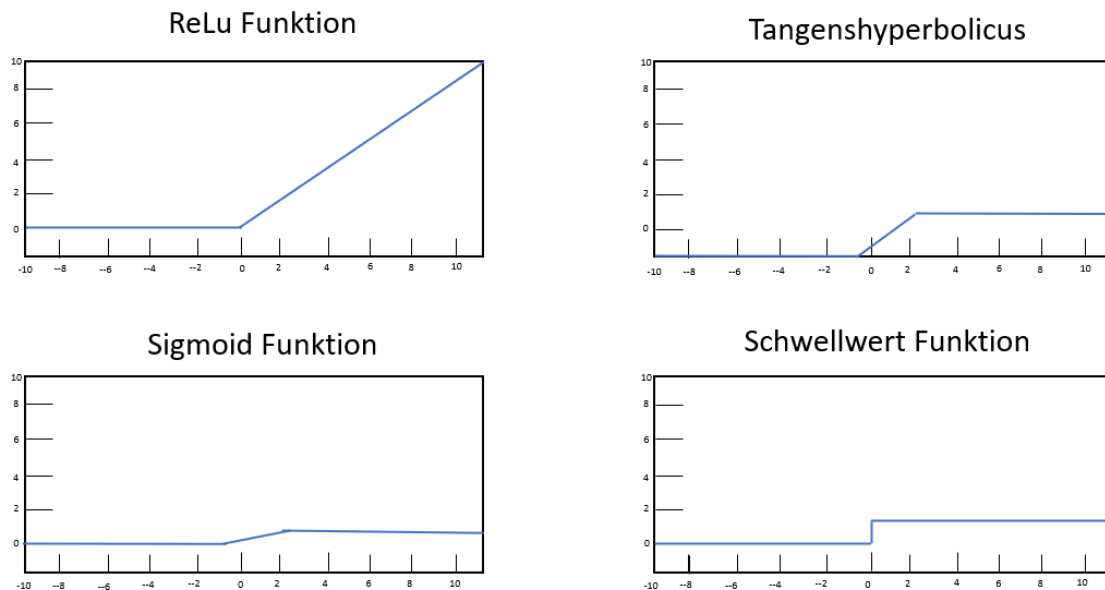


Abbildung 10: Funktionsverlauf der Aktivierungsfunktion im Intervall -10 bis 10 (Eigene Darstellung, analog zu: Neiden, Analyse hochdimensionaler MALDI-Daten mittels Convolutional Neural Networks, 2017)

Die Sigmoidfunktion deckt den Bereich zwischen  $[0,1]$  ab und die ReLu zwischen  $[0,\infty]$ . Aus diesem Grund wird bei Modellierungen von Wahrscheinlichkeiten meistens die Sigmoidfunktion benutzt, da diese Funktion im Gegensatz zu der ReLu Funktion zwischen 0 und 1 beschränkt ist. Mit dem ReLu werden alle positiv reellen Zahlen modelliert. Der wichtige Vorteil dieser Funktion ist, dass bei der Berechnung von Convolutional Neural Networks keine Komplikationen mit dem „Schwinden“ des Gradienten existieren. Außerdem hat sich herausgestellt, dass Convolutional Neural Networks mittels ReLu effizienter trainiert werden können. (Fischer, Convolutional Neural Networks am Beispiel eines selbstfahrenden Roboters 0.1 Dokumentation, 2017)

An dieser Stelle ist die Konsequenz erkennbar, wenn es zu einem Schwinden des Gradienten kommt. Um dies zu verstehen, wird zunächst das Gradientenabstiegsverfahren, oder auch Verfahren des steilsten Abstiegs genannt, erklärt. Das Ziel dieses Verfahren ist die Suche nach dem optimalen Minimum für eine gewisse ausgewählte Funktion und dass die Gewichte bei jeder Iteration durch das Netz in Richtung eines Fehlerminimums angepasst werden. Die allgemein Problematik liegt darin, dass die Trainingsdaten hochdimensional sein können. Je hochdimensionaler die Daten sind, desto höher ist die Gefahr ein lokales, statt ein globales Minimum zu finden. Der Fokus liegt daher auf der Suche nach dem globalen Minimum. (Sauer, Wiesche, & Krimling, 2018, S. 226)

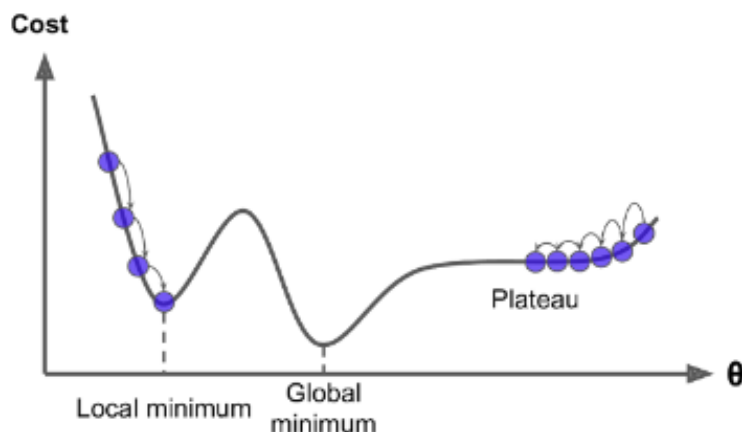


Abbildung 11: Grafische Abbildung vom lokales und globales Fehlerminimum beim Gradientenabstiegsverfahren (Sauer, Wiesche, & Krimmling, Management digitalier Plattformen – Datengetriebene Geschäftsmodelle für Mobilität in der Smart City 2018)

Nun wird ein Vergleich zu der Sigmoid-Aktivierungsfunktion gezogen. Bei dieser Funktion ist der Gradient typischerweise ein Bruchteil zwischen 0 und 1. Bei einem mehrschichtigen Convolutional Neural Network multiplizieren sich diese und erzeugen exponentiell kleine Gradienten. Jeder Schritt des Gradientenabstiegs führt unweigerlich zu einer minimalen Änderung der Gewichte. Diese Änderung hat eine langsame Konvergenz zur Folge. Im Vergleich zu der ReLu-Aktivierung ist der Gradient der ReLu entweder 0 oder 1. Dies führt dazu, dass nach vielen Schichten der Gradient häufig das Produkt aus einer Reihe von 1en enthält und somit der Gesamtgradient nicht zu klein oder zu groß ist. Bei dieser Performance liegt die ReLu Funktion ebenfalls vorne. Diese Funktion ist viel schneller als andere Sigmoid-Aktivierungsfunktionen, die einen Exponenten verwenden, welcher bei häufigem Gebrauch rechenintensiv sein kann. Die Sigmoid-Aktivierungsfunktion war im Bereich des Trainings neuronale Netze früher sehr beliebt, allerdings hat sie das Problem der verschwindenden Gradienten. Daher lautet die allgemeine Empfehlung für das Convolutional Neural Network Training, dass man in den meisten Fällen die Rectified Linear Unit Funktion verwendet werden sollte. (Keras, 2022)

Es ist hervorzuheben, wo die Problematik bezüglich der neuronalen Netze liegt und dementsprechend für das Convolutional Neural Network Modell. Falls es zu einem Verschwinden von Gradienten kommt, ist dies insbesondere ein Problem bei wiederkehrenden neuronalen Netzen, da bei der Aktualisierung des Netzwerkes das Netzwerk für jeden Eingabezeitschritt abgewickelt wird. Dadurch entsteht ein instabiles Netzwerk, welches Gewichtsaktualisierungen erfordert. (Babel, 2021, S. 26)

Der Gradient kann bei neuronalen Netzen mit vielen Schichten nicht nur verschwinden, sondern auch explodieren, wobei der Gradient exponentiell zunehmen kann, wenn er sich durch das Netzwerk rückwärts ausbreitet. Die Explosion des Gradienten, welche als Fehlgradient bezeichnet wird, führt dazu, dass das Modell instabil wird. Dies hat zur Folge, dass es nicht mehr aus den hinzugefügten Trainingsdaten lernen kann. (Brutalk, 2021)

## 2.3 Convolutional Neural Network und seine Komponenten

Die Geschichte von Neuronalen Netzwerken geht weiter zurück als die Künstliche Intelligenz. In den 1940ern Jahren untersuchten Forscher, wie Neuronen im Gehirn arbeiten. Es wurden Experimente durchgeführt, in denen Neuronen in Form von Schaltungen nachgebildet wurden. Den ersten Durchbruch gab es 1957 mit dem „Mark I Perceptron“, einer Maschine, die mit Hilfe eines einzelnen „Neurons“ Eingabedaten in zwei Klassen aufteilen konnte. Die Maschine lernte aus den Fehlern der Vergangenheit, um die folgenden Versuche optimaler zu gestalten. Der Forscher des Deep-Learning Yann LeCun entwickelte in den 90er Jahren die ersten Convolutional Neural Network Modelle, um handgeschriebene Zahlen zu erkennen. Die damalige neue Form von neuronalen Netzwerken war besonders dazu geeignet, bestimmte Sachen auf Bildern zu erkennen. Danach bekamen allerdings bis 2012 neuronale Netze mit Convolutional Neural Networks kaum Beachtung. (Hertwig, 2018)

Ab 2012 kamen neuronale Netze mit Convolutional Neural Networks wieder in den Vordergrund, als der Experte für maschinelles Lernen Alex Krizhevsky sie einsetzte, um den ersten Preis im ImageNet-Wettbewerb zu erhalten. Er konnte den Rekord bei den Klassifizierungsfehlern deutlich senken und erreichte damit 15 Prozent. Hierbei handelt es sich um eine Verbesserung gegenüber dem alten Rekord von 26 Prozent. (Data Science Team, 2020)

Nachdem die Geschichte erläutert wurde, wird nun auf das Convolutional Neural Network und seine Komponenten eingegangen. Das Convolutional Neural Network (zu Deutsch: faltendes neuronales Netzwerk) ist eine besondere Variante von neuronalen Netzwerken. Es dient zu einer automatisierten Klassifikation von digitalen Bildern. Das Convolutional Neural Network dient der vereinfachten Darstellung von digitalen Bildern, in dem es das Quellbild mit jeder Schicht in neue Merkmalskarten umwandelt, auch Pooling genannt. (Matthias, 2021, S. 245)

Die erste Komponenten von Convolutional Neural Network ist das Convolutional Layer. Das Convolutional Layer stellt eine Arbeitsweise dar, in der lokale Merkmale aus den Eingabedaten extrahiert werden. Die Größe des lokalen Merkmals entspricht der Größe

des Filters und wird als rezeptives Feld bezeichnet. Es ist eine Anwendung identischer Operationen auf kleine Bereiche. Es hat die Aufgabe, die Faltungen zu berechnen, indem ein Fenster pixelweise über das Input-Bild geschoben wird. Die Position des Fensters hat viele pixelartige Positionen, und an jeder wird eine gewichtete Summe der Pixel berechnet, die innerhalb der Fenster liegen. Die können durch Neuronen dargestellt werden. Die Aktivierungen werden auch als Feature-Map (Feature = Filter) bezeichnet. Das Convolutional Layer ist ein wesentlicher Bestandteil vom Convolutional Neural Network, da eine Verkettung von mehreren Convolutional Layers im Endeffekt das Convolutional Neural Network darstellt. Diese Methodik findet in der Praxis häufig in der Audio- und Bildanalyse Verwendung. (Proff, Brand, & Schramm, 2020, S. 87)

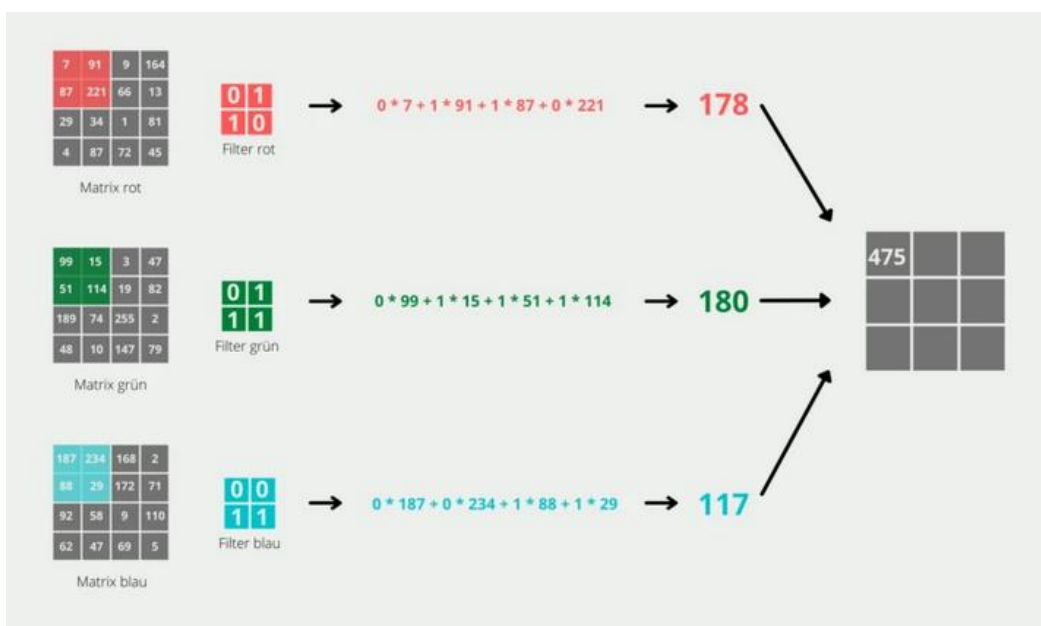


Abbildung 12: Convolutional Layer (Online: <https://databasecamp.de/ki/convolutional-neural-network>; Abgerufen am 25.04.2022; 19:29 Uhr)

Um die Funktionsweise vom Convolutional Layer näher darzustellen, wird hier ein Praxisbeispiel aufgezeigt. In Abbildung 12 sollen die Dimensionen eines Bildes im Convolutional Layer verringert werden. Für diesen Fall werden drei verschiedene Farbfiler mit einer Dimension von 2x2 definiert. In den Farbfilern ist immer eine 1 oder eine 0 vorhanden. Die Zahlen bedeuteten eine Schrittlänge von 1, was bedeutet, dass der Filter nach jedem Berechnungsschritt um einen Pixel weiterverschoben wird. Das Ziel der Schrittlänge 1 besteht darin, dass die Details des Bildes erhalten bleiben. Es hat nicht das Ziel, die Dimensionen stark zu verringern. Bei einer Abwanderung der 4x4 Matrix mit einer 2x2 Matrix, wie in Abbildung 12 zu sehen ist, hat das Convolutional Layer eine 3x3 Matrix als Output. Die Berechnung besteht darin, dass die einzelnen Werte der 4x4 Matrix mit den Farbfilern multipliziert werden. Die Summe ergibt sich aus der Addition

aller drei Ergebnisse. Im nächsten Schritt wird die Komplexität mit dem Pooling Layer erneut reduziert. (Lang, 2021)

Anschließend folgt nach dem Convolutional Layer das Pooling Layer als zweite Komponente vom Convolutional Neural Network. Das Pooling Layer wird benötigt, um zu selektieren, welche Informationen und Daten überflüssig sind und auf welche es sich zu fokussieren gilt. Beispielhaft kann hier wieder die Objekterkennung genannt werden, in der es nicht wichtig ist, an welcher exakten Position ein Feature oder eine Kante ist, sondern die ungefähre Lokalisierung ausreicht. Unterschieden wird hier zwischen zwei unterschiedlichen Varianten: Max Pooling und dem Avg Pooling. (Selle, 2018, S. 36)

Der Unterschied vom Pooling Layer zum Convolutional Layer liegt darin, dass beim Pooling Layer vom Ergebnis je nach Anwendung entweder nur der Maximal- oder der Durchschnittswert übernommen wird. Durch diese Funktionen bleiben Features in wenigen Pixeln erhalten, die für die Aufgabenlösung entscheiden sind. (Lang, 2021)

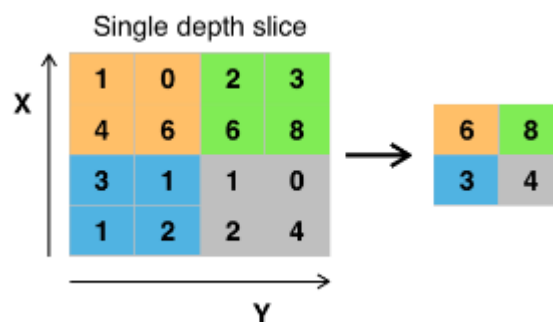


Abbildung 13: Max Pooling mit 2x2-Filter und Stride=2 (Selle, Künstliche Neuronale Netzwerke und Deep Learning, 2018)

In Abbildung 13 ist ein Max Pooling Layer mit 2x2 Filter zu sehen. Es ist ein Beispiel für einen basierten Diskretisierungsprozess. Die Diskretisierung ist ein Teil der angewandten Mathematik und dient der Übertragung kontinuierlicher Funktionen. Das Ziel besteht darin, eine Eingabedarstellung (Bild-, Hidden-Layer-Ausgangsmatrix etc.) abzutasten und die Dimensionalität zu reduzieren. Dadurch werden Annahmen über die in den Unterregionen enthaltenen Features erstellt. Beim Max Pooling wird die Anzahl der zu erlernenden Parameter reduziert, in dem man sich auf das Maximum vom Input  $n$  im jeweiligen Bereich konzentriert. Dadurch wird auch der Rechenaufwand reduziert. (Fischer & Pochwyt, Convolutional Neural Networks am Beispiel eines selbstfahrenden Roboters, 2017)

Bei der Verwendung der Avg Pooling Variante werden ausschließlich Mittelwerte der Aktivitäten verwendet. Mit Hilfe des Poolings, auch Down Sampling oder Subsampling genannt, lässt sich also die Dimension und damit die Komplexität reduzieren. Dadurch

wird eine Überanpassung vorgebeugt, die Berechnungsgeschwindigkeit steigt und der Speicherbedarf sinkt. Pooling ist eine wichtige Voraussetzung und unabdingbar, wenn sehr tiefe Convolutional Neural Network Modelle konstruiert werden sollen, um komplexe Probleme zu lösen. Die nächste Variante, die auf das Pooling Layer folgt, ist das Fully Connected Layer. (Selle, 2018, S. 37-38).

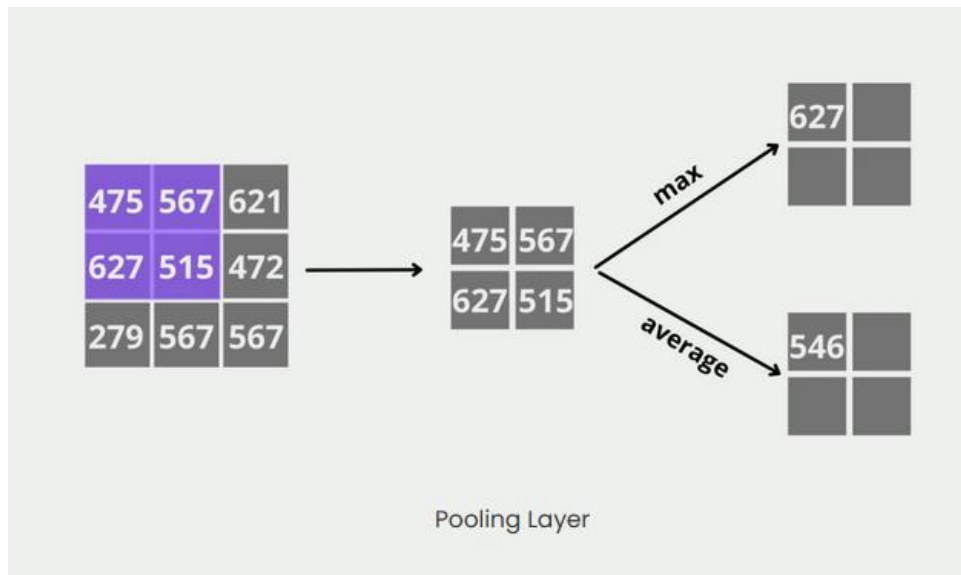


Abbildung 14: Pooling Layer übernimmt die 3x3 Matrix der Convolutional Layer als Input – Dimensionsreduktion (Quelle: <https://databasecamp.de/ki/convolutional-neural-network>; online abgerufen am 26. April, 22:00 Uhr)

Abschließend folgt die letzte Komponente vom Convolutional Neural Network, nämlich das Fully-Connected Layer. Die Neuronen im Fully-Connected Layer (auch Dense Layer genannt) besitzen alle Verbindungen zu den Neuronen des vorrangigen Layers. Sie dient in erster Linie zur Klassifikation. (Brinkmann, 2016, S. 13-14)

$$\begin{bmatrix} w_{11} & \dots & w_{1m} \\ \cdot & & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & & \cdot \\ w_{n1} & \dots & w_{nm} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix} + \begin{bmatrix} b_1 \\ \cdot \\ \cdot \\ b_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ y_n \end{bmatrix}$$

In der obigen Gleichung ist die allgemeine Formel zu sehen, in der der Ausgabewert eines Fully-Connected Layer berechnet wird. Das  $n$  steht für die  $\Sigma$  der Neuronen im Layer und  $m$  ist die Anzahl der Eingabewerte des vorherigen Layers. Der Fully Connected Layer hat Verbindungen zu allen Neuronen von den vorherigen Schichten (Convolutional Layer, Pooling Layer) ohne Beschränkungen. Die Funktionsweise vom Fully-Connected Layer dient dazu, die vorherigen komplexen berechneten Merkmale für die gesamte Eingabe zu berechnen. Aus den komplexen Merkmalen kann dann, falls es zu

Problemen bei der Klassifikation kommt, mit Hilfe von SoftMax Layer Classifiern eine Klassifikation bestimmt werden. (Brinkmann, 2016, S. 14)

Der Fully-Connected Layer wird immer am Ende eines Netzwerkes verwendet, um die gewünschte Anzahl an Outputs zu erreichen. Die gewünschte Anzahl an Outputs wird dann für die Entscheidungsfindung genutzt. (Neiden, 2017, S. 22-23)

Falls es zu Klassifizierungsproblemen vom Fully-Connected Layer kommt, wird das SoftMax Layer verwendet, die dann statt dem Fully-Connected Layer am Ende der Netzwerkarchitektur steht. Der SoftMax Layer berechnet vom vorherigen Layer die Pseudowahrscheinlichkeiten für Klassenzugehörigkeiten. Die Eingabe vom SoftMax Layer hat eine konkrete Dimensionalität mit  $1 \times k$  Werten bei  $K$  Objektklassen. Allerdings haben die Werte die Herausforderung, dass die beliebig klein oder groß sein können. Diese Werte werden mithilfe der SoftMax Funktion

$$\sigma(z) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$$

korrespondierend der Höhe ihres vorherigen Wertes im Bereich  $[0,1]$  gebracht, so dass alle Klassen den Wert 1 haben. Mit dieser Berechnung lässt sich die Ausgabe des neuronalen Netzes, da nach der Normalisierung die Summe aller  $k$  Werte 1 ergibt, intuitiv als Wahrscheinlichkeit angesehen werden. (Brinkmann, 2016, S. 14)

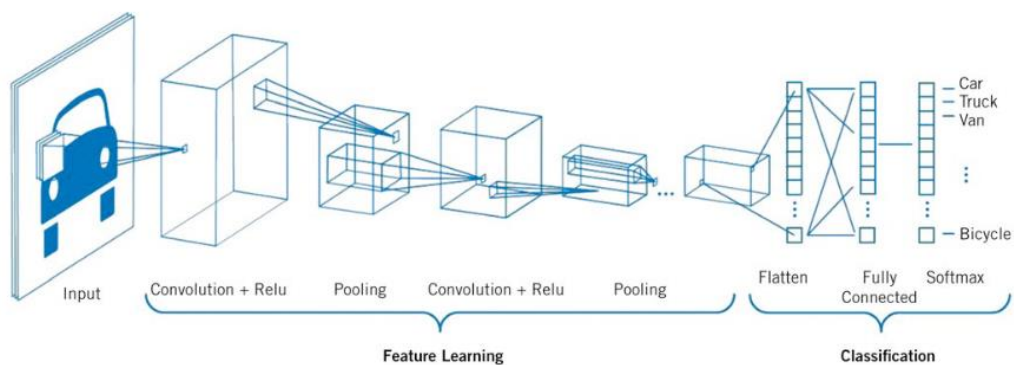


Abbildung 15: Allgemeiner Aufbau eines CNN in Anlehnung an MathWorks (Neiden, Analyse hochdimensionaler MALDI-Daten mittels Convolutional Neural Networks, 2017)

Um die Funktionsweise der Komponenten vom Convolutional Neural Network besser zu verdeutlichen, wird in Abbildung 15 ein Praxisbeispiel im Bereich der Automobile gezeigt. Die gesammelten Daten (Input Layer = Straßenverkehrsschilder und die Umgebung) vom lokalen Merkmal gelangen in das Convolutional Layer. Das Convolutional Layer sorgt dafür, die jeweiligen Eingabedaten einzelner Merkmale (Form, Farbe vom Verkehrsschild etc.) zu erkennen und diese dann zu extrahieren. Der nächste Schritt in Abbildung 15 ist das Pooling Layer. Das Pooling Layer sorgt dafür, wie oben bereits



beschrieben, für eine Selektion der Informationen (Umgebung wie blauer Himmel werden als nicht bedeutsame Informationen selektiert). Es werden überflüssige Informationen nicht betrachtet. Als nächstes erfolgt erneut das Convolutional Layer, indem es alle Daten, die das Pooling Layer für wichtig erhält, extrahiert. Diese Vorgänge werden über Dutzende oder Hunderte von Schichten wiederholt, und in jeder Schicht lernt das Convolutional Neural Network, unterschiedliche Merkmale zu identifizieren. Das wird auch Feature Learning bis zum Fully-Connected Layer genannt. Nachdem das Erlernen im Feature Learning Bereich durchgezogen worden ist, geht die Convolutional Neural Network-Architektur zum Fully-Connected Layer (Klassifikation) über. Das Fully-Connected Layer sorgt dann für die jeweilige Klassifikation. Falls es zu Problemen mit der Klassifikation kommt, wird die SoftMax Funktion verwendet, so dass alle Klassen den Wert 1 haben. In Abbildung 16 ist die vereinfachte Form des Convolutional Neural Networks dargestellt. (MathWorks, 2021)

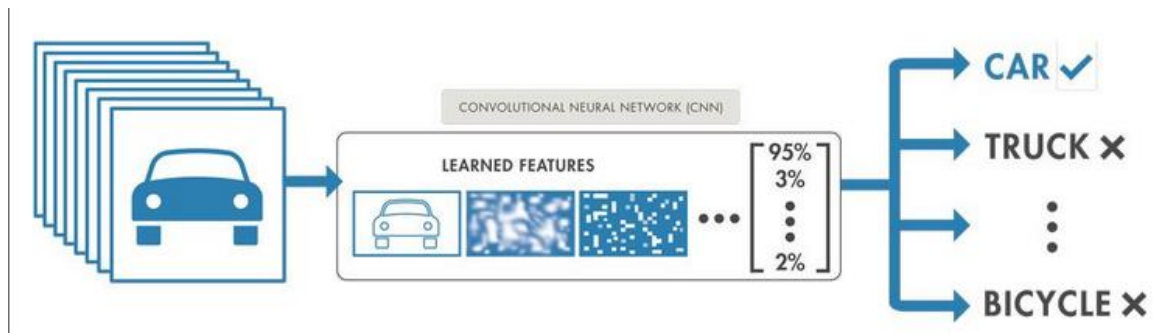


Abbildung 16: Aufbau - vereinfacht dargestellt (Online abgerufen am 28.05.22, 21:10 Uhr : <https://de.mathworks.com/discovery/convolutional-neural-network-matlab.html>; MathWorks, 2021)

## 2.4 Convolutional Neural Network und seine Wichtigkeit im Bereich des autonomen Fahrens

Die Funktionsweise von maschinelles Lernen und Convolutional Neural Network wurden schlüssig aufgezeigt. In diesem Unterkapitel geht es nun um den Zusammenhang mit dem autonomen Fahren, und welche wichtige Rolle das Convolutional Neural Network in diesem Bereich spielt. Im Bereich der Bilderkennung wurden in den vergangenen Jahren durch das Deep Learning enorme technologische Fortschritte erzielt. Durch das Convolutional Neural Network können Straßenszenen in eine semantische Segmentierung übersetzt werden. Diese semantische Segmentierung bildet einen Baustein im Zusammenspiel mehrerer redundanter Systeme, welche das autonome Fahren ermöglichen sollen. (Leyendecker, 2020, S. 100)

Durch die Segmentierung von Convolutional Neural Network werden gewünschte Merkmale bei der Bilderkennung hervorgehoben und die irrelevanten Informationen, die nicht notwendig sind, ignoriert. (Dariti, 2016, S. 5)

Ein weiteres Merkmal, das zeigt, wieso das Convolutional Neural Network besonders wichtig für das autonome Fahren sein kann, ist die Tatsache, dass in der Bilderkennung mit Convolutional Neural Network keine Verzerrungen auftreten. Das bedeutet, dass das Convolutional Neural Network unempfindlich und damit besonders robust auf die Veränderung der Kameralinse oder unterschiedliche Lichtverhältnisse reagiert. (Rowen & Kumar, 2016, S. 3)

Nachdem die Wichtigkeit aufgezeigt wurde, weshalb das Convolutional Neural Network wichtig sein könnte für das autonome Fahren, ist die Frage zu beantworten, was bedeutet überhaupt autonomes Fahren? Das autonome Fahren ist die Fortbewegung von mobilen Robotern, Fahrzeugen und fahrerlosen Transportsystemen, die sich ohne menschliche Hilfe weitgehend autonom bewegen. Dort greifen verschiedene Assistenzfunktionen wie Motor (Beschleunigung), Bremse und Lenkung autonom in das System ein. (Scheffels, 2022)

Das autonome Fahren ist insgesamt in fünf Stufen unterteilt:

- Ohne Assistenz: Der Fahrer besitzt im Auto kein eingreifendes Fahrzeugsystem beziehungsweise keinerlei automatisierte integrierte Fahrfunktionen. Der Fahrer übernimmt die volle Verantwortung. Aufgrund dessen, dass kein Assistenz vorhanden ist, spielt der erste Schritt für die fünf Stufen keine relevante Rolle und hat deshalb keine Stufe (Stufe 0) von den fünf relevanten Stufen für das autonome Fahren.
- Assistiertes Fahren: Der Fahrer übernimmt bei der ersten Stufe die Quer- und Längsführung seines Autos, dabei greift kein Fahrzeugsystem ein. Der Fahrer übernimmt bei der ersten Stufe hiermit die volle Verantwortung.
- Teilautomatisiertes Fahren: Der Fahrer führt bei der zweiten Stufe dauerhaft die Quer- und Längsführung des Autos. Das Auto beziehungsweise die Systeme übernehmen jeweils andere Funktionen. Der Fahrer hat dennoch bei der zweiten Stufe die volle Verantwortung.
- Hochautomatisiertes Fahren: Der Fahrer überwacht bei der dritten Stufe das System dauerhaft. Das Auto übernimmt aber die Quer- und Längsführung für eine gewisse Zeit oder in bestimmten Situationen. Der Fahrer hat bei der dritten Stufe die volle Verantwortung.
- Vollautomatisiertes Fahren: Der Fahrer muss bei der vierten Stufe das System nicht dauerhaft überwachen und ist nicht mehr in spezifischen Anwendungsfällen

erforderlich. Das Auto übernimmt weiterhin die Quer- und Längsführung für eine gewisse Zeit oder in bestimmten Situationen. Das Auto erkennt und respektiert Systemgrenzen, die Zeiten für die Übernahme durch den Fahrer sind relativ lang. Der Fahrer hat bei der vierten Stufe nicht mehr die volle Verantwortung.

- Fahrerlos: Das Auto übernimmt bei der letzten Stufe in allen Verkehrssituationen im öffentlichen Straßenverkehr die komplette Fahraufgabe. Der Fahrer muss nicht überwachen, regeln oder steuern. Es ist bis heute nicht definiert, wer bei einem Verkehrsunfall oder sonstigen Schäden die finale Verantwortung trägt. (Rößger, 2015, S. 51)

Stufe 0 ohne Assistenz	Stufe 1 assistiert	Stufe 2 teilautomatisiert	Stufe 3 hochautomatisiert	Stufe 4 vollautomatisiert	Stufe 5 autonom
Fahrer führt dauerhaft Längs- und Querführung aus.	Fahrer führt dauerhaft Längs- oder Querführung aus.	Fahrer muss das System dauerhaft überwachen.	Fahrer muss das System nicht dauerhaft überwachen.	Kein Fahrer im spezifischen Anwendungsfall erforderlich.	Vollständig Fahrerlos
Keine eingreifendes Fahrzeugsystem	System übernimmt die jeweils anderen Funktionen.	System übernimmt Längs- und Querführung in einem spezifischen Anwendungsfall*.	System übernimmt Längs- und Querführung in einem spezifischen Anwendungsfall*. Systemgrenzen werden erkannt.	System kann im spezifischen Anwendungsfall* alle Situationen automatisch bewältigen.	Das System übernimmt die Fahraufgabe vollumfänglich bei allen Straßentypen, Geschwindigkeitsbereichen und Umfeld Bedingungen.

Abbildung 17: Die fünf Stufen des automatisierten Fahrens (Online abgerufen: <https://blog.dou-bleslash.de/vision-zero-durch-autonomes-fahren/>; am 19.06.2022; 18:40 Uhr)

## 2.5 Zielbild bei der Klassifikation Muster- und Bilderkennung im öffentlichen Straßenverkehr

Wie könnte nun ein Zielbild mit Convolutional Neural Network im Bereich autonomes Fahren aussehen? In den folgenden zwei Abbildungen habe ich selbst Verkehrssituationen in Mülheim fotografiert. Anhand meiner eigenen bearbeiteten Fotos will ich durch die Bearbeitung vereinfacht darstellen, wie das Zielbild von Convolutional Neural Network bei der Klassifikation beim autonomen Fahren an den öffentlichen Straßenverkehren aussehen soll. Bei Abbildung 18 werden die wichtigen Daten klassifiziert in einer Matrixform angezeigt (Auto, Verkehrsschilder, Ampeln) und durch die Segmentierung hervorgehoben. Die Objekte werden lokalisiert. Irrelevante Informationen wie zum Beispiel blauer Himmel oder die Architektur der jeweiligen Gebäude an den Straßenrändern

werden vom Convolutional Neural Network für das autonome Fahren ignoriert. Hier ist es in vereinfachter Form dargestellt. In der zweiten selbst erstellten Abbildung wird das Zielbild komplizierter. Es ist eine weitere Klasse (Bus) dazugekommen, sowie mehrere Autos und Verkehrsschilder sind zu sehen. Ebenfalls befindet man sich nun in einer Kreuzung mit mehreren Autofahrern. Hier wird die Problematik ersichtlich, nämlich das Unmengen von Daten benötigt werden, damit das autonome Fahrzeug in jeder Situation im öffentlichen Straßenverkehr zurechtkommt. (Lynn, 2019, S. 13-14)

Hier wird die Problematik des ganzen ersichtlich, nämlich das Unmengen von Daten von verschiedenen Verkehrssituationen benötigt werden, damit das Convolutional Neural Network Modell trainiert werden kann, damit das autonome Fahrzeug in jeder Situation zurechtkommt. Wie die Quantität von Daten für die vielen möglichen Straßenverkehrssituationen gesammelt, gespeichert und verarbeitet werden, wird im nächsten Kapitel ausführlich erläutert.

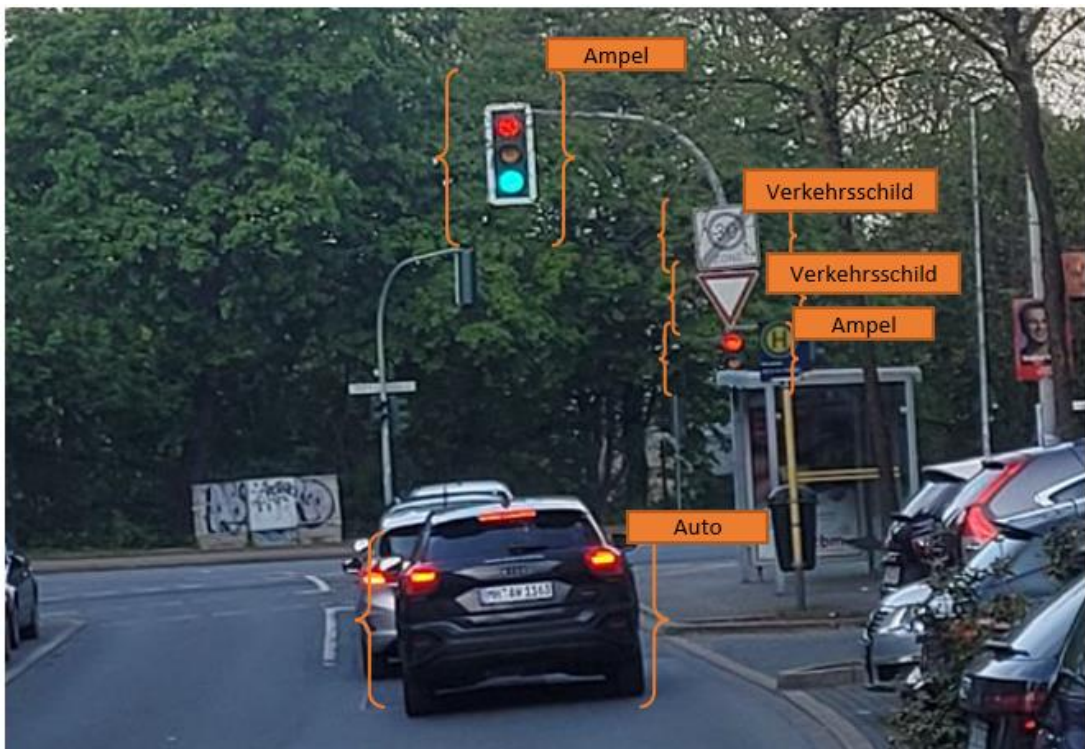


Abbildung 18: Eigenes bearbeitetes Foto – Zielbild CNN im öffentlichen Straßenverkehr Mülheim (1) [Analog zu: <https://artemoppermann.com/de/ki-use-case-computer-vision/>]

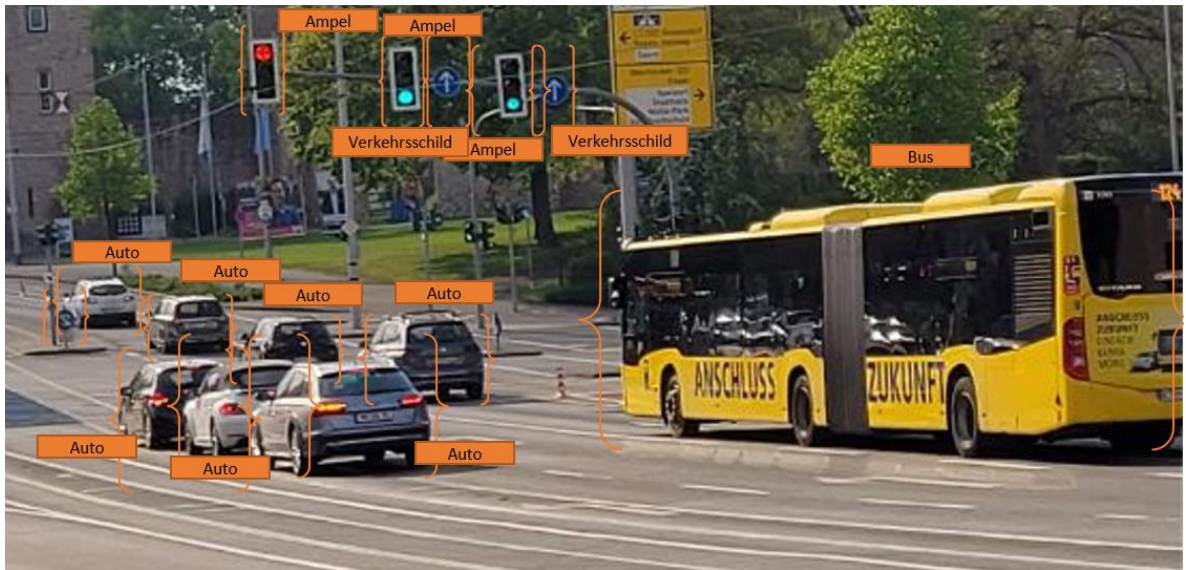


Abbildung 19: Eigenes bearbeitetes Foto – Zielbild CNN im öffentlichen Straßenverkehr Mülheim (2) [Analog zu: <https://artemoppermann.com/de/ki-use-case-computer-vision/>]

## **3 Daten für den Einsatz von Convolutional Neural Network und für das autonome Fahren**

Im vorherigen Abschnitt wurde das maschinelle Lernen und die Funktionsweise von Convolutional Neural Network und dessen Komponenten erläutert. Für das bessere bildliche Verständnis für die Funktionsweise vom Convolutional Neural Network im Bereich autonomes Fahren wurde der vorherige Abschnitt mit einem Zielbild in einem öffentlichen Straßenverkehr dargestellt. In diesem Teil der Ausarbeitung geht es um die Bedeutung von Daten für das Convolutional Neural Network Modell und für das autonome Fahren. Der folgende Abschnitt soll aufzeigen, wie die Daten für das Training von Convolutional Neural Network für das autonome Fahren gewonnen, gespeichert und verarbeitet werden können und welche Herausforderungen dabei auftreten beziehungsweise gemeistert werden müssen.

### **3.1 Bedeutung von Daten und die Datenstruktur für das Convolutional Neural Network**

Der Zusammenhang mit dem vorherigen Kapitel besteht darin, dass die Künstliche Intelligenz und das Machine Learning nur dann funktionieren, wenn auch die entsprechenden Daten vorhanden sind. Die Daten bilden das Fundament für neuronale Netze für das erwünschte Ergebnis. Die Daten können in unterschiedliche Strukturen unterteilt werden. (Herrmann, 2021, S. 32)

Daten können in unstrukturierte und strukturierte Daten unterteilt werden. Strukturierte Daten verfügen als Grundlage über ein Ordnungssystem, in dem Wertigkeiten und Begriffe anhand einer Tabelle systematisch angeordnet werden können. Beispiele für strukturierte Daten sind Öffnungszeiten, Produktpreise, Bewertungen, etc. (Klaus, 2019)

Jeder Datensatz von strukturierten Daten entspricht einer Zeile der Tabelle. Diese Tabellen werden in Unternehmen in Datenbanken gespeichert, die auch unter SQL-Datenbanken bekannt sind. Somit ist das entscheidende Merkmal von strukturierten Daten, dass sie eine Zuweisung zu einem festgelegten Platz einer Tabelle haben. Dadurch lässt sich jede Information leicht wiederfinden. (Herrmann, 2021, S. 32-33)



Merkmal	a1	a2	a3	...	an
Datensatz					
1					
2					
...					
m					

Tabelle 1: Strukturierte Daten (Betriebswirtschaftliche KI-Anwendungen, Herrmann, 2021, S.33)

Unstrukturierte Daten besitzen kein Ordnungssystem und werden dementsprechend auch nicht in Datenbanken abgelegt. Die werden erst dann bearbeitet, wenn sie für bestimmte Zwecke verwendet werden müssen. Unstrukturierte Daten bei der Datengewinnung spielen für das Convolutional Neural Network eine große Rolle. Beispiele für unstrukturierte Daten sind Bilder, Sound-Files, Videos sowie Text-Daten und Artikel. (Klaus, 2019, S. 1)

Daten im Kontext von Convolutional Neural Network eröffnen neue Anwendungspotenziale für Unternehmen im Bereich der Entscheidungsunterstützung und im Management, da unstrukturierte Daten wie zum Beispiel Bilder- und Videodaten attribuiert werden können. So wird einem Unternehmen bspw. die Möglichkeit gegeben, unstrukturierte Daten, wie die von Nutzern in einem Netzdienst hochgeladenen Bilder auf Motive oder die Anzahl der zu sehenden Personen zu bestimmten Zeiten zu analysieren. (Kemper, 2021, S. 229-230)

### 3.2 Daten und die Dateninfrastruktur für das autonome Fahren

Nachdem aufgezeigt wurde, was Daten überhaupt sind und welche wirtschaftliche Bedeutung Daten für Unternehmen, maschinelles Lernen und Convolutional Neural Networks haben, geht es in diesem Kapitel um Daten und die Dateninfrastruktur für das autonome Fahren.

In kaum einem anderen Bereich werden in Zukunft so große Datenmengen gesammelt und verarbeitet wie beim autonomen Fahren. Dazu gehören vor allem Bilddaten aus der Umgebung und Bewegungsdaten aus dem eigenen Fahrzeug und von anderen Verkehrsteilnehmern. (Audi, 2021)

Zuerst muss die Frage geklärt werden, weshalb Daten für das autonome Fahren benötigt werden. Das autonome Fahren kann nur möglich gemacht werden, wenn Wissen generiert wird. Hierfür beruft man sich auf eine große und repräsentative Datenmenge,

die in sechs Kategorien gegliedert werden: (Bundesministerium für Wirtschaft und Energie, 2020, S. 23)

- Kundenwünsche, -erwartungen: Daten, die zu den Kundenwünschen des autonomen Fahrens gehören sind zum Beispiel, dass die Kunden jederzeit eingreifen und die Kontrolle übernehmen können. Weitere Kundenwünsche sind neue Entwicklungen, die Kunden ausprobieren möchten. (Schriller, Rothmann, & Götze, 2016, S. 6-9)
- Technische Anforderungen: Daten über Speicher, Rechenleistung, Sensorik. Dazu zählt auch das Systemdesign, Architektur und die jeweiligen technischen Kosten. (Bundesministerium für Wirtschaft und Energie, 2020, S. 23)
- Performance erzeugen und bewerten: Die Performance wird durch die gesammelten Lerndaten von der Künstlichen Intelligenz erzeugt.
- Anforderungen erfüllen: Es muss eine Verifikation & Validierung von Funktionen vorhanden sein, um sicher zu stellen, ob die Anforderungen, die gestellt werden, garantiert sind. Ebenfalls müssen die Anforderungen durch Szenarien-basiertes Testen kontrolliert werden. (Bundesministerium für Wirtschaft und Energie, 2020, S. 24)
- Sicherheit: Bei der Sicherheit handelt es sich um den Sicherheitsnachweis von automatisierten Fahrfunktionen. Rund 65 Prozent der Befragten befürchten, wie in Abbildung 2 dargestellt, dass die Technik nicht zuverlässig ist. (Schriller, Rothmann, & Götze, 2016, S. 8)
- Sicherheit sicherstellen: Daten über die Qualität und das Sicherheitsmonitoring müssen generiert werden. (Bundesministerium für Wirtschaft und Energie, 2020, S. 25)

Somit haben die Daten, die gesammelt werden müssen, unterschiedliche Anforderungen in den verschiedenen Kategorien zu erfüllen. Für das autonome Fahren auf der technischen Ebene allerdings müssen Trainingsdaten gesammelt werden. Trainingsdaten sind Daten, die den Lernalgorithmus trainieren. Es dient dazu, einen Modell zu erlernen. (Bundesministerium für Wirtschaft und Energie, 2020, S. 25-28)

Dementsprechend ergibt sich die Frage, wo die Trainingsdaten, die notwendig sind für das autonome Fahren gesammelt, gespeichert und verarbeitet werden. Es ist eine Dateninfrastruktur notwendig. Die Daten, die beim autonomen Fahren gespeichert werden, werden an ein Rechenzentrum übermittelt. (SkySystems, 2020)



Ein Rechenzentrum ist ein physischer Ort, an dem gesammelte Informationen und Daten gespeichert, verwaltet und verarbeitet werden. Das Rechenzentrum ist ein wichtiger Bestandteil der Dateninfrastruktur für das autonome Fahren. Die Aufgaben eines Rechenzentrums sind vielfältig und umfassen Tätigkeiten wie die Datensicherung bis zur Ausführung von IT-Prozessen. (Kröger, 2021)



Abbildung 20: Professionelles Rechenzentrum [Online abgerufen: <https://www.it-daily.net/it-management/data-center/wie-sieht-ein-modernes-monitoring-der-rechenzentrums-infrastruktur-aus>, am 10.05.22; 13:30 Uhr]

Das autonome Fahren beginnt nicht mit den technologischen Fortschritten der Bilderkennung, sondern im Rechenzentrum. Damit das autonome Fahren in komplexen Umgebungen zur Realität wird, werden zunächst Unmengen von Daten gesammelt. Ohne diese Unmengen an Daten und die sichere Aufbewahrung dieser sind autonom fahrende Fahrzeuge nicht zu betreiben, da die Systeme anhand der vorhandenen Daten trainiert werden müssen. Aus diesem Grund ist die höchste Priorität für die Realisierung von autonomen Fahrzeugen ein Rechenzentrum mit genug Datenkapazität. (Nguyen, 2020, S. 1-2)

Das autonome Fahren wird ohne Rechentechnik immer in weiter Ferne liegen und nicht erreichbar sein. Um zu verdeutlichen, weshalb ein Rechenzentrum mit einer hohen Datenkapazität benötigt wird, wird das Praxisbeispiel von BMW aufgezeigt. Das Data Driven Development für das autonome Fahren von BMW setzt eine hohe Anzahl von Daten voraus, die in einem bestimmten Zeitrahmen verarbeitet werden müssen. Dabei werden über 100.000 Prozessor-Kerne, mehr als 200 GPUs und über 230 Petabyte Storage eingesetzt. Zwischen den Stationen ermöglichen Glasfaser-Kabel eine Datenübertragung

mit 3,75 Terabyte pro Sekunde. An einem Tag werden dabei über 1.500 Terabyte neue Rohdaten gesammelt. (Schmidt, 2019)



Abbildung 21: Der Weg der Daten von der Straße ins Rechenzentrum und in die Cloud (eigene Darstellung) (Analog zu: Nguyen, Autonomes Fahren beginnt im Rechenzentrum, 2020)

Um einen Vergleichswert zu haben und die Wichtigkeit des Rechenzentrums für die Daten zu verdeutlichen, wird auch ein Praxisbeispiel von Tesla aufgezeigt. Dazu hat Tesla GPU-Cluster NVIDIA A100 Grafikprozessoren verwendet, um leistungsstarke Rechenzentren der Welt zu betreiben. Der GPU-Cluster ist Teil des vertikal integrierten Ansatzes von Tesla für das autonome Fahren, bei dem mehr als 1 Million Fahrzeuge, die bereits auf der Straße fahren, zur Verfeinerung und Entwicklung neuer Funktionen für eine kontinuierliche Verbesserung genutzt werden. Diese Vorhersagen werden aufgezeichnet, und alle Fehler oder Fehleinschätzungen werden dementsprechend protokolliert. Die Ingenieure von Tesla verwenden diese Instanzen dann, um einen Trainingsdatensatz mit unterschiedlichen und schwierigen Szenarien zu erstellen und das Convolutional Neural Network zu verfeinern. Das Ergebnis ist eine Sammlung von etwa 1 Million 10-Sekunden-Clips, die mit 36 Bildern pro Sekunde aufgezeichnet wurden und insgesamt 1,5 Petabyte an Daten umfassen. Das Convolutional Neural Network Modell wird dann im Rechenzentrum immer wieder durch diese Szenarien geführt, bis es fehlerfrei funktioniert. Hierbei handelt es sich um einen iterativen Prozess. Das Training von Convolutional Neural Network auf dieser Anzahl von Daten erfordert eine riesige Leistung an das Rechenzentrum, damit das Training möglich ist. Das hat Tesla dazu veranlasst, die aktuelle Generation von Supercomputern mit leistungsstarken A100-GPUs zu bauen und einzusetzen. (NVIDIA, 2021)

### 3.2.1 Datengewinnung

Bei der Datengewinnung muss die Frage beantwortet werden, wie die Daten für das autonome Fahren zustande kommen. Die Autos verwenden dabei eine Vielzahl von Sensoren, um ihre Umgebung wahrnehmen zu können. Es gibt mehrere Sensoren, die für

die Gewinnung von Daten genutzt werden. Durch die Vor- und Nachteile der verschiedenen Sensoren können mehrere Datenarten (Positionsdaten des Fahrzeuges, Zustandsdaten des Fahrzeugs wie Richtung, Geschwindigkeit, Lenkeinschlag) gewonnen werden. Die Sensoren sind Kamera-, Radar-, Ultraschall- und LiDAR-Sensoren. (Petit, 2020)

- Kamera: Die Kamera hat die Funktion, eine visuelle Repräsentation der Welt in Farbbildern zu präsentieren. Neben den Farben werden auch Textur- und Kontrastinformationen gewonnen. Die gewonnenen Daten können mithilfe von Software-Lösungen ausgewertet werden, um zum Beispiel ein Verkehrsschild oder eine Fahrbahnmarkierung zu identifizieren. Durch die Kameras werden viele Objektdaten erfasst. (Petit, 2020)
- Radar: Die Radar-Sensortechnologie ist robust und kostengünstig. Die Vorteile von Radar-Sensoren sind, dass sie bei trotz schlechtem Wetter zuverlässige Daten liefern. Radar-Sensoren haben allerdings eine geringe Auflösung und besitzen keine Farbinformation. Aus diesem Grund sind sie nicht geeignet für die Identifikation und Unterscheidung von Objekten. (Maurer, Lenz, & Gerdes, 2015, S. 433)
- Ultraschall: Ultraschall-Sensoren dienen für die Erfassung von Abstandsinformationen. Die Vorteile von Ultraschall-Sensoren liegen darin, dass sie bei Nacht und Nebel präzise Daten liefern und Objekte problemlos erfassen können. Die Schwäche liegt aufgrund seines Energiespektrums an der Reichweite, die weniger als 10 Meter beträgt. (Komarnicki, Haubrock, & Styczynski, 2020, S. 229)
- LiDAR: Die LiDAR-Sensoren haben eine hohe Auflösung bei großer Reichweite. Die Punktwolken, die durch LiDAR-Sensoren erfasst werden, sind so detailliert, dass Objekte erkannt und kategorisiert werden können. Ein Beispiel wäre, dass durch die Sensoren ein Fußgänger von einem Fahrradfahrer unterschieden werden kann. Die Sensoren sind unabhängig von Umwelteinflüssen, dadurch können die Sensoren auch bei schlechten Wetterbedingungen zuverlässige Daten liefern. Aktuell sind die Sensoren sehr teuer aufgrund von dem komplexen und wartungsintensiven Aufbau. (Petit, 2020)

Durch die Vielzahl von Sensoren werden Informationen gesammelt, die später ausgewertet und interpretiert werden können. Die Informationen aus den verschiedenen Sensoren müssen gebündelt werden, um ein übersichtliches Bild aller Daten und Datenarten wie Kante der Objekte, Abstand und Umgebung zu erhalten. Die Aufnahmen stellen den Anfang der Verarbeitungskette auf der Softwareebene dar. Nachdem erläutert wurde, wie die Daten mit den Autos gesammelt werden, wird erläutert, wo die Datengewinnung erfolgen kann. (shaip, 2022)

Die Datengewinnung kann in realen technisch aufgerüsteten Teststrecken oder durch virtuelle Simulationen erfolgen. Hier wird ein Praxisbeispiel aufgezeigt, wie Daten in realen Teststrecken gesammelt werden können. Um Eingangsdaten in realen Teststrecken zu erhalten, haben Unternehmen wie ZF Teststrecken gebaut. Die Testfahrzeuge in Friedrichshafen sind für den normalen Straßenverkehr zugelassen und werden von einem erfahrenen Testingenieur gefahren. Der erste Schritt ist die Sammlung von Daten, mitunter auch die Datensätze kritischer Situationen (auch Grenzfälle genannt), wie zum Beispiel Baustellen oder Unfälle. Da bei jeder Testfahrt weitere Situationen dazugewonnen werden, können die Ingenieure mit Hilfe dieser Daten kontinuierlich weitere Simulationen entwickeln. Die Fahrzeuge werden ebenfalls darauf trainiert, mit komplexen Verkehrssituationen umzugehen. Die Testfahrer nutzen in diesem Praxisbeispiel dafür zunächst die Friedrichshafener Basisstrecke. Es befinden sich auf der 5,5 Kilometer langen Strecke unterschiedliche Straßentypen: Eine Bundesstraße sowie Tempo-50- und Tempo-30-Zonen. Auch Kreisverkehre, mehrspurige Abbiegungen oder unmarkierte Straßen werden Computer-Chauffeure aufgrund ihrer Komplexität später vor besondere Aufgaben stellen. Anfang 2019 ist noch eine Route durch die Fußgängerzone ergänzt worden. Um den selbstfahrenden Autos die Orientierung zu erleichtern, werden die ausgewählten Straßen technisch aufgerüstet. In der ersten Aufbaustufe kommunizieren in Friedrichshafen insgesamt neun Ampelanlagen über „Road Side Units“ mit den Testfahrzeugen und übermitteln, ob sie gerade rote oder grüne Leuchtzeichen geben. Außerdem hat der Verkehrscomputer der Leitstelle mehr Rechenpower bekommen, um die Ampeltechnik zu überwachen und die Daten zu verwalten. Der Übergang zu echten autonomen Fahrversuchen soll damit dann fließend werden. Mit den Daten aus der Teststrecke wird das Rechenzentrum gefüttert, der die autonomen Funktionen im Fahrzeug steuert. Das hat das Ziel, dass die einzelnen Funktionen wie zum Beispiel die Reaktionszeit anhand dieser Daten überprüft werden kann. So wird sich dem autonomen Fahren schrittweise angenähert. Dieser Prozess der Datengewinnung nimmt einige Zeit in Anspruch. Die Datengewinnung bei Testfahrten wird später auch in Innenstädten oder Straßen, die in Wohngebieten liegen, durchgeführt. Den Zugang zu diesen Bereichen regeln in Deutschland die Kommunen – und es muss für jede Stadt individuell angefordert werden. Eine entsprechende Kooperation mit der Stadt Friedrichshafen wurde Ende 2018 beschlossen. Gemeinsam mit dem Start-up e.GO bringt ZF den e.GO Mover auf den Markt. Der e.GO Mover ist ein autonomer Shuttlebus. Beim autonomen Fahren ist die höchste Priorität die Datengewinnung, denn je mehr Daten gesammelt und ausgewertet werden können, desto sicherer werden die Fahrzeuge später auf der Straße unterwegs sein. Allerdings ist bei der Datenverarbeitung nicht nur die Anzahl der Daten entscheidend, sondern ebenfalls die Qualität der Daten. Die Teststrecke in Friedrichshafen zeigt:

ZF will weiter eine führende Rolle in der Entwicklung von Technologien zum autonomen Fahren spielen und die Entwicklungskompetenz am Standort bündeln. (Schrahe, 2021)

Um die Datengewinnung zu ermöglichen, wurden an mehreren Standorten weltweit Teststrecken eingerichtet, um ein autonomes Fahrzeug für die Datengewinnung zu trainieren. Hier wird ersichtlich, dass der Markt für autonomes Fahren im Vorwärtstrend ist, da in den letzten Jahren einige technisch aufgerüstete Teststrecken gebaut wurden. Die folgenden fünf Standorte sind:

- Michigan, USA: Diese Teststrecke ist seit 2015 in Betrieb und hierbei handelt es sich um eine künstliche Stadt, in denen Situationen erprobt werden, wo autonom fahrende Autos verschiedene Verkehrsteilnehmer wie Fußgänger und Radfahrern begegnen. Um es noch realistischer zu gestalten, wurden Gebäudefassaden mit bis zu zwei Stockwerken Höhe neben den Straßen gebaut.
- Düsseldorf, Deutschland: In Düsseldorf wurde im September 2018 eine 20 Kilometer lange Teststrecke für automatisierte und vernetzte Autos eingerichtet. Hierbei handelt es sich um ein Projekt Namens Kooperative Mobilität im digitalen Testfeld Düsseldorf (KoMoD).
- Südkorea: Hier wurde ein Testgelände Namens „K-City“ gebaut. Es umfasst 320.000 Quadratmeter mit mehreren Straßentypen sowie Zebrastreifen, Mautstellen, Fußgängerüberwegen und Haltestellen.
- Aachen, Deutschland: Im Oktober 2018 hat die Mikroelektronikforschung die CER Mcity in Aachen eröffnet. Die nachempfundene Stadt bietet auf 50.000 Quadratmetern Parkbereiche, Kreuzungen, Haltestellen, Kreisverkehre und Häuserfronten. Die komplette Infrastruktur ist vernetzt und kommuniziert miteinander. Das Projekt wurde vom Bundesministerium für Bildung und Forschung mit 80 Prozent unterstützt. Das sind insgesamt 3,3 Millionen Euro.
- Peking: In der Hauptstadt von China wurde ebenfalls ein Testzentrum zur Datengewinnung für das autonome Fahren eröffnet. Die Eröffnung fand 2018 statt. Auf dieser Teststrecke testet der Carsharing-Anbieter Pand Autos, die mit der ZF-Baidu-Technologie ausgerüstet sind. (ZF, 2021)

Die Teststrecken benötigen für die Datengewinnung ebenfalls eine Technische Ausstattung. Um ein Beispiel zu nennen, wie eine Technische Ausstattung aussieht, wird hier das Testgelände der Universität der Bundeswehr München, die eine eigene Teststrecke betreibt, genommen. Die Teststrecke besitzt eine RTK-GNSS-Korrektursignal von lokaler Referenzstation. Das RTK sorgt für eine präzise Bestimmung von Positionskordinaten. Das GNSS ist ein Globales Navigationssatellitensystem und dient der Positionsbe-

stimmung und Navigation. Die Technik des RTK Signals ermöglicht es Nutzern, das zielgenaue Fahren von Spur zu Spur zu ermöglichen. Eine weitere Technische Ausstattung sind mehrere RTK-GNSS Empfängern. Der RTK-GNSS Empfänger ist eine Technik zur Erhöhung der Genauigkeit von GNSS-Positionen. Die sorgt dafür, dass sich die Korrekturdaten an einen bewegenden Empfänger sendet. Das wäre in diesem Beispiel das autonome Auto. Die komplette Teststrecke ist ebenfalls mit einer WLAN-Abdeckung abgesichert. (Universität der Bundeswehr München, 2020)

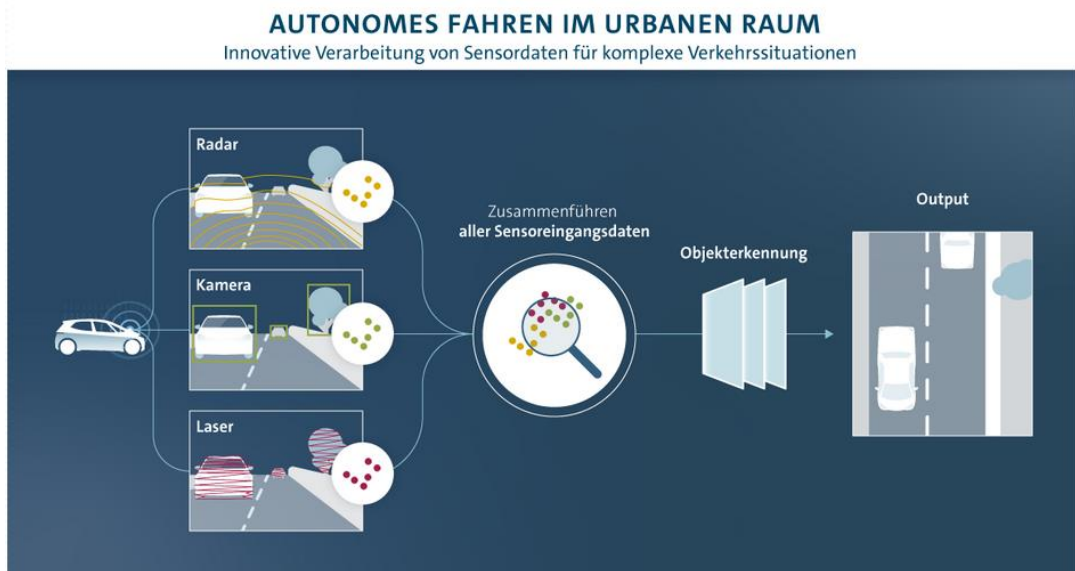


Abbildung 22: Autonomes Fahren im Urbanen Raum - Innovative Verarbeitung von Sensordaten für komplexe Verkehrssituationen (Volkswagen, <https://www.volkswagenag.com/de/news/stories/2019/12/the-champions-league-of-autonomous-driving.html#>, online abgerufen am 08. Mai

Eine weitere Möglichkeit der Datenextraktion bietet die virtuelle Simulation, wie im Spiel GTA IV. Convolutional Neural Network Modelle haben sich bei der Bildklassifizierung als äußerst leistungsfähig erwiesen. Der Nachteil, den Convolutional Neural Networks, wie auch andere überwachte maschinelle Lernmodelle, oft unpraktisch macht, ist der Mangel an Daten. Ein Mangel an Daten liegt dann vor, wenn die gesammelte Datenmenge nicht ausreicht für eine qualitative Bewertung. Es ist nicht nur der Mangel an Daten, sondern auch die Qualität der Daten: Um effektiv zu sein, müssen Modelle des maschinellen Lernens auf einer großen Menge an kommentierten Daten trainiert werden, die genügend Informationen für die Modelle lernen können. Das größte Hindernis bei der Beschaffung eines brauchbaren Datensatzes ist der hohe Kostenaufwand. Bei der Verwendung von realen Daten für diese Aufgabe gibt es mehrere große Hindernisse. Das größte Hindernis ist die Anzahl der benötigten Datenpunkte. (Anthony Martinez, et al., 2017, S. 5-6)

Für die virtuelle Datengewinnung wurden ebenfalls, wie bei den realen Teststrecken, Testfelder errichtet, nur in digitaler Form. Das BMVi zählt 14 digitale Testfelder für das automatisierte und vernetzte Fahren in Deutschland. Die digitalen Testfelder haben alle ihre eigene Straßenkategorie wie zum Beispiel Autobahn oder Hauptverkehrsstraßen, um die Vielfalt an der Datengewinnung voranzutreiben. Jedes virtuelle Testfeld hat verschiedene Fokussierungen in den Bereichen:

- Erprobungsschwerpunkte: Beispiele für Erprobungsschwerpunkte sind Automatisierte Fahrfunktionen und Intelligenter Straßenausstattung.
- Ausstattungsmerkmale: Beispiele für Ausstattungsmerkmale sind Lückenlose Abdeckung mit LTE-Highspeed-Mobilfunk und Schnelle LAN-Verbindung.
- Charakteristika: Beispiele für Charakteristika sind Autobahnen und Städtischer Raum.
- Vorhaben (Status): Beispiele für Vorhaben sind Tank- und Rastanlage und Intelligente Glättevorhersage.
- Ansprechstelle: Beispiele für Ansprechstellen sind Kommunikations- und Koordinierungsplattform und Senatsverwaltung für Umwelt, Verkehr und Klimaschutz. (BMVi, 2020)

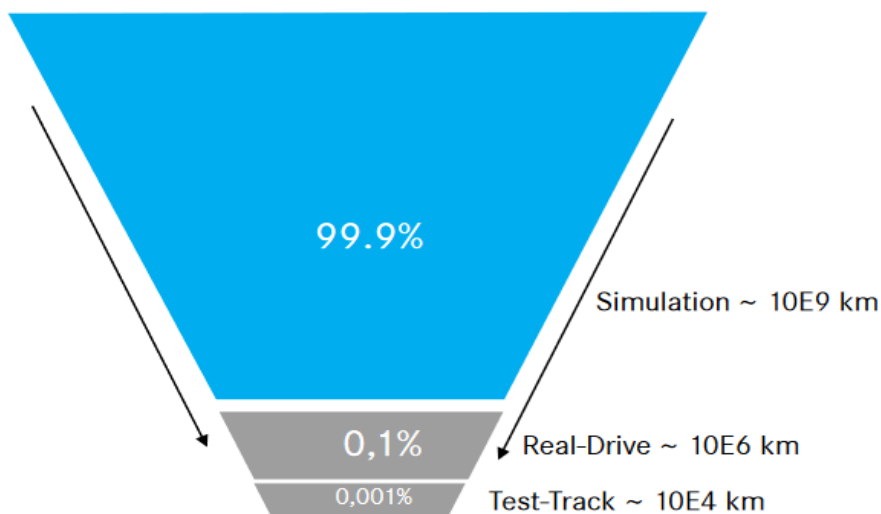


Abbildung 23: Nur wenige Teile der Daten können real sein [Online abgerufen: [https://www.bmwk.de/Redaktion/DE/Downloads/W/ws5-praesentation-ki-und-daten.pdf?\\_\\_blob=publicationFile&v=4;](https://www.bmwk.de/Redaktion/DE/Downloads/W/ws5-praesentation-ki-und-daten.pdf?__blob=publicationFile&v=4;) am 17.05.2022, 20:55 Uhr]

In der Abbildung 23 ist zu erkennen, wie die Datengewinnung aufgeteilt ist. Nur wenige Teile der Daten können von realen Teststrecken sein, da Unmengen an Möglichkeiten, die im öffentlichen Straßenverkehr auftreten, nicht durch eine Teststrecke oder Real-Drive nachgebildet werden können. Aus diesem Grund werden bei einer Simulation ca. 10.000.000.000 km gefahren, um 99,9% der Daten zu bekommen, die für das autonome



Fahren notwendig sind. Für die Annotation der einzelnen 10E9 Key-Frames werden ca. 20 Minuten bis 1 Stunde und 30 Minuten benötigt. Folgend werden zu 0,1% Daten von Real-Drive gewonnen, das entspricht 10.000.000 km. Die restlichen 0,001% (100.000 km) werden aus den Teststrecken gewonnen. (Bundesministerium für Wirtschaft und Energie, 2020, S. 8)

Bei der Datengewinnung ist noch unklar, wie viel Testkilometer nötig sein werden, um automatisierte Systeme zuverlässig zu validieren. Die Einschätzung variiert hier zwischen 240 Millionen und mehreren Milliarden km. Ein Mitarbeiter von Elektrobit zeigt die Dauer auf, die dieser Vorgang in Anspruch nehmen wird. Elektrobit ist eine deutsche Tochterfirma der Continental AG, das im Bereich der Automotive Software tätig ist. Bei Realfahrten würden 100 Fahrzeuge für ca. 240 Millionen Kilometer rund 25 Jahre benötigen. Bei der Simulation, wie von Bundesministerium für Wirtschaft aufgezeigt, was 99,9% der Daten entspricht, würde es bei 100 Computern fünf Jahre dauern, allerdings wären die 240 Million Kilometer in der Cloud in nur zwei Wochen simuliert. Das Problem in der Cloud besteht allerdings darin, dass es sehr kostenintensiv ist. Hier zeigen sich die großen Herausforderungen beim Aufwand der Datengewinnung. (Meyer, Dieckmann, & Laßmann, 2019, S. 80-81)

### **3.2.2 Datenspeicherung und die Herausforderungen**

Nach der Datengewinnung durch reale oder virtuelle Testfahrten müssen die gesammelten Daten gespeichert werden, um im nächsten Schritt verarbeiten werden zu können. Die anfallenden Gigabyte an Daten müssen sicher, problemlos und schnell in das Rechenzentrum transferiert werden können. Für diese Zwecke gibt es zwei mögliche Maßnahmen, um den Datentransfer in das Rechenzentrum zu ermöglichen. Zur ersten Maßnahme zählen die wechselbaren Speichermedien. Wechselbare Speichermedien sind dann eine geeignete Maßnahme, wenn keine schnelle Verbindung verfügbar ist. In diesem Fall können die wechselbaren Speichermedien wie Cartridges aus dem Fahrzeug entnommen, versendet und dem Verantwortlichen im Rechenzentrum übergeben werden. Falls eine ausreichende Verbindung bei der Datengewinnung gegeben ist - das wäre im Idealfall 128 Gbit/s – können die Messdaten direkt aus dem Fahrzeug in das Rechenzentrum übertragen werden. (Friedel, 2020)



Bei der Datenspeicherung für das autonome Fahren im Rechenzentrum ergeben sich folgenden Herausforderungen, die gelöst werden müssen:

- Die Effizienz der Datenspeicherung und -verwaltung: Der Stromverbrauch führt bei größeren Datenmengen zu einer Kostenerhöhung. Die Hauptursache sind die Stromkosten für Speicher, Server und Netzwerke, sowie die wachsenden Rechenleistungen und der zunehmende Bedarf an Dantespeicherung. (Industrie- und Handelskammer Nürnberg für Mittelfranken, 2018)

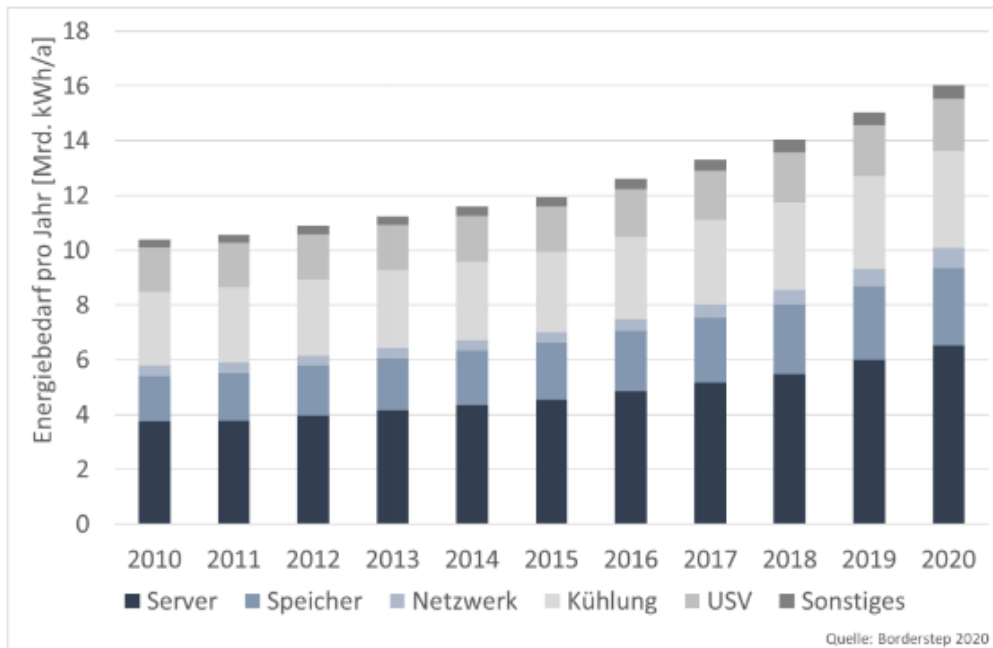


Abbildung 24: Energiebedarf pro Jahr [Mrd. kWh/a] (online abgerufen: <https://www.bundestag.de/resource/blob/863850/423c11968fcb5c9995e9ef9090edf9e6/WD-8-070-21-pdf-data.pdf>; am 20.06.2022; 23:10 Uhr)

Um die Wichtigkeit der Effizienz der Datenspeicherung und den Anstieg der Stromkosten zu verdeutlichen, wird Deutschland als Beispiel genommen. Der jährliche Energiebedarf an Rechenzentren beziehungsweise an den Rechenzentrumstrukturen steigt stetig an, wie in Abbildung 24 zu erkennen ist. Die Ursache liegt am steigenden Strombedarf. Die IT Komponenten (Storage, Server, Netzwerk) benötigten im Vergleich mit 10 Mrd. kWh im Jahr 2020 etwa 75% mehr Energie als im Jahr 2010 (5,8 Mrd. kWh). Die Leistung der Rechenzentren gemessen an der Verarbeitungsmenge bzw. die Auslastung eines Prozesses wurde verzehnfacht. (Deutscher Bundestag, 2021, S. 5-7)

- Die Kapazität muss gewährleistet sein: Das heutige Datenwachstum im informellen Zeitalter hat als Konsequenz, dass eine große Datenspeicherinfrastruktur gewährleistet sein muss, um alle wichtigen gesammelten Daten sicher aufbewahren zu können. (Quantum, 2016, S. 43)

- Die Skalierbarkeit der Leistungsfähigkeit: Der Datenspeicher muss sich leicht an eine erhöhte Arbeitsbelastung angepasst werden können, wenn am Ende festgestellt wird, dass bei der Datensammlung die Größe oder das Volumen die Erwartungen übersteigt. (Hoogenraad, 2021)
- Es werden reichlich an Daten gespeichert: Die Datenspeicherung bei einem vernetzten Auto beträgt durchschnittlich jährlich 280 Petabyte. Aus diesem Grund muss es langfristig eine kostengünstige Speicherung unter Berücksichtigung der gesetzlichen Auflagen gewährleistet sein. (Ziegler, Krämer, & Haustein, 2020, S. 43)

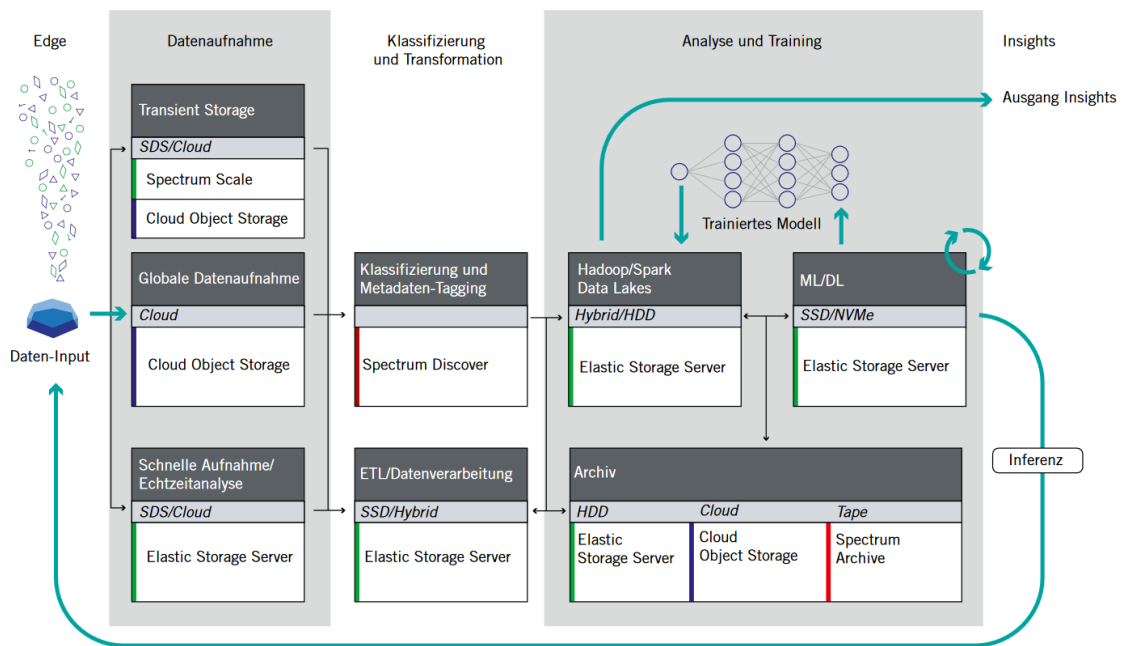


Abbildung 25: Speicheroptionen für die KI-Datenpipeline (Ziegler, Krämer, & Haustein, KI und Big-Data-Management für autonomes Fahren, 2020, S.43)

In Abbildung 25 ist ein Lösungsansatz für die Herausforderungen der Datenspeicherung zu sehen. Die Daten, die bei den Simulationen und Testfahrten gewonnen werden, gelangen in das Rechenzentrum bzw. in die Cloud. Hierbei handelt es sich um unstrukturierte Daten. Aus diesem Grund wird das Cloud Object Storage verwendet. (Ziegler, Krämer, & Haustein, 2020, S. 43)

Das Cloud Object Storage ist ein Format für die Speicherung unstrukturierter Daten in der Cloud. Objektspeicher gelten als gut geeignet für die Cloud, da sie elastisch und flexibel sind und sich leichter auf mehrere Petabytes skalieren lassen, um ein unbegrenztes Datenwachstum von unstrukturierten Daten zu unterstützen. (Lelii, 2020)

Im nächsten Schritt der mittleren Ebene werden nach der globalen Datenaufnahme die Daten klassifiziert. Entsprechend der Klassifizierung kommt das Elastic Storage Server System zum Einsatz. (Ziegler, Krämer, & Haustein, 2020, S. 43)

Das Elastic Storage Server System ist eine mögliche Software-Lösung für die Datenspeicherung und Reduzierung der langfristigen Kosten. Es ist ein softwaredefinierter Speicher. Es dient der Vereinfachung des Daten- und Infrastrukturmanagements mit Automatisierung und vereinheitlichtem Cloud-Betrieb. (International Business Machines Corporation, 2022)

Auf der unteren Ebene der Abbildung 25 im vorletzten Schritt werden die Daten verarbeitet. Im letzten Schritt „Analyse und Training“ ist zu sehen, dass die verarbeiteten Daten gespeichert werden. Die Datenspeicherung ist in mehreren technischen Varianten möglich. Die Daten können in HDD, Clouds oder in Tapes archiviert werden. (Ziegler, Krämer, & Haustein, 2020, S. 43-44)

	HDD	Tape	Cloud
Vorteile	<ul style="list-style-type: none"> <li>• Hohe Geschwindigkeit</li> <li>• Kostengünstig Im Vergleich zu der Cloud</li> </ul>	<ul style="list-style-type: none"> <li>• Sehr hohe Speicherkapazität (sechs Terabyte realisierbar)</li> <li>• Kosten sind günstig im Vergleich zu der Cloud</li> </ul>	<ul style="list-style-type: none"> <li>• Schnelle Netzwerkverbindung</li> <li>• Gute Geschwindigkeit</li> <li>• Keine Obergrenze beim Speicherplatz</li> </ul>
Nachteile	<ul style="list-style-type: none"> <li>• Der Durchsatz bei einer Wiederherstellung duplizierter Daten ist beschränkt</li> <li>• Kurze Lebensdauer</li> </ul>	<ul style="list-style-type: none"> <li>• Die gewonnen Daten sind nicht sofort verfügbar</li> <li>• Die Bänder müssen bewegt und an die richtige Stelle gepult werden</li> </ul>	<ul style="list-style-type: none"> <li>• Wiederherstellung kann Zeit in Anspruch nehmen bei einer schlechten Datenverbindung</li> <li>• Daten können an Unbekannte übergeben werden</li> <li>• Intransparente Datenspeicherung weltweit</li> <li>• Kostenintensiv</li> </ul>

Tabelle 2: Vor- und Nachteile der verschiedenen Datenspeichermöglichkeiten (eigene Darstellung) [online abgerufen: <https://www.cio.de/a/aktuelle-datensicherungsmethoden-im-ueberblick,3564528>; Püschel, 2017)

Auf der mittleben Ebene im letzten Schnitt ist dann das Training aus den verarbeiteten Daten für das Modell zu erkennen. Durch das Elastic Storage Server System (ESS), maschinelles Lernen (ML) und Deep learning (DL) wird am Ende das Zielbild für das autonome Auto dargestellt. (Ziegler, Krämer, & Haustein, 2020, S. 43-44)

### 3.2.3 Datenverarbeitung und die Herausforderungen

Bei mehreren Testfahrten werden viele Daten gesammelt. Die nützlichen und repräsentativen Daten zu extrahieren, erfordert einen hohen Aufwand. Die Daten, die anhand von

mehreren Sensoren (Kamera, Radar, Ultraschall, LiDAR) gesammelt werden, werden begutachtet und mit vorhandenen Datensätzen abgeglichen. Die Problematik besteht darin, dass auch nicht relevante oder auch Daten mit Gemeinsamkeiten gesammelt werden, die am Ende nicht zu gebrauchen sind. Aus diesem Grund müssen die Daten verarbeitet werden. Basierend auf verschiedene Nutzungssituationen, können die Daten entsprechend gefiltert werden: (Nguyen, 2020)

### 1. Kennzeichnung und Verarbeitung der Daten

Die Trainingsdaten werden von den Ingenieuren analysiert und ein definierter Datensatz von Metainformationen wird erstellt und verarbeitet. Beispiele sind Kennzeichnungen wie die Fahrumgebung, Kreuzungssituationen, Wetterbedingungen oder die Anwesenheit anderer Verkehrsteilnehmer.

### 2. Grenzfälle erkennen und verarbeiten

Bei großen Datenmengen ist es zeit- und arbeitsaufwendig, die vielen Sonderfälle zu erkennen, wie zum Beispiel, dass Kinder mit Kreide auf der Straße malen, die während einer Fahrt auftreten können. Es sind jedoch diese Grenzfälle, wie die Kinder oder Rollstuhlfahrer, die für das gezielte Training der Fahrerassistenz-Systeme unter Umständen benötigt werden, um alle notwendigen Daten für das autonome Fahren zu haben. Die Verarbeitung der Grenzfälle erfolgt durch das Sammeln, Aufzeichnen, Ordnen, Speichern und Betrachten dieser Daten. (Nguyen, 2020, S. 36-37)

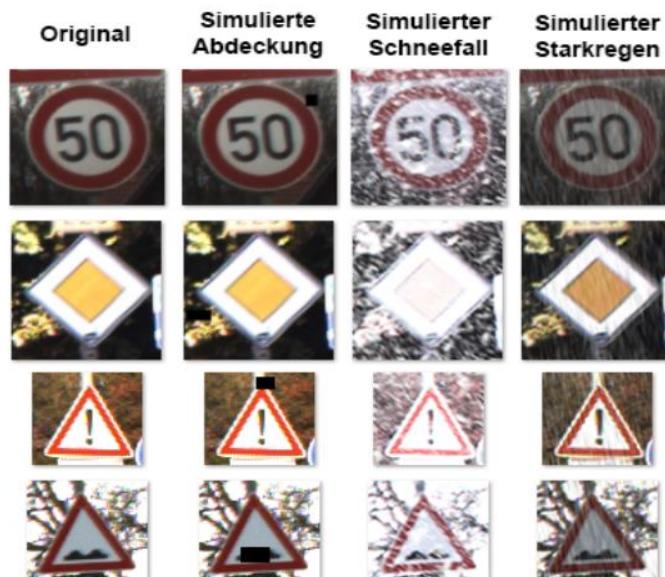


Abbildung 26: Datensätze mit mehreren Simulationen (Berghoff, Erkennung und Klassifizierung von Verkehrsschildern unter schwierigen Verhältnissen mit Convolutional Neural Network (CNN), 2020, S.5)

Um zu verdeutlichen, welche Grenzfälle es im öffentlichen Straßenverkehr gibt, wird hier ein simuliertes Praxisbeispiel anhand von Verkehrsschildern aufgezeigt.

In Abbildung 26 ist zu sehen, wie die verschiedenen Simulationen durchgeführt werden, um an alle für das Training benötigte Daten in Grenzfälle zu gelangen. Die Verkehrsschilder befinden sich im „Original“ im normalen Status, in der das Convolutional Neural Network in der Objekterkennung keine Probleme bekommen sollte. Es sind allerdings mehrere Grenzfälle möglich, die für die Datenverarbeitung berücksichtigt werden müssen, wie zum Beispiel Abdeckungen, Schneefall oder Starkregen. Grenzfälle verfügen über viele unterschiedliche Eigenschaften, weswegen es sinnvoll ist, mehrere Varianten dieser Grenzfälle zu simulieren, damit gewährleistet werden kann, dass auch alle Varianten abgedeckt sind. Jede Variante hat unterschiedliche Auswirkungen auf die Videoaufnahmen bei der virtuellen Simulation. Der simulierte Schneefall sorgt dafür, dass die Kamera vom Fahrzeug beeinträchtigt wird. Durch steigende Reflexion und der damit einhergehenden Erhöhung der Helligkeit folgt eine Reduzierung der Sättigung von Videoaufnahmen. Beim simulierten Starkregen sorgt es nicht für eine Erhöhung der Helligkeit, sondern für eine Reduzierung. Die Reduzierung der Helligkeit sorgt ebenfalls für eine Reduzierung des maximalen Kontrasts von Videoaufnahmen. (Berghoff, 2020, S. 4-5)

### 3. Intelligente Datenverarbeitung

Bei Tests in einer Stadt oder sogar in mehreren Städten derselben Region treten häufig gemeinsame Merkmale mit bestehenden, für Trainingszwecke markierten Daten zutage. Das hat zur Folge, dass zwei folgende Fragen bei der Intelligenten Datenverarbeitung beantwortet werden müssen:

- Welche aufgezeichneten Daten bei den Testfahrten sind von besonderem Interesse für das Convolutional Neural Network Modell?
- Welche Daten innerhalb der Aufzeichnungen sind nützlicher als die vorherigen gekennzeichneten Daten? (Nguyen, 2020, S. 36-37)

Um die Herausforderungen des Prozesses für die Intelligente Datenverarbeitung zu erleichtern, können Software-Lösungen wie zum Beispiel das Smart Data Harvesting für das autonome Fahren verwendet werden. Bei Verwendung der Software Smart Data Harvesting in Testfahrten mit dem Tool Intelligent Recording zeichnen die Sensoren alle Daten selektiv auf. Ohne bestimmte Software-Anwendungen werden normalerweise die Daten ungefiltert gesammelt. Die Software ist mit einem Algorithmus ausgestattet, die bereits während der Testfahrt benötigte Daten erkennt, analysiert und aufzeichnet. Somit ist eine frühe Erkennung und Selektierung der Intelligenten Datenauswahl gewährleistet. Es wird bei der Datenverarbeitung Zeit gespart für nicht wertschöpfende Tätigkeiten, da irrelevante Szenen nicht doppelt aufgenommen werden. Das Smart Data Harvesting ist nur eine Möglichkeit, um die Intelligente Datenauswahl voranzutreiben. (b-plus, 2022)

Es gibt viele verschiedene Software-Lösungen für eine Intelligenteren Datenauswahl und Datenverarbeitung. Das Convolutional Neural Network benötigt für sein Training eine große Anzahl an Daten zum Trainieren, um bei der Klassifikation eine gute Performance für das autonome Fahren zu erzielen. Es wird ebenfalls eine hohe Qualität an den Datensätzen vorausgesetzt, so dass Algorithmen wichtige Merkmale von den Verkehrsschildern lernen und diese auf unbekannte Daten anwenden können. Durch die Smart Data Harvesting Software kann eine hohe Qualität des Datensatzes gewährleistet sein. (Dariti, 2016, S. 18)

### **3.3 Potenzial für ein Convolutional Neural Network Modell mit den Datensätzen von GTSRB**

Nachdem aufgezeigt wurde, wie Daten per virtuelle Simulation oder einer realen Testfahrt gewonnen werden können und wie Datenspeicherung- und Verarbeitung funktioniert, wird in diesem Unterkapitel das Potenzial von Convolutional Neural Network mit den Daten der German Traffic Sign Recognition Benchmark aufgezeigt. Hier wird nochmal verdeutlicht, welche Potenziale mit dem Convolutional Neural Network ausgeschöpft werden können und weshalb ein gut trainiertes Convolutional Neural Network für das autonome Fahren geeignet sein kann.

German Traffic Sign Recognition Benchmark kann als Grunddatensatz verwendet werden, da es ein umfangreiches Datenmaterial besitzt, was zu einer Ersparnis von Zeit und Kosten führt. Somit wird die Arbeit erleichtert und eine Klassifizierung (Kennzeichnung der Daten) ist ebenfalls vorhanden (siehe Anhang A). Der GTSRB ist ein Wettbewerb zur Klassifizierung von Einzelbildern mit mehreren Klassen, der auf der International Joint Conference on Neural Network (IJCNN) 2011 stattfand. Im GTSRB befinden sich mehr als 40 Klassen und es stehen über 50.000 Bilder in einer naturgetreuen Datenbank für das Training für das Convolutional Neural Network Modell zur Verfügung. Allerdings beinhaltet das GTSRB keine Grenzfälle (kritische Situationen), wie zum Beispiel verdeckte Verkehrsschilder durch Regen oder Schnee. Aus diesem Grund dient es nur als Grundlage, wie eine Klassifizierung von Verkehrsschildern aussehen könnte. (Institut für Neuroinformatik, 2011)

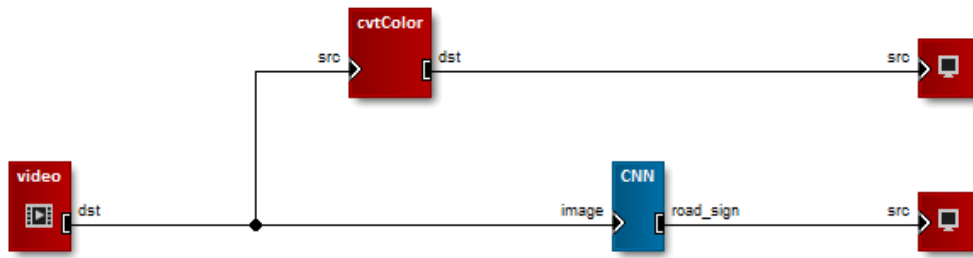


Abbildung 27: Funktionsweise CNN anhand von deutschen Verkehrsschildern (Online abgerufen: <http://www.durak-bolat.de/RSRwithCNN.html>; am 10. Juni 2022, 14:50 Uhr)

Der Datensatz von deutschen Verkehrsschildern von GTSRB kann zum Beispiel für das Training eines neuronalen Netzwerks als Fallbeispiel genutzt werden, um das Verständnis zu schärfen. Hierbei ist ein einfaches System ausreichend, um die Funktionsweise für die Abbildung 27 zu erläutern. Das Netzwerk soll in ein separates Modul, Convolutional Neural Network genannt, integriert und über eine Quelle mit Straßenaufnahmen von Verkehrsschildern versorgt werden. Am Ausgang emittiert das Convolutional Neural Network Modul synthetische Bilder der geschätzten Verkehrsschilder. Hierbei handelt es sich um realistische Bilder, die von einer künstlichen Intelligenz hergestellt wurden. Der obere Bereich dient zur Darstellung der Originalbilder, um die Schätzung mit diesen vergleichen zu können. Sind in beiden Bildern identische Verkehrsschilder zu sehen, dann hat das neuronale Netzwerk den Inhalt der Aufnahme mit einer 100-prozentigen Trefferquote geschätzt. Ist dies nicht der Fall, dann liegt eine Fehlklassifikation der Verkehrsschilder vor. Im Wettbewerb wurden verschiedene Methoden von Convolutional Neural Network angewendet, dessen Ergebnisse werden in Abbildung 28 vorgestellt. (Bolot, 2020)

TEAM	METHOD	TOTAL	SUBSET
[156] DeepKnowledge Seville	CNN with 3 Spatial Transformers	99.71%	99.71%
[3] IDSIA ★	Committee of CNNs	99.46%	99.46%
[155] COSFIRE	Color-blob-based COSFIRE filters for object recogn	98.97%	98.97%
[1] INI-RTCV ★	Human Performance	98.84%	98.84%
[4] sermanet ★	Multi-Scale CNNs	98.31%	98.31%
[2] CAOR ★	Random Forests	96.14%	96.14%
[6] INI-RTCV	LDA on HOG 2	95.68%	95.68%
[5] INI-RTCV	LDA on HOG 1	93.18%	93.18%
[7] INI-RTCV	LDA on HOG 3	92.34%	92.34%

Abbildung 28: GTSRB Results 2011 ([https://benchmark.ini.rub.de/gtsrb\\_results.html](https://benchmark.ini.rub.de/gtsrb_results.html), online abgerufen am 29.05.2022, 18:20 Uhr)

In Abbildung 28 ist zu erkennen, dass das Convolutional Neural Network mit drei räumlichen Transformatoren im Wettbewerb die höchste Trefferquote mit 99,71% bei der Erkennung von Verkehrsschildern erzielt hat. An zweiter Stelle belegt die Methode „Committee of CNNs“ mit einer Performance von 99,48%. Die menschliche Performance liegt an dritter Stelle bei lediglich 98,84%. Die Teams IDISIA, INI-RTCV, sermanet und CAOR sind mit einem Stern gekennzeichnet. Dieser Stern bedeutet, dass die Personen der jeweiligen Teams an der letzten Wettbewerbssitzung, die auf der IJCNN 2011 stattfand, teilgenommen haben. Hier ist deutlich zu erkennen, dass die Performance von Convolutional Neural Network von verschiedenen Methoden abhängt, um eine bessere Performance im Vergleich zu der menschlichen zu erwarten. Hier wird das Potenzial von Convolutional Neural Network eindeutig und ersichtlich dargestellt. (Ruhr-Universität-Bochum, 2011)

Nicht nur die Ruhr-Universität-Bochum hat diverse Tests in dem Bereich des Convolutional Neural Networks durchgeführt, sondern auch die technische Universität Eindhoven entwickelte ein solches Convolutional Neural Network Modell zur Erkennung von Verkehrsschildern für die Geschwindigkeitsbegrenzung. Dieses Network konnte mit einer Auflösung von 1280x720 und einer Wiederholungsrate von 35,7 Bilder/s die Verkehrsschilder mit einer hohen Trefferquote von 99,81% erkennen. Die menschliche lag lediglich bei 98,81%. Das trainierte Convolutional Neural Network Modell hat in diesem Praxisbeispiel von der technischen Universität Eindhoven ebenfalls eine höhere Performance bewiesen. (Technische Universität Eindhoven, 2012, S. 2)



### 3.4 Anwendungsbeispiel: Der gesamte Datenprozess für das Convolutional Neural Network

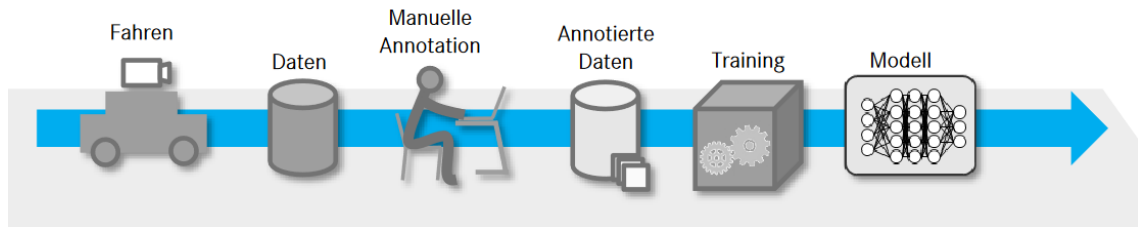


Abbildung 29: Der Weg zum Modell [Online abgerufen: [https://www.bmwk.de/Redaktion/DE/Downloads/W/ws5-praesentation-ki-und-daten.pdf?\\_\\_blob=publicationFile&v=4](https://www.bmwk.de/Redaktion/DE/Downloads/W/ws5-praesentation-ki-und-daten.pdf?__blob=publicationFile&v=4); Bundesministerium für Wirtschaft und Energie, 2020, S. 2)

In diesem Unterkapitel wird der komplette Prozess praxisnah von Anfang an von der Datensammlung bis hin zum Training des Convolutional Neural Network Modells schlüssig aufgezeigt.

In Abbildung 29 ist nochmal die komplette Prozesskette von der Datensammlung (reale Testfahrten oder virtuelle Simulationen) bis zum Training für das Convolutional Neural Network Modell dargestellt. Durch die Testfahrten (real oder virtuelle Simulation) werden die Daten gesammelt (Datengewinnung). (Bundesministerium für Wirtschaft und Energie, 2020, S. 25-28)

Für die Datensammlung können echte Autos auf den technisch aufgerüsteten Teststrecken, oder virtuelle Autos in den digitalen Teststrecken benutzt werden, die mit Sensoren (Kamera, Radar, Ultraschall und LiDAR-Sensoren) ausgestattet werden. Die Autos auf den realen Teststrecken werden von Testingenieuren gefahren. Die Ausstattung der Autos beinhaltet für die Datengewinnung ebenfalls Speichermedien wie Tape, HDD oder auch die Cloud. Um eine bessere Entscheidung treffen zu können, welche Speichermedien verwendet werden können, wurde aus diesem Grund die Vor- und Nachteile der verschiedenen Speichermedien in *Kapitel 3.3.2 Datenspeicherung* erläutert. (ZF, 2021)

Die Daten, die bei realen Teststrecken, Real-Drive und virtuelle Simulationen gesammelt werden, sind unstrukturierte Daten, weil sie in einer unverarbeiteten Rohform vorliegen. Bei realen Testfahrten können die Daten aus den genannten Teststrecken gewonnen werden. Bei virtuellen Testfahrten werden die Daten aus den virtuellen Testfeldern gewonnen. Der Fokus wird nicht bei den realen Teststrecken oder Real-Drive gesetzt, da es zu häufig zu unvorhersehbaren Verkehrssituationen kommen kann. Diese hohe

Quantität an Daten kann nicht durch reale Teststrecken oder Real-Drive nachgebildet werden, da es zu aufwändig wäre. Der Fokus bei der Datengewinnung liegt bei den virtuellen Testfeldern. Das Bundesministerium für Wirtschaft und Energie hat anhand einer Beispielrechnung aufgezeigt, wie die Aufteilung des Datenvolumens aussehen könnte. Das Datenvolumen an Teststrecken beträgt bei der Beispielrechnung lediglich 0,0001%, das sind ca. 100.000 km an reiner gefahrener Strecke. Das Datenvolumen von Real-Drive beträgt bei der Rechnung 0,1 %, das sind ca. 10.000.000 km an reiner gefahrener Strecke, wie im *Kapitel 3.3.1 Datengewinnung* aufgezeigt wird. Die meisten Daten werden durch virtuelle Simulationen gewonnen. Für die virtuelle Simulation wurden ebenfalls digitale Teststrecken gebaut. Das Datenvolumen durch virtuelle Simulationen bei den virtuellen Testfeldern beträgt bei der Rechnung 99,99%, das sind ca. 10.000.000.000 km. (Bundesministerium für Wirtschaft und Energie, 2020, S. 8)

Bei den Daten, die gewonnen werden, müssen ebenfalls die Vielzahl an Grenzfällen (kritische Situationen) im öffentlichen Straßenverkehr berücksichtigt werden, die auftreten können. In unserem Anwendungsbeispiel der Verkehrsschilder könnte es sich zum Beispiel um Grenzfälle wie verdeckte Verkehrsschilder durch Schneefall, Starkregen oder eine generelle Abdeckung handeln, wie bereits im vorherigen *Kapitel 3.3.3 Datenverarbeitung* erwähnt. (Berghoff, 2020, S. 4-5)

Alle gesammelten Daten werden dann in Rechenzentren gespeichert. Hier ergeben sich bei der Datenspeicherung der unstrukturierten Daten viele Herausforderungen. Die Herausforderung ist zum Beispiel, dass die Kapazität des Rechenzentrums für die Anhäufung an Daten vorhanden sein muss. Eine weitere Herausforderung ist zum Beispiel die Effizienz der Datenspeicherung, da die Kosten für den Stromverbrauch bei einem hohen Datenvolumen in die Höhe getrieben werden können. (Quantum, 2016, S. 43) Für einen möglichen Lösungsansatz für die Datenspeicherung kann zum Beispiel das Cloud Object Storage verwendet werden, da es ein Format für die Speicherung unstrukturierter Daten in der Cloud ist. Das Cloud Object Storage ist elastisch und flexibel und kann auf mehrere Petabytes skaliert werden, um ein unbegrenztes Datenwachstum von unstrukturierten Daten zu unterstützen. (Lelii, 2020)

Die aufgezeichneten Daten werden von Ingenieuren analysiert und ein definierter Datensatz von Metainformationen wird erstellt und hinzugefügt. Die Daten werden dann von den Ingenieuren klassifiziert. Ein Praxisbeispiel für eine mögliche Klassifizierung in einen kleinen Rahmen, wären die Daten von der German Traffic Sign Recognition Benchmark (GTSRB). (Institut für Neuroinformatik, 2011). Nachdem die Daten von den Ingenieuren klassifiziert wurden, können die Input-Daten in eine Software zur Analyse

des Convolutional Neural Network Modells transferiert werden. (International Business Machines Corporation, 2022)

Beim autonomen Fahren spielt es eine wichtige Rolle, dass das Convolutional Neural Network Modell bei der Objekterkennung im öffentlichen Straßenverkehr die Abhängigkeiten und Zusammenhänge der Datensätze (Vereinfachtes Beispiel: Ampel ist grün = losfahren) erkennt. Für das autonome Fahren wird aus diesem Grund die erste Variante des maschinellen Lernens, nämlich das überwachte Lernen eingesetzt. (Müller, 2021)

In dem iterativen Prozess evaluieren die Ingenieure verschiedenste Parameter des neuronalen Netzwerkes, um Modellierungsentscheidungen für das Convolutional Neural Network Modell, wie zum Beispiel die Schwellwertfunktion oder Aktivierungsfunktionen zu treffen. Hierbei besitzt die ReLu Funktion von den Aktivierungsfunktionen die meisten Vorteile für ein Convolutional Neural Network Modell, da es keine obere Beschränkung besitzt ( $\varphi(v) = \max(0, v)$ ) und dadurch ein effizienteres Training möglich ist. Weshalb die ReLu Funktion im Gegensatz zu den anderen Modellierungsentscheidungen allgemein für das Convolutional Neural Network Modell empfohlen wird, ist im *Kapitel 2.2 Künstliche Neuronale Netzwerke* ausführlich erläutert wurden. Nach der Modellierungsentscheidung und dem Iterativen Prozess wird das Convolutional Neural Network Modell für das autonome Fahren eingesetzt. (Lorenz, 2020, S. 123-124)

## **4 Forschungsstand, Problematik, weitere Herausforderungen und Ausblick**

In dem vorherigen Unterkapitel wurde der gesamte Datenprozess und das Training des Convolutional Neural Network Modells für das autonome Fahren dargestellt. Das Kapitel 4 konzentriert sich nun auf den Forschungsstand, die Problematik und die einhergehenden Herausforderungen, weshalb Stufe vier und fünf noch nicht im öffentlichen Straßenverkehr möglich sein können. Ein möglicher Ausblick für das autonome Fahren wird hier ebenfalls dargestellt.

Durch die Literaturlarbeit wurde der aktuelle Forschungsstand einschließlich der Möglichkeiten durch trainierte Convolutional Neural Network Modelle aufgezeigt. Das autonome Fahren der vierten und fünften Stufe im öffentlichen Straßenverkehr ist wünschenswert. Hierbei gibt es allerdings noch große Grenzen. Diese zeichnen sich durch die Fehlerquoten von trainierten Convolutional Neural Network Modellen und dem Mangel an Daten für die vierte und fünfte Stufe aus, wie anhand der Literaturquellen aufgezeigt wurde.

Laut dem Prof. Dr. Philip Slusallek besteht ein Datendefizit für das autonome Fahren der Stufe vier und fünf aufgrund mangelnder Kooperation der globalen Unternehmen. Philipp Slusallek ist ein Informatiker und Professor für Computergrafik an der deutschen Universität des Saarlandes. Weltweit muss es Kooperationen für die Sammlung von aufgezeichneten Daten geben. So werden viele Modelle in verschiedenen Ländern über den öffentlichen Straßenverkehr gebildet, wie zum Beispiel Bewegungen von Fahrradfahrern und Fußgängern. Ohne eine digitale Realität von mehreren Ländern wird es perspektivisch nicht möglich sein, autonome Fahrzeuge der Stufe vier und fünf im öffentlichen Straßenverkehr zu verwirklichen. Eine internationale Kraftanstrengung bei der Verteilung von lokalen Fahrdaten wird für eine Entwicklung robusterer Algorithmen sorgen. (Science media center - germany, 2021)

Neben dem Convolutional Neural Network spielen beim aktuellen Forschungsstand auch die Mikrocontroller eine bedeutsame Rolle. Mikrocontroller sind sehr kleine Computer, welche auf einem Chip angebracht werden und dadurch ein geschlossenes System bilden. Die Aurix Mikrocontroller haben weltweit eine Zertifizierung für das autonome Fahren erhalten. Laut aktuellen Stand bieten sie das höchste funktionale Sicherheitslevel sowie die schnellstmögliche Verbindung und Bewertung von Datenpaketen im Auto an. Der Aurix Mikrocontroller stellt vollständige Skalierbarkeit und Unterstützung für hochautomatisierte Autos ab Stufe zwei bis Stufe fünf bereit. (Schäfer, 2019)

Um neben dem aktuellen Forschungsstand auch die Problematik aufzuzeigen, wird hier ein Beispiel von Tesla aufgezeigt. Bereits 2022 sollen autonome Fahrzeuge in Deutschland der Stufe vier erlaubt werden, allerdings nicht im öffentlichen Straßenverkehr. Bei der vierten Stufe ist der Fahrer noch am Lenkrad und kann notfalls die autonome Fahrt unterbinden, jedoch hat die Regierung eine Begrenzung eingeführt. Diese sieht vor, dass der Einsatz der Fahrzeuge nur auf bestimmte Kurzstreckenbereiche wie zum Beispiel Shuttle-Verkehre, oder bei automatisierten Personentransporten auf kurzen Strecken agiert. Ein Beispiel hierfür ist das autonome Fahren von Tesla-Fahrzeugen, wovon einige in der Vergangenheit einen Unfall durch den verwendeten Autopilot, falsch reagierende Algorithmen, Datenmängel und Missverständnisse in den USA verursachten. Das Wort Missverständnis definiert den Begriff „Full Self-Driving“. Es suggeriert dem Fahrer, dass er auf die Technik im Gesamten vertrauen kann. Allerdings kommt diese Technik in vielen komplexen Verkehrssituationen nicht nach. Das ist einer der Hauptprobleme, weshalb die vierte und fünfte Entwicklungsstufe im Bereich des autonomen Fahrens im öffentlichen Straßenverkehr nicht möglich sein kann. (FINN, 2022)

Nun wird die Problematik in der aktuellen Gesetzeslage erläutert. Diese sieht vor, dass der Fahrer die gesamte Fahrt vom Ziel A bis B überwachen muss. Im Vergleich hierzu hat die fünfte Entwicklungsstufe des autonomen Fahrens das Ziel, dass der Fahrer nicht eingreifen und daher keine Überwachung durch ihn erfolgen muss. Aktuell gibt es noch keine autonomen Fahrzeuge der Stufe fünf. Sofern es diese bereits geben würde, dürften laut der aktuellen Gesetzeslage Fahrzeuge der Stufe fünf im deutschen öffentlichen Straßenverkehr nicht erlaubt sein. Die Gesetzeslage muss in diesem Fall ebenfalls eine Änderung vornehmen, so dass der Fahrer bei einem Fahrzeuge der Stufe fünf nicht die gesamte Fahrt überwachen muss. Im rechtlichen Kontext fehlen allerdings bestimmte Verordnungen, wie zum Beispiel die exakten Bauartvorschriften. Bis zur vierten und fünften Stufe müssen noch viele Regulierungen umgesetzt werden, damit die ersten autonomen Fahrzeuge im öffentlichen Straßenverkehr eingesetzt werden dürfen. (ADAC, 2022)

Weitere Barrieren der vierten und fünften Stufe im Bereich des autonomen Fahrens sind die stetig steigenden Energiekosten bei der Rechenleistung. Wie in Abbildung 24 zu erkennen ist, steigt der Energiebedarf bei erhöhtem Strombedarf stetig an. Die Fahrzeuge der vierten und fünften Stufe benötigen mehr Rechenleistung als Fahrzeuge bis zur dritten Stufe, da sich die Anforderungen und die Datenmenge erhöhen. Dies führt abermals zu einem erhöhten Energiebedarf, jedoch müssen in der vierten und fünften Stufe die erhöhten Anforderungen einer funktionalen Sicherheit gegeben sein. (Grai, 2019)

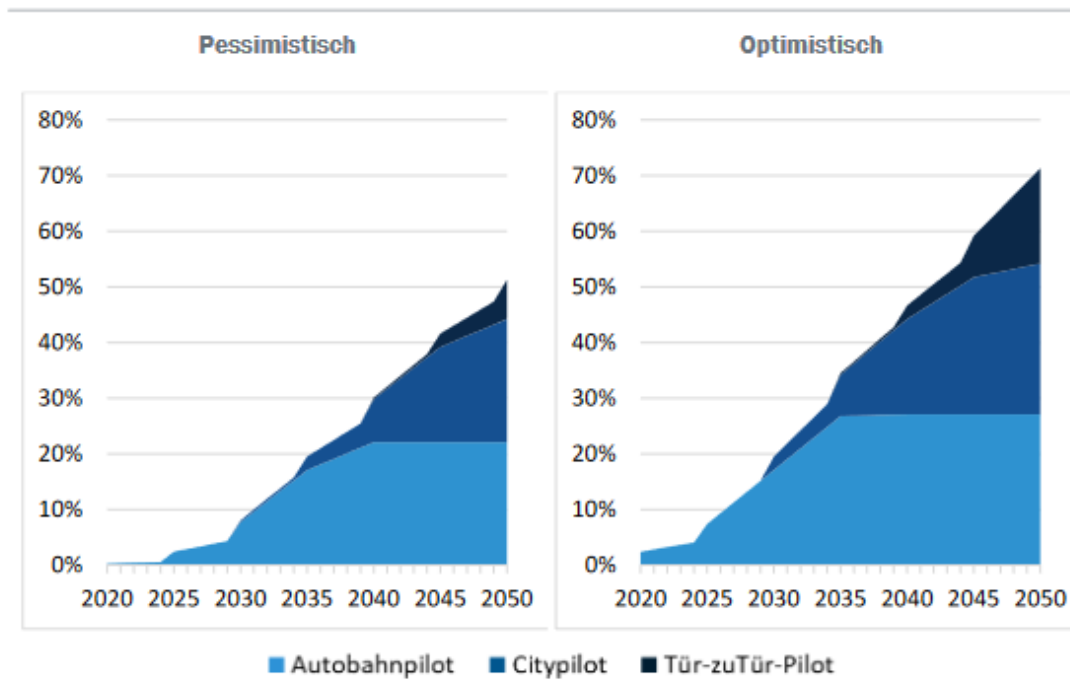


Abbildung 30: Anteil der Fahrzeuge mit Automatisierungsfunktionen an den Neuzulassungen (Prognos-Forschungsinstitut. Einführung von Automatisierungsfunktion in der Pkw-Flotte, 2018, S.15)

Der aktuelle Ausblick in Abbildung 30 bezieht sich auf die Studie des Prognos-Forschungsinstituts zum autonomen Fahren. Sie prognostiziert, dass sich das autonome Fahren nur langsam in der Zukunft durchsetzen wird. Der Begriff Citypilot repräsentiert die Bewegung des autonomen Autos auf den Autobahnen und in der Stadt. Tür-zu-Tür Pilot bedeutet, dass sich das Auto auch auf Landstraßen ohne einen Fahrer autonom fortbewegen kann. Das Prognos-Forschungsinstitut hat hierfür ein pessimistisches und optimistisches Szenario erstellt. Laut des Instituts wird beim optimistischen Szenario der Anteil von Neufahrzeugen im Jahr 2050 auf 70 geschätzt, bei denen der Fahrer auf Autobahnen, Städten und Landstraßen auf die gesamte Steuerung des Autos durch die Technik vertrauen kann. Im Vergleich dazu geht das pessimistische Szenario im Jahr 2050 von lediglich 50 Prozent aus. Laut dieser Statistik lässt sich schlussfolgern, dass das vollautomatisierte Fahren noch in ferner Zukunft liegt. Es liegt nahe, dass bald einige Fahrzeuge neben vollautomatisierten Fahrzeugen unterwegs sein werden. (prognos-Forschungsinstitut, 2018, S. 14-16)

Wie in Abbildung 30 zu erkennen ist, wird die Automatisierung im Straßenverkehr bis 2050 stetig zunehmen. Durch den Einsatz lernender Systeme, wie dem Convolutional Neural Network, wird die Fahrzeugautomatisierung zukünftig empathischer werden. Das autonome Fahren durch vernetzte Fahrzeuge kann sich nicht nur den öffentlichen Straßenverkehr effizienter gestalten, sondern zugleich auch den Personen- und Güterverkehr. Allerdings ergeben sich dabei einige Herausforderungen, zum Beispiel, dass sich

neben autonomen auch viele normale Fahrzeuge ohne Automatisierungstechniken, oder nur der Stufe eins und zwei im Straßenverkehr bewegen werden. Dadurch entsteht ein sogenannter Mischverkehr. Die Herausforderungen besteht darin, dass die Menge an Daten und die Datenqualität so hinreichend ist, damit die Interaktionen der lernenden Systeme mit anderen autonomen und nicht autonomen Autos sicher darstellbar sind, insbesondere für den Wechsel der Spur, das Eingliedern in den Straßenverkehr oder das Überqueren von Fahrbahnen. Ein Beispielszenario wäre ein hochautomatisiertes Fahrzeug, welches einem Menschen im Straßenverkehr Vorfahrt gewähren will. So muss das automatisierte System in einer gewissen Form die Vorfahrt gegenüber dem Menschen kommunizieren können. (AG Mobilität und Intelligente Systeme, 2019, S. 13-14)

Das Bundesministerium für Digitales und Verkehr hat in naher Zukunft neue Projekte zur Untersuchung des Mischverkehrs geplant. Das Ziel der Untersuchung besteht darin, viele Daten bezüglich Unfälle, Verkehrsmengen, Umweltdaten und Streckeninformationen zu erhalten. Es soll auf einem Streckenabschnitt der Autobahn A8 mit hochautomatisierten und autonomen Fahrzeugen (Stufe 3 und höher) in festgelegten zeitlichen Intervallen untersucht werden. Der Ausblick der Bundesregierung besteht darin, dass die Unfallzahlen drastisch minimiert werden, sobald autonomes Fahren der Stufe vier und fünf im öffentlichen Straßenverkehr möglich ist. Aktuell sind 90 Prozent der Unfälle im öffentlichen Straßenverkehr auf menschliches Fehlverhalten zurückzuführen. (Bundesministerium für Digitales und Verkehr, 2021)

## 5 Fazit

Die vorliegende Bachelorarbeit ging der Frage nach, warum das Convolutional Neural Network für das autonome Fahren geeignet sein kann, welche Daten dafür benötigt werden und woher die Datengewinnung erfolgt. Ebenfalls wurden die allgemeinen Herausforderungen und die Herausforderungen im gesamten Prozess der Datensammlung, -Speicherung und -Verarbeitung dargestellt. Für die Beantwortung der Forschungsfragen wurde eine quantitative Literaturliteraturarbeit auf Grundlage der neuesten wissenschaftlichen Erkenntnissen durchgeführt.

Aus den bisherigen Ergebnissen, wie in Abbildung 28 dargestellt, lässt sich schließen, dass bei einer guten Datensammlung, Datenspeicherung und Datenverarbeitung eine höhere Erfolgsquote bei der Objekterkennung erzielt werden kann im Vergleich zu einer einzelnen Person. Dies ergibt sich aus den jeweiligen Rechenzentren und der Weiterentwicklung dieser durch ein trainiertes Convolutional Neural Network Modell.

Durch die quantitative Literaturliteraturarbeit wurde somit gezeigt, dass das Convolutional Neural Network für das autonome Fahren geeignet sein kann. Das Convolutional Neural Network hat mit der richtigen Dateninfrastruktur, Datenmanagement und Methode einschließlich der jeweiligen Datensätze eine höhere Trefferquote bei der Objekterkennung im Vergleich zu einem Menschen. Somit kann ein trainiertes Convolutional Neural Network Modell beim autonomen Fahren ein geringeres Unfallrisiko im Straßenverkehr aufweisen. Sowohl in den Experimenten der technischen Universität Eindhoven als auch im Wettbewerb IJCNN 2011, erzielte die Mehrheit der teilgenommenen Kandidaten eine höhere Trefferquote bei der Erkennung von Verkehrsschildern aufgrund der Integration ihrer Convolutional Neural Networks.

Damit im Vergleich zum Menschen in der Objekterkennung im öffentlichen Straßenverkehr für die breite Masse höhere Trefferquoten erzielt werden können, müssen die auftretenden Herausforderungen in der Datensammlung, Datenspeicherung und Datenverarbeitung bewältigt werden. Wie bereits in dieser Abschlussarbeit erläutert, muss die hohe Quantität an gesammelten Daten und die damit einhergehenden Herausforderungen in der effizienten Nutzung der Kapazität in den Rechenzentren in Zukunft angegangen werden. Es führt kein Weg vorbei an leistungsfähigen Rechenzentren für den gesamten aufwendigen Datenprozess.

Um die Herausforderungen bezüglich der Datensammlung und Datenspeicherung im Prozess zu bewältigen, kann auf verschiedene Software-Lösungen zurückgegriffen werden, die letztlich für Effizienz sorgen.



Es muss jedoch berücksichtigt werden, dass auch dort keine 100 prozentige Trefferquote bei der Objekterkennung erreicht wurde. Obwohl eine bessere Performance vorhanden ist, besteht immer noch ein gewisses Unfallrisiko im Straßenverkehr. Das autonome Fahren kann mit dem aktuellen Forschungsstand lediglich bis zur dritten Stufe im öffentlichen Straßenverkehr zugelassen werden. Es darf dabei nicht unerwähnt bleiben, dass Unfälle bisher noch nicht zu 100 % auf der vierten und fünften Stufe ausgeschlossen werden können. Aus diesem Grund gestaltet sich die Umsetzung bisher als problematisch. Da autonome Fahrzeuge der Stufe vier und fünf bisher noch keine Zulassung für den öffentlichen Straßenverkehr haben, können hier aufgrund fehlender Informationen keine Vergleiche zu der Unfallwahrscheinlichkeit mit der menschlichen Performance gezogen werden, sodass die Forschung in diesem Bereich eingeschränkt ist. Das Ziel des Convolutional Neural Networks besteht allerdings nicht darin, die 100 prozentige Erfolgsquote zu erreichen, da eine Perfektion in diesem Sinne nicht erreichbar ist. Das Ziel besteht darin, eine bessere Performance als ein Mensch im öffentlichen Straßenverkehr zu erzielen. Dadurch können die beiden Stufen vier und fünf im öffentlichen Straßenverkehr möglich gemacht werden, wodurch eine Reduzierung der Unfallzahlen erreicht werden kann. In diesem Bereich kann sich anschließend weiterer Forschungsbedarf ergeben.

In der aktuellen Gesetzeslage ist das autonome Fahren lediglich bis zur dritten Stufe im öffentlichen Straßenverkehr erlaubt. Da mögliche Unfälle beim autonomen Fahren auf der vierten und fünften Stufe vom autonomen Systemen und nicht von Menschen verursacht werden würden, sind juristische Grundlagen wie Verantwortlichkeiten und Haftbarkeiten zu definieren. In diesem Bereich müssen weitere Forschungen durchgeführt werden.

Zusammenfassend lässt sich sagen, dass das trainierte Convolutional Neural Network Modell mit der richtigen Datenqualität und dem optimalen Training für das autonome Fahren geeignet sein kann. In der Zukunft können im optimistischen Szenario 70 Prozent der Fahrzeuge bis zum Jahr 2050 im Straßenverkehr autonom unterwegs sein, wie die Studie Prognos Forschungsinstitut aufgezeigt hat. Um dies zu ermöglichen, werden beim gesamten Datenprozess hoch qualifizierte Ingenieure gebraucht. Mit den umfangreichen Software-Lösungen, die im Markt angeboten werden, kann der gesamte Datenprozess für das autonome Fahren in Zukunft beschleunigt werden.

## Literaturverzeichnis

- ADAC. (04. Januar 2022). *Gesetz zum autonomen Fahren: Diese Regeln gelten*. Abgerufen am 17. Juni 2022 von Autonomes Fahren - Verordnungen: <https://www.adac.de/rund-ums-fahrzeug/ausstattung-technik-zubehoer/autonomes-fahren/recht/gesetz-zum-autonomen-fahren/#hochautomatisiertes-fahren>
- AG Mobilität und Intelligente Systeme. (2019). *Auf dem Weg zu einem intelligenten Mobilitätsraum - Handlungsfelder, Chancen und Herausforderungen*. (L. S.-D. Intelligenz, Hrsg.) München. Abgerufen am 21. Juni 2022 von [https://www.plattform-lernende-systeme.de/umfeldszenario-intelligent-vernetzt-unterwegs.html?file=files/Downloads/Publikationen/AG5\\_Mobilitaet\\_web.pdf](https://www.plattform-lernende-systeme.de/umfeldszenario-intelligent-vernetzt-unterwegs.html?file=files/Downloads/Publikationen/AG5_Mobilitaet_web.pdf)
- Aichele, C., & Jörg, H. (2021). *Betriebswirtschaftliche KI-Anwendungen - Digitale Geschäftsmodelle auf Basis Künstlicher Intelligenz* (1. Ausg.). Hochschule Kaiserslautern, Zweibrücken, Deutschland: Springer Vieweg.
- Anthony Martinez, M., Sitawarin, C., Finch, K., Meincke, L., Yablonski, A., & Kornhauser, A. (2017). *Beyond Grand Theft Auto V for Training, Testing and Enhancing Deep Learning in Self Driving Cars*. Princeton, New Jersey, Vereinigte Staaten: Princeton University. Abgerufen am 8. Mai 2022 von <https://arxiv.org/pdf/1712.01397.pdf>
- Audi. (2021). *Autonomes Fahren auf dem Weg zur gesellschaftlichen Akzeptanz*. Fachbereich Künstliche Intelligenz. Ingolstadt: Audi SocAlty Studie. Abgerufen am 13. Mai 2022 von [https://www.audi.com/content/dam/gbp2/company/research/audi-beyond/2021/AUDI\\_SocAlty\\_Studie\\_dgtl\\_1206\\_Deutsch\\_small.pdf](https://www.audi.com/content/dam/gbp2/company/research/audi-beyond/2021/AUDI_SocAlty_Studie_dgtl_1206_Deutsch_small.pdf)
- Audi. (2. Dezember 2021). *Summary: Die wichtigsten Erkenntnisse*. (AUDI, Herausgeber) Abgerufen am 13. Mai 2022 von Summary zur SocAlty Studie 2021 der Audi Initiative: <https://www.audi.com/de/company/research/and-audi-initiative/audi-autonomous-vehicles-socalty-study-2021-summary.html>
- Babel, W. (2021). *Industrie 4.0, China 2025, IoT - Der Hype um die Welt der Automatisierung* (1. Ausg., Bd. I). Wiesbaden, Deutschland: Springer Vieweg. Abgerufen am 13. Juni 2022 von <https://link.springer.com/content/pdf/10.1007/978-3-658-34718-5.pdf>

- Berghoff, S. (2020). *Erkennung und Klassifizierung von Verkehrsschildern unter schwierigen Verhältnissen mit Convolutional Neural Network (CNN)* (Bd. I). Hamburg, Hamburg, Deutschland: Hochschule für Angewandte Wissenschaften Hamburg. Abgerufen am 20. Mai 2022 von <https://autonomesysteme.informatik.haw-hamburg.de/papers/2020SebastianBerghoff.pdf>
- BMVi. (10. September 2020). *Digitale Testfelder für das automatisierte und vernetzte Fahren im Realverkehr in Deutschland*. Abgerufen am 20. Juni 2022 von autonomes Fahren: [https://www.bmvi.de/SharedDocs/DE/Anlage/DG/Digitales/uebersicht-digitale-testfelder-avf-bmvi.pdf?\\_\\_blob=publicationFile](https://www.bmvi.de/SharedDocs/DE/Anlage/DG/Digitales/uebersicht-digitale-testfelder-avf-bmvi.pdf?__blob=publicationFile)
- Bolat, D. (10. Februar 2020). *Bildverarbeitung und Mustererkennung.*, 1. Abgerufen am 08. Juni 2022 von Schilderererkennung mit Convolutional Neural Network (CNN): <http://www.durak-bolat.de/RSRwithCNN.html>
- b-plus. (1. Januar 2022). *Smart Data Harvesting*. (b-plus, Herausgeber) Abgerufen am 16. Mai 2022 von Daten-Qualität als Basis für sichere Funktionen von automatisiertem Fahren: <https://www.b-plus.com/de/loesungen/smart-data-harvesting>
- Breitfuß, M. (2020). *Extraktion und Klassifikation realer Verkehrsszenarien zur Absicherung autonomer Fahrzeuge* (Bd. I). Maria Alm, Österreich: Technischen Universität München. Abgerufen am 2 von <https://mediatum.ub.tum.de/doc/1553700/file.pdf>
- Breyer-Mayländer, T. (2022). *Industrie 4.0 bei Hidden Champions - Digitalisierungsstrategien und ihre Anwendungsfelder in innovativen Industrieunternehmen* (1. Ausg., Bd. I). (T. Breyer-Mayländer, Hrsg.) Hochschule Offenburg, Baden-Württemberg, Deutschland: Springer Gabler. Abgerufen am 15. April 2022
- Brinkmann, N. (2016). *Objekterkennung in natürlichen Szenen mittels Region-based Convolutional Neural Networks*. Dortmund, Nordrhein-Westfalen, Deutschland: technische Universität dortmund. Abgerufen am 10. Juni 2022 von [https://patrec.cs.tu-dortmund.de/pubs/theses/ba\\_brinkmann.pdf](https://patrec.cs.tu-dortmund.de/pubs/theses/ba_brinkmann.pdf)
- Brutalk. (10. Februar 2021). *So beheben Sie das Problem mit verschwindenden Verläufen mithilfe der ReLU*. Abgerufen am 13. Juni 2022 von ReLu: <https://www.brutalk.com/de/nachrichten/brutalk-blogg/visa/so-beheben-sie-das-problem-mit-verschwindenden-verlaeufen-mithilfe-der-relu-604702ee0f56d>

- Bundesministerium für Digitales und Verkehr. (20. Juli 2021). *Neue Projekte zur Untersuchung des Mischverkehrs geplant*. (bmvi, Herausgeber) Abgerufen am 21. Juni 2022 von Mischverkehr: <https://www.bmvi.de/SharedDocs/DE/Artikel/DG/mobilitaet-der-zukunft-zwei-weitere-testfelder.html>
- Bundesministerium für Wirtschaft und Energie. (2020). KI und Daten - Herausforderungen auf dem Weg zum autonomen Fahren. In eu20.de (Hrsg.), *Workshopreihe zum Förderprogramm "Zukunftsinvestitionen Fahrzeughersteller und Zulieferindustrie" (Konjunkturpaket 35c). I*, S. 43. Berlin: eu20.de. Abgerufen am 15. Mai 2022 von [https://www.bmwk.de/Redaktion/DE/Downloads/W/ws5-praesentation-ki-und-daten.pdf?\\_\\_blob=publicationFile&v=4](https://www.bmwk.de/Redaktion/DE/Downloads/W/ws5-praesentation-ki-und-daten.pdf?__blob=publicationFile&v=4)
- Capgemini. (15. November 2019). *Datengetriebene Unternehmen erzielen 70 Prozent mehr Umsatz je Mitarbeiter*. Abgerufen am 15. April 2022 von Capgemini: <https://www.capgemini.com/de-de/news/studie-datengetriebene-unternehmen/>
- Dariti, M. (2016). *Bachelorarbeit - Erkennung von Verkehrszeichen mit Konvolutionalen Neuronalen Netzen*. Hamburg, Hamburg, Deutschland: Hochschule für Angewandte Wissenschaften Hamburg. Abgerufen am 30. April 2022
- Data Science Team. (10. Februar 2020). *Faltungsneurale Netzwerke (Convolutional Neural Network)*. Abgerufen am 21. Juni 2022 von CNN: <https://datascience.eu/de/computer-vision/faltungsneurale-netzwerke/>
- Deutscher Bundestag. (2021). *Energieverbrauch von Rechenzentren*. Energie. Berlin: Deutscher Bundestag. Abgerufen am 20. Juni 2022 von <https://www.bundestag.de/resource/blob/863850/423c11968fcb5c9995e9ef9090edf9e6/WD-8-070-21-pdf-data.pdf>
- Dimitri, G. (10. April 2016). *Maschinelle Bilderkennung mit Big Data und Deep Learning*, 4. Abgerufen am 20. Mai 2022 von [entwickler.de](https://entwickler.de/machine-learning/einblicke-in-die-konigsdisziplin-001): <https://entwickler.de/machine-learning/einblicke-in-die-konigsdisziplin-001>
- DIVISIO. (24. Mai 2019). *Überwachtes & unüberwachtes Lernen im ML*. Abgerufen am 26. Mai 2022 von Blog: <https://divis.io/2019/05/ki-leicht-erklart-teil-5-ueberwachtes-unueberwachtes-lernen-ml/>
- Ege, B., & Paschke, A. (2020). Semantische Datenintelligenz im Einsatz. In B. Ege, & A. Paschke (Hrsg.). Berlin; Wien, Deutschland; Österreich: Springer Vieweg. Abgerufen am 25. April 2022

- Ellen, L. (4. Januar 2022). *Der Unterschied zwischen Daten, Information und Wissen*. (Gesundheitswesen, Herausgeber, & Digital Leadership) Abgerufen am 15. April 2022 von <https://www.informationsdienste-langenstein.de/2018/11/21/wissensgesellschaft-i-unterschied-zwischen-daten-informationen-und-wissen-in-datenbest%C3%A4nden/>
- Ertel, W. (2008). *Grundkurs Künstliche Intelligenz*.
- Ertel, W. (2008). *Grundkurs Künstliche Intelligenz*. Wiesbaden, Hessen, Deutschland: Springer.
- Ertel, W. (2021). *Grundkurz Künstliche Intelligenz - Eine praxisorientierte Einführung* (5. Ausg., Bd. VI). Weingarten, Baden-Württemberg, Deutschland: Springer Vieweg.
- FINN. (31. Januar 2022). *Autonom fahren – ab wann ist es möglich?* Abgerufen am 17. Juni 2022 von autonomes Fahren: <https://www.finn.auto/guide/when-will-cars-drive-autonomously>
- Fischer, J. (10. Februar 2017). *Neuronale Netze*, 1.5.3. (J. Fischer, Produzent) Abgerufen am 05. August 2021 von Neuronale Netze: [https://user.phil.hhu.de/~petersen/SoSe17\\_Teamprojekt/AR/neuronalenetze.html](https://user.phil.hhu.de/~petersen/SoSe17_Teamprojekt/AR/neuronalenetze.html)
- Fischer, J., & Pochwyt, K. (10. Mai 2017). *Convolutional Neural Networks am Beispiel eines selbstfahrenden Roboters*. Abgerufen am 23. April 2022 von Convolutional Neural Networks am Beispiel eines selbstfahrenden Roboters: [https://user.phil.hhu.de/~petersen/SoSe17\\_Teamprojekt/AR/neuronalenetze.html](https://user.phil.hhu.de/~petersen/SoSe17_Teamprojekt/AR/neuronalenetze.html)
- Folkers, A. (2019). *Steuerung eines autonomen Fahrzeugs durch Deep Reinforcement Learning*. Bremen, Bremen, Deutschland: Springer Spektrum. Abgerufen am 15. April 2022
- Franken, R. (2020). *Wissen, Lernen und Innovationen im digitalen Unternehmen* (2. Ausg., Bd. II). (S. Franken, Hrsg.) Technische Hochschule Köln, Nordrhein-Westfalen, Deutschland: Springer Gabler. Abgerufen am 25. April 2022
- Freeden, W., & Rummel, R. (2017). *Photogrammetrie und Fernerkundung* (Bd. I). (C. Heipke, Hrsg.) Hannover, Deutschland: Springer Spektrum.
- Friedel, D. (21. August 2020). *Autonomes Fahren*, 1. Abgerufen am 16. Mai 2022 von DIGITALE WELT: <https://digitaleweltmagazin.de/fachbeitrag/autonomes-fahren-beginnt-im-rechenzentrum/>

- Grai, T. (25. November 2019). *Die 5 Herausforderungen auf dem Weg zum autonomen Fahrzeug*. (all-electronics, Herausgeber) Abgerufen am 23. Juni 2022 von Herausforderungen: <https://www.all-electronics.de/automotive-transportation/die-5-herausforderungen-auf-dem-weg-zum-autonomen-fahrzeug.html>
- Grünwald, R. (3. August 2020). *Künstliches neuronales Netz einfach erklärt: Lernen im Data Mining*, 1. (Novustat, Herausgeber) Abgerufen am 25. April 2022 von Künstliches neuronales Netz einfach erklärt: Lernen im Data Mining: <https://novustat.com/statistik-blog/kuenstliches-neuronales-netz-einfach-erklaert.html>
- Heinrich, L., Heinzl, A., & Riedl, R. (2011). *Wirtschaftsinformatik - Einführung und Grundlagen* (4. Ausg., Bd. IV). Mannheim, Baden-Württemberg, Deutschland: Springer. Abgerufen am 9. Mai 2022
- Herrmann, J. (2021). *Betriebswirtschaftliche KI-Anwendungen*. (J. Herrmann, Hrsg.) Hochschule Kaiserslautern, Rheinland-Pfalz, Deutschland: Springer Vieweg. Abgerufen am 15. April 2022
- Hertwig, F. (11. Juli 2018). *Mailborn Wolf - Mensch IT*. (F. Hertwig, Herausgeber, & F. Hertwig, Produzent) Abgerufen am 20. Juli 2021 von Mailborn Wolf - Mensch IT: <https://www.maibornwolff.de/blog/geschichte-von-neural-networks-und-deep-learning>
- Hildebrand, K., Gebauer, M., & Mielke, M. (2021). *Daten- und Informationspolitik* (5. Ausg., Bd. V). (M. Mielke, Hrsg.) Frankfurt am Main, Hessen, Deutschland: Springer Vieweg. Abgerufen am 09. Mai 2022
- Hölscher, M. (2021). *Künstliche Neuronale Netze als Instrument der Werttreiberanalyse im Kontext der wertorientierten Unternehmenssteuerung am Beispiel des EVA* (1 Ausg.). Universität Bielefeld, Deutschland, Nordrhein-Westfalen, Deutschland: Springer Gabler. Abgerufen am 10. März 2022
- Hoogenraad, W. (20. Juli 2021). *Skalierbarkeit als Softwareanforderung, Bedeutung und Definition*. Abgerufen am 26. Mai 2022 von ITpedia: <https://de.itpedia.nl/2021/07/20/schaalbaarheid-als-software-requirement-betekenis-en-definitie/>
- Horsting, T. (03. September 2019). *Datenbanken: Wozu man sie braucht und welche Arten es gibt*. Abgerufen am 21. Juni 2022 von Datenbank: <https://www.ionos.de/digitalguide/hosting/hosting-technik/datenbanken/>

- Industrie- und Handelskammer Nürnberg für Mittelfranken. (20. Dezember 2018). *Netze für Nachhaltigkeit Nordbayern - Nürnberg*. (I. u. Mittelfranken, Herausgeber) Abgerufen am 28. Mai 2022 von Energieeffizienz im Rechenzentrum - Optimieren Sie jetzt den Stromverbrauch Ihrer IT: <https://www.nuernberger-netze.de/energieeffizienz-im-rechenzentrum-optimieren-sie-jetzt-den-stromverbrauch-ihrer-it/>
- Institut für Neuroinformatik. (16. September 2011). *The German Traffic Sign Recognition Benchmark*, 1. Abgerufen am 20. Mai 2022 von The German Traffic Sign Benchmark: [https://benchmark.ini.rub.de/gtsrb\\_news.html](https://benchmark.ini.rub.de/gtsrb_news.html)
- International Business Machines Corporation. (4. Januar 2022). *IBM FlashSystem: Flash-Datenspeicher*. (IBM, Herausgeber) Abgerufen am 16. Mai 2022 von Speicher mit Cyber-Resilienz : [https://www.ibm.com/de-de/storage/flash?utm\\_content=SRCWW&p1=Search&p4=43700068028905254&p5=e&gclid=CjwKCAjw7leUBhBbEiwADhiEMWBQGwIRC52koubnvWKAF4n9-13BLA9-En5Y5qAclurO9dd7X1yYtxoCb5EQAvD\\_BwE&gclsrc=aw.ds](https://www.ibm.com/de-de/storage/flash?utm_content=SRCWW&p1=Search&p4=43700068028905254&p5=e&gclid=CjwKCAjw7leUBhBbEiwADhiEMWBQGwIRC52koubnvWKAF4n9-13BLA9-En5Y5qAclurO9dd7X1yYtxoCb5EQAvD_BwE&gclsrc=aw.ds)
- Jesgulke, O. (23. August 2016). *Toll Collect - service on the road*. Abgerufen am 30. April 2022 von Autonomes Fahren auf der Überholspur: <https://blog.toll-collect.de/autonomes-fahren-auf-der-ueberholspur/>
- Kemper, H.-G. (2021). *Business Intelligence & Analytics - Grundlagen und praktische Anwendungen* (1. Ausg.). Universität Stuttgart, Baden-Württemberg, Deutschland: Springer Vieweg. Abgerufen am 14. April 2022
- Keras. (26. März 2022). *How ReLU works in convolutional neural network*. Abgerufen am 17. Juni 2022 von ReLu: <https://androidkt.com/how-relu-works-in-convolutional-neural-network/>
- Klaus, L. (10. Juli 2019). *ITWissen.info*. (D. B. GmbH, Herausgeber) Abgerufen am 15. April 2022 von ITWissen.info: <https://www.itwissen.info/Daten-data.html>
- Koch, K. (2018). *Autonomes Fahren. Ist die autonomie Fahrzeugführung aus technischer, rechtlicher und gesellschaftlicher Sicht realisierbar?* (1. Ausg., Bd. I). Nürtingen, Baden-Württemberg, Deutschland: Hochschule für Wirtschaft und Umwelt Nürtingen-Geislingen. Abgerufen am 11. Mai 2022
- Komarnicki, P., Haubrock, J., & Styczynski, Z. (2020). *Elektromobilität und Sektorenkopplung - Infrastruktur- und Systemkomponenten* (2. Ausg.). Bielefeld, Nordrhein-Westfalen, Deutschland: SpringerVieweg.

- Kröger, J. (14. Juni 2021). *IT-Service.Network*, 1. Abgerufen am 12. Mai 2022 von Was ist ein Rechenzentrum?: <https://it-service.network/blog/2021/06/14/was-ist-ein-rechenzentrum/>
- Krol, B. (2021). *Explainable Artificial Intelligence: Analyse und Visualisierung des Lernprozesses eines Convolutional Neural Network zur Erkennung deutscher Straßenverkehrsschilder* (Bd. XXVI). Essen, Nordrhein-Westfalen, Deutschland: FOM Hochschule. Abgerufen am 20. Juni 2022
- Künstliche Neuronale Netzwerke und Deep Learning, S. S. (2018). *Künstliche Neuronale Netzwerke und Deep Learning*.
- Lang, N. (28. November 2021). *Data Base Camp*. Abgerufen am 25. April 2022 von Was sind Convolutional Neural Networks (CNN)?: [https://databasecamp.de/ki/convolutional-neural-network#Recap\\_Neuronale\\_Netzwerke\\_Fully-Connected](https://databasecamp.de/ki/convolutional-neural-network#Recap_Neuronale_Netzwerke_Fully-Connected)
- Laurenz, W. (13. September 2020). *Reinforcement Learning: Wenn KI auf Belohnungen reagiert - Deep Learning Grundlagen*. Abgerufen am 26. April 2022 von datasolut: <https://datasolut.com/reinforcement-learning/>
- Lelii, S. (10. Februar 2020). *Techtarget*, 1. Abgerufen am 16. Mai 2022 von cloud object storage : <https://www.techtargget.com/searchstorage/definition/cloud-object-storage>
- Leyendecker, B. (2020). *Qualitätsmanagement in den 20er Jahren - Trends und Perspektiven* (Bd. I). Koblenz, Rheinland-Pfalz, Deutschland: Springer Vieweg. Abgerufen am 30. April 2020
- Lorenz, U. (2020). *Reinforcement Learning - Aktuelle Ansätze verstehen - mit Beispielen in Java un Greenfoot* (1. Ausg.). Neckargemünd, Baden-Württemberg, Deutschland: Springer Vieweg. Abgerufen am 13. Juni 2022 von <https://link.springer.com/content/pdf/10.1007/978-3-662-61651-2.pdf>
- Lynn, J. (2019). *Analyse und Implementation von Methoden zur Spurmarkierungserkennung aus Bilddaten zum Generieren von virtuellen Teststrecken* (1. Ausg., Bd. I). Berlin, Berlin, Deutschland: Freie Universität Berlin. Abgerufen am 11. Mai 2022
- MathWorks. (10. Mai 2021). *Convolutional Neural Network - Drei Dinge, die Sie über Convolutional Neural Networks wissen sollten*. Abgerufen am 28. Mai 2022 von CNN: <https://de.mathworks.com/discovery/convolutional-neural-network-matlab.html>



- Matthias, P. (2021). *Data Science - Grundlagen, Statistik und maschinelles Lernen* (1. Ausg.). (S. Spektrum, Hrsg.) Berlin, Berlin, Deutschland: Springer Spektrum. Abgerufen am 15. April 2022
- Matzka, S. (2021). *Künstliche Intelligenz in den Ingenieurwissenschaften - Maschinelles Lernen verstehen und bewerten* (1. Ausg.). Berlin, Berlin, Deutschland: Springer Vieweg. Abgerufen am 24. April 2022
- Maurer, M., Lenz, B., & Gerdes, C. (2015). *Autonomes Fahren - Technische, rechtliche und gesellschaftliche Aspekte*. (H. Winner, Hrsg.) Braunschweig, Deutschland: Springer Open. Abgerufen am 20. Juni 2022 von <https://library.oapen.org/bitstream/handle/20.500.12657/27813/1002192.pdf?sequence=1#page=376>
- Meyer, H., Dieckmann, T., & Laßmann, P. (3. April 2019). Automatisiertes Fahren ganzheitlich betrachtet. *Tagungsbericht*, 80-81. Abgerufen am 19. Mai 2022 von <https://link.springer.com/content/pdf/10.1007/s35148-019-0074-3.pdf>
- Mockenhaupt, A. (2021). *Digitalisierung und Künstliche Intelligenz in der Produktion - Grundlagen und Anwendung* (1. Ausg.). Sigmaringen, Baden-Württemberg, Deutschland: Springer Vieweg. Abgerufen am 24. April 2022
- Müller, S. (2021). *Supervised Learning*. (datasolut, Hrsg.) Köln, Nordrhein-Westfalen, Deutschland: datasolut. Abgerufen am 13. Juni 2022 von <https://datasolut.com/wiki/supervised-learning/>
- Neiden, M. (2017). *Analyse hochdimensionaler MALDI-Daten mittels Convolutional Neural Networks* (1. Ausg.). Universität zu Lübeck, Institute of Mathematics and image Computing, Schleswig-Holstein, Deutschland: Universität zu Lübeck.
- Nguyen, T. (05. April 2020). Autonomes Fahren beginnt im Rechenzentrum. *DIGITALE WELT*, 1-2. Abgerufen am 8. Mai 2022
- NVIDIA. (22. Juni 2021). *Tesla Unveils Top AV Training Supercomputer Powered by NVIDIA A100 GPUs*. (NVIDIA, Herausgeber) Abgerufen am 18. Mai 2022 von "Incredible" GPU cluster powers AL development for Autopilot and full self-driving: <https://blogs.nvidia.com/blog/2021/06/22/tesla-av-training-supercomputer-nvidia-a100-gpus/>
- Oracle University. (02. April 2020). *Was ist das IoT?* (O. D. KG, Herausgeber) Abgerufen am 05. Mai 2022 von Was ist das IoT?: <https://www.oracle.com/de/internet-of-things/what-is-iot/>

- Peter Buxmann, H. S. (2019). *Künstliche Intelligenz - Mit Algorithmen zum wirtschaftlichen Erfolg*. Technische Universität Darmstadt, Deutschland: Springer Gabler.
- Petit, F. (17. Juni 2020). *Sensorfusion – Schlüsselkomponenten für autonomes Fahren*. (BlickfeldBlog, Herausgeber) Abgerufen am 21. Juni 2022 von Sensoren: <https://www.blickfeld.com/de/blog/sensorfusion/>
- Plaue, M. (2021). *Data Science - Grundlagen, Statistik und maschinelles Lernen*. Berlin, Berlin, Deutschland: Springer Spektrum. Abgerufen am 25. April 2022
- Preuss, P. (2017). *In-Memory-Datenbank SAP HANA*. Stuttgart: Springer Gabler.
- Proff, H., Brand, M., & Schramm, D. (2020). *Altersgerechte Fahrerassistenzsysteme - Technische, psychologische und betriebswirtschaftliche Aspekte*. Universität Duisburg-Essen, Nordrhein-Westfalen, Deutschland: Springer Gabler. Abgerufen am 15. April 2022
- prognos-Forschungsinstitut. (2018). *Einführung von Automatisierungsfunktionen in der Pkw-Flotte*. (A. e. V., Hrsg.) Basel, Schweiz: ADAC e. V. Abgerufen am 17. Juni 2022 von [https://www.adac.de/-/media/pdf/motorwelt/prognos\\_automatisierungsfunktionen.pdf?la=de-de&hash=4FE03D2842A22A8F900AE176AFCA6887](https://www.adac.de/-/media/pdf/motorwelt/prognos_automatisierungsfunktionen.pdf?la=de-de&hash=4FE03D2842A22A8F900AE176AFCA6887)
- Püschel, N. (2. November 2017). *Aktuelle Datensicherungsmethoden im Überblick*. (J. Triebel, Herausgeber) Abgerufen am 16. Mai 2022 von IDG Tech Media GmbH: <https://www.cio.de/a/aktuelle-datensicherungsmethoden-im-ueberblick,3564528>
- Quantum. (2016). Ein anderer Ansatz der Datenspeicherung. *Quantum* (S. 43). München: Quantum.com. Abgerufen am 26. Mai 2022 von [https://www.also.com/ec/cms5/media/documents/1010\\_anbieter/quantum\\_1/qua16032\\_st\\_saag\\_st01285v5\\_cmyk\\_web\\_v07.pdf](https://www.also.com/ec/cms5/media/documents/1010_anbieter/quantum_1/qua16032_st_saag_st01285v5_cmyk_web_v07.pdf)
- Raich, M., & Müller-Seeger, J. (2020). *Symposium Qualitative Forschung 2018 - Verantwortungsvolle Entscheidungen auf Basis qualitativer Daten* (1. Ausg., Bd. I). (J. Müller-Seeger, Hrsg.) Hall in Tirol, Österreich: Springer Gabler. Abgerufen am 9. Mai 2022
- Rößger, P. (2015). *Autonomes Fahren - Wie viel Autonomie das Autofahren verträgt* (Bd. II). Stuttgart, Baden-Württemberg, Deutschland: Springer. Abgerufen am 30. April 2022
- Rowen, C., & Kumar, R. (9. Mai 2016). *elektroniknet.de*. (W. F. GmbH, Herausgeber) Abgerufen am 1. Mai 2022 von Neuronale Netze für die Bilderkennung - Lernen

statt programmieren:

<https://www.elektroniknet.de/automotive/assistenzsysteme/lernen-statt-programmieren.130142.html>

Ruhr-Universität-Bochum. (10. September 2011). *Results*. Abgerufen am 28. Mai 2022 von Ruhr-Universität-Bochum: [https://benchmark.ini.rub.de/gtsrb\\_results.html](https://benchmark.ini.rub.de/gtsrb_results.html)

Sauer, P., Wiesche, M., & Krimmling, J. (2018). *Management digitalier Plattformen - Datengetriebene Geschäftsmodelle für Mobilität in der Smart City* (Bd. I). (H. Krcmar, Hrsg.) Berlin, Berlin, Deutschland: Springer Gabler. Abgerufen am 13. Juni 2022 von <https://link.springer.com/content/pdf/10.1007/978-3-658-21214-8.pdf>

Schäfer, P. (16. Januar 2019). *Infineon und TTTech Auto zeigen neuen Aurix-Mikrocontroller*. Abgerufen am 25. Juni 2022 von Aurix-Mikrocontroller: <https://www.springerprofessional.de/automatisiertes-fahren/automobilelektronik--software/infineon-und-tttech-auto-zeigen-neuen-aurix-mikrocontroller/16398042>

Schaust, F. (2019). *Vorhersage von Gefechts-Ausgängen im Echtzeit-Strategiespiel StarCraft II mittels Convolutional Neural Networks* (1. Ausg., Bd. I). Koblenz, Rheinland-Pfalz, Deutschland: Universität Koblenz - Bachelorarbeit.

Scheffels, G. (4. Februar 2022). *Automobil Industrie*. Abgerufen am 30. April 2022 von Autonomes Fahren: Wie es funktioniert, wo die Technik steht: <https://www.automobil-industrie.vogel.de/autonomes-fahren-wie-es-funktioniert-wo-die-technik-steht-a-786184/>

Schiefkofer, P., Erdogan, Y., Schindler, S., & Wendl, M. (2019). *Maschinelles Lernen für das automatisierte Fahren - Entwicklung automatisiertes Fahren*. DELO - Industrie Klebstoffe. Frankfurt am Main: DELO. Abgerufen am 05. Mai 2022 von <https://link.springer.com/content/pdf/10.1007/s35148-019-0151-7.pdf>

Schmidt, B. (29. März 2019). *BMW Data Driven Development: Mit D3 zum Autonomen Fahren*. Abgerufen am 8. Mai 2022 von Bimmer Today: <https://www.bimmertoday.de/2019/03/28/bmw-data-driven-development-mit-d3-zum-autonomen-fahren/#image-6>

Schrahe, S. (2. September 2021). *Datensammeln im Alltagsverkehr*. (Z. F. AG, Herausgeber) Abgerufen am 8. Mai 2022 von Autonomes Fahren: [https://www.zf.com/mobile/de/stories\\_14848.html](https://www.zf.com/mobile/de/stories_14848.html)

- Schriller, T., Rothmann, K., & Götze, K. (2016). *Autonomes Fahren in Deutschland - wie Kunden überzeugt werden*. Analytic. Düsseldorf: Deloitte - Analytics Institute.
- Science media center - germany. (15. April 2021). *Wie weit ist die Forschung mit dem autonomen Fahren?* . Abgerufen am 25. Juni 2022 von Forschungsstand: <https://www.sciencemediacenter.de/alle-angebote/science-response/details/news/wie-weit-ist-die-forschung-mit-dem-autonomen-fahren/#Frage3>
- Selle, S. (2018). *Künstliche Neuronale Netzwerke und Deep Learning*. Hochschule für Technik und Wirtschaft des Saarlandes, Saarbrücken: wirtschafts wissenschafte htw saar.
- shaip. (19. April 2022). *Qualitativ hochwertige Trainingsdaten treiben leistungsstarke autonome Fahrzeuge an*, 1. Abgerufen am 09. Juni 2022 von Trainingsdaten: <https://de.shaip.com/blog/training-data-for-autonomous-vehicles/>
- Sigurd Schacht, C. L. (2019). *Blockchain und maschinelles Lernen - Wie das maschinelle Lernen und die Distributed-Ledger-Technologie voneinander profitieren* (1. Ausg.). Hochschule Heilbronn, Deutschland: Springer Vieweg.
- SkySystems. (10. Juli 2020). *R wie Rechenzentrum*, 1. Abgerufen am 15. Mai 2022 von Sky Systems: <https://www.skysystems.it/r-wie-rechenzentrum/>
- Statista. (10. Juli 2019). *Definition Bias - Verzerrung* . Abgerufen am 10. Juni 2022 von Bias: [https://de.statista.com/statistik/lexikon/definition/35/bias\\_verzerrung/](https://de.statista.com/statistik/lexikon/definition/35/bias_verzerrung/)
- Süllwold, L.-J. (2020). *Verkehrsschilderkennung mit künstlichen Neuronalen Netzen*. Hannover, Niedersachsen, Deutschland: Universität Hannover. Abgerufen am 30. April 2022
- Technische Universität Eindhoven. (20. März 2012). *Parallel Architecture Research Eindhoven* . Abgerufen am 08. Juni 2022 von CNN: <http://parse.ele.tue.nl/research/projects>
- Tesla. (2021; 2020). *Tesla Vehicle Safety Report*. Autonomes Fahren. USA: Tesla. Abgerufen am 13. Mai 2022 von <https://www.tesla.com/VehicleSafetyReport>
- TIBCO Software Inc. (10. Juli 2020). *Was ist Überwachtes Lernen?* (T. S. Inc., Herausgeber, & TIBCO Software Inc.) Abgerufen am 24. April 2022 von Was ist Überwachtes Lernen?: <https://www.tibco.com/de/reference-center/what-is-supervised-learning>

- Töbel, L., & Kropp, M. (9. März 2018). "Association Rule Learning" - Assoziationsanalyse. *Machine Learning*, 1(1.), 17. Abgerufen am 25. April 2022 von [http://www.sigs.de/public/js/2017/01/kropp\\_toebel\\_JS\\_01\\_17.pdf](http://www.sigs.de/public/js/2017/01/kropp_toebel_JS_01_17.pdf)
- Universität der Bundeswehr München. (10. Februar 2020). *Testgelände autonomes Fahren*. Abgerufen am 20. Juni 2022 von Testgelände: <https://www.unibw.de/tas/ausstattung/testgelaende>
- Volkswagen. (6. Dezember 2019). *Gemeinsam mit IBM und internationalen Hochschulen forscht Volkswagen an der Königsdisziplin des autonomen Fahrens: selbstfahrende Autos im Stadtverkehr*. . (Volkswagen, Herausgeber) Abgerufen am 9. Mai 2022 von <https://www.volkswagenag.com/de/news/stories/2019/12/the-champions-league-of-autonomous-driving.html#>
- Wennker, P. (2020). *Künstliche Intelligenz in der Praxis - Anwendung in Unternehmen und Branchen: KI wettbewerbs- und zukunftsorientiertes einsetzen*. Bochum, Nordrhein-Westfalen, Deutschland: Springer Fachmedien Wiesbaden GmbH.
- Witte, F. (2020). *Strategie, Planung und Organisation von Testprozessen - Basis für erfolgreiche Projektabwicklung im Softwaretest*. Landshut, Bayern, Deutschland: Springer Vieweg. Abgerufen am 14. Juni 2022 von <https://link.springer.com/content/pdf/10.1007/978-3-658-31228-2.pdf>
- Wuttke, L. (11. August 2020). *Was ist Supervised Learning (Überwachtes Lernen)?* Abgerufen am 24. April 2022 von Was ist Supervised Learning (Überwachtes Lernen)?: <https://datasolut.com/wiki/supervised-learning/>
- Wuttke, L. (10. August 2020). *Was ist Unsupervised Learning (Unüberwachtes Lernen)*. (datasolut, Herausgeber) Abgerufen am 25. April 2022 von <https://datasolut.com/wiki/unsupervised-learning/>
- ZF. (14. Februar 2021). *Trainingscampus für autonome Autos*, 1. (ZF, Herausgeber) Abgerufen am 13. Mai 2022 von ZF: [https://www.zf.com/mobile/de/stories\\_14976.html](https://www.zf.com/mobile/de/stories_14976.html)
- Ziegler, L., Krämer, F., & Haustein, N. (2020). *KI und Big-Data-Management für autonomes Fahren* (1. Ausg., Bd. I). Kelsterbach, Hessen, Deutschland: IBM Deutschland. Abgerufen am 16. Mai 2022

## Anhang A:



## Stichwortverzeichnis

### A

Aktivierungsfunktion V, 12, 13, 15  
Aktivierungspotentials  
    Aktivierungspotenzial 12  
Algorithmen  
    Algorithmus 5, 9, 39, 44, 54  
Algorithmus 8, 9  
angewandten Mathematik  
    angewandte Mathematik 18  
Annotation 33  
Assoziationsanalyse 9  
Aufnahme 40  
Aufzeichnungen 38  
Ausgabeschicht 11  
Ausgabewert 6, 8, 19  
autonomen Fahrzeuge 44

### B

**Bestärkendes Lernen** 7  
Bias IX, 19  
Bildererkennung 10, 22

### C

Cloud VI, 29, 34, 35, 36, 42, 43  
**Clusteranalyse** 9  
Convolutional Layer V, 16, 17, 18, 19, 20  
Convolutional Neural Network 1, I, III, VI, VIII, 1,  
    3, 4, 5, 8, 15, 16, 17, 18, 19, 21, 22, 23, 24,  
    25, 29, 32, 37, 38, 39, 40, 41, 42, 43, 44, 46,  
    47, 48, 49, 53

### D

Daten I, III, V, VI, VII, IX, 1, 4, 5, 7, 8, 9, 10, 11,  
    14, 17, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30,  
    32, 33, 34, 35, 36, 37, 38, 39, 40, 42, 43, 46,  
    47, 49, 50, 51, 52, 54, 55  
Datenaffinität 24  
Datenauswahl 38, 39

Datenbank IX, 39  
Datengewinnung III, 30, 31, 33, 34, 39, 42, 46  
Datenkapazität 28  
Datenmengen 4, 9, 26, 34, 37  
Datenprozess I, III, 4, 42, 44, 46, 47  
Datenqualität 24, 47  
Datensammlung 1, III, 1, 4, 24, 35, 42, 46  
Datensätze  
    Datensatz 8  
Datenspeicherung I, III, 34, 35, 36, 39, 42, 43,  
    46, 55  
Datenverarbeitung I, III, 31, 36, 37, 38, 39, 43,  
    46, 47  
Datenwachstum 35, 43  
Deep Learning V, IX, 10, 18, 21, 48, 50, 52, 53,  
    56  
Dense Layer 19  
Diagramm 13  
Dimension 18  
**Dimensionsreduktion** 9  
Diskretisierungsprozesses  
    Diskretisierungsprozess 18  
Down Sampling 18

### E

Eingabedaten 8, 16  
Eingabeparameter 8  
Entwicklungsstufe 21, 22, 44

### F

Fahren 1, I, III, V, VI, 1, 2, 3, 4, 5, 8, 9, 21, 22,  
    24, 26, 27, 28, 29, 30, 31, 32, 33, 35, 38, 39,  
    43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54,  
    55, 56, 57  
Fahrerassistenz 37  
Fahrzeuge VI, 1, 2, 28, 29, 30, 34, 44, 45, 47,  
    49, 56  
Forschungsstand III, 3, 44, 47  
Fully-connected Layer 19  
Fully-Connected Layer 19, 20

Funktion VIII, 10, 13, 14, 15, 44  
Funktionsweise VI, 3, 4, 10, 17, 20, 21, 24, 40

## G

German Traffic Sign Benchmark 39, 43, 52  
Gesamtergebnis 11  
Gewicht 11, 19  
Gradient 15  
Gradientenabstiegsverfahren V, 14, 15  
Grenzfälle 30, 37, 38, 39, 43

## H

HDD VIII, 36, 42  
Hidden Layer 11

## I

individuell 11  
Informationen III, IX, 5, 11, 17, 20, 21, 22, 25,  
26, 27, 28, 32, 47  
Ingenieure 29, 30, 37, 43, 44, 47  
Input Layer 11, 20

## K

**Klassifikation** 8, 16, 19, 22  
Komponente 10, 17, 19  
Komponenten III, 3, 4, 16, 20, 24  
Kostenaufwand 32  
Künstliche Intelligenz V, VIII, IX, 10, 11, 12, 13,  
16, 24, 48, 50, 53, 54, 57  
Künstliche Neuronale Netzwerke III, V, 11, 18,  
44, 52, 56

## L

leistungsfähig 32  
Lernen III, V, IX, 4, 5, 7, 8, 9, 10, 21, 24, 26, 36,  
43, 50, 51, 53, 54, 55, 56, 57  
Lernprozess 6

## M

maschinellen Lernen 6, 10  
Max Pooling 18  
Metainformationen 37, 43

Methode 9, 40, 46, 47  
Modelle 4, 10, 16, 18, 32  
Modellen  
Modell 12

## N

Nachbarneuronen 12  
Netzwerk IX, 10, 15, 16, 40  
Netzwerke  
Netzwerk 10  
Neuronen 10, 11, 16, 19

## O

Objekterkennung 1, 5, 17, 27, 38, 43, 46, 47  
Output Layer 11

## P

Parameter 18, 19  
Patente V, 1, 2  
Performance 15, 27, 39, 40, 41, 47  
Pool 36  
Pooling Layer V, 17, 18, 19, 20  
Prozess 4, 5, 10, 30, 38, 42, 44, 46

## R

Rechenzentrum VI, 4, 28, 29, 30, 34, 35, 52,  
54, 56  
Rechenzentren 4, 29, 43, 46, 47  
Rectified Linear Unit VIII, 13  
**Regression** 8  
ReLu VIII, 14, 15, 44, 49, 52

## S

Schicht 11, 16  
Schneefall 38, 43  
Schwellwertfunktion 13, 44  
Sigmoid Funktion 13  
Simulation 32, 33, 34, 38, 39, 42  
Simulationen VI, 30, 35, 37, 38, 42, 43  
Software-Lösungen 38, 39, 46, 47  
Speichermedien 34, 42  
Starkregen 38, 43



Straßenverkehr III, VI, 2, 22, 23, 27, 30, 33, 38,  
42, 43, 44, 46, 47  
strukturierte 24  
Subsampling 18  
System VIII, 8, 21, 22, 36, 40

## T

Tangens Hyperbolicus 13  
Tape 36  
Technik IX, 2, 3, 4, 44, 45, 47, 55, 56  
Teilnehmer 41  
Teststrecken 30, 31, 33, 42, 43, 53  
Tool 38  
Training I, 15, 24, 29, 36, 37, 38, 39, 40, 42, 44,  
47, 48, 54  
Trefferquote 6, 40, 41, 44, 46, 47  
Trefferquoten 46

## U

**Überwachtes Lernen** 6

Unfallrisiko 9, 46, 47  
unstrukturierte 24, 25, 35, 43  
unstrukturierte Daten 25, 35, 43  
unüberwachte Lernen 8, 9  
**Unüberwachtes Lernen** 6

## V

Validierung 27  
Variante 7, 16, 18, 38, 43  
Varianten 10, 17, 36, 38  
Verkehrsschilder VI, 3, 5, 8, 22, 37, 38, 39, 40,  
41, 43, 46, 48

## W

Wettbewerb 39, 40, 46  
Wissen 5, 26, 27, 50, 51  
wechselbare Speichermedien  
wechselbaren Speichermedien 34

## Erklärung

Hiermit erkläre ich, dass ich die vorliegende Bachelorarbeit selbständig angefertigt habe. Es wurden nur die in der Arbeit ausdrücklich benannten Quellen und Hilfsmittel benutzt. Wörtlich oder sinngemäß übernommenes Gedankengut habe ich als solches kenntlich gemacht. Die vorgelegte Arbeit hat weder in der gegenwärtigen noch in einer anderen Fassung schon einem anderen Fachbereich der Hochschule Ruhr West oder einer anderen wissenschaftlichen Hochschule vorgelegen.

Mülheim, 28.06.2022

Ort, Datum

A. Kaya

Unterschrift