



Innovative Lösungsstrategien für die
Personenidentifizierung im Fahrzeug:
Evaluierung, Auswahl und prototypische
Umsetzung mit vorhandener
Fahrzeugtechnologie.

Maximilian Molzahn

Matrikelnummer: 15888520

Bachelorarbeit Informatik

Prüfer:

Prof. Dr. Lars Wischhof, Hochschule München

Betreuer:

Dr. Thomas Kupka, BMW Group

08.04.2024

Gender-Hinweis

Zur besseren Lesbarkeit wird in dieser Bachelorarbeit das generische Maskulinum verwendet. Die in dieser Arbeit verwendeten Personenbezeichnungen beziehen sich – sofern nicht anders kenntlich gemacht – auf alle Geschlechter.

Erklärung

Maximilian Molzahn, geb. 03.07.2000 (IF7, WS 2023/24)

Matrikelnummer: 15888520

Hiermit erkläre ich, dass ich die Bachelorarbeit selbständig verfasst, noch nicht anderweitig für Prüfungszwecke vorgelegt, keine anderen als die angegebenen Quellen oder Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate als solche gekennzeichnet habe.

München, 08.04.2024

.....

Unterschrift

Abstract

The aim of this bachelor thesis is to deal with the identification of passengers in a vehicle. The following research question is raised: How can passenger identification be implemented taking into account data protection, usability and functionality? In order to answer this question, some basic principles of passenger identification in a vehicle context were presented. A market analysis was first carried out to determine what technologies are already in use across the market. This showed that there are already fundamental approaches to innovative identification technologies, but they relate only to the driver. Three possible identification technologies were then investigated: Identification via a mobile phone app, facial recognition and voice recognition. For each technology, a brainstorming session for a concept design was carried out to assess its advantages and disadvantages. Voice recognition was selected as the identification technology for a prototype implementation in the automotive context. For this purpose, a four-tier architecture was designed and developed, which was used for further investigations in the context of experiments. The results of the experiments were evaluated. Under ideal conditions, the voice recognition worked accurately. However, it turned out that the setup procedure for the speakers was too time consuming and that the results of the voice recognition took too long. This means that voice recognition is not suitable as a technology for identifying passengers without further effort. The research question was answered through the implementation and evaluation. Finally, the results were discussed and an outlook on further research was given.

Zusammenfassung

Das Ziel dieser Bachelorarbeit ist es, sich mit der Insassenidentifizierung in einem Fahrzeug zu beschäftigen. Dazu wird die Forschungsfrage gestellt: Wie kann die Insassenidentifizierung unter Berücksichtigung von Datenschutz, Nutzbarkeit und Funktionalität umgesetzt werden? Um diese Frage zu beantworten, werden zu Beginn einige Grundlagen erläutert, die sich auf die Insassenidentifikation im Fahrzeugkontext beziehen. Zunächst wurde eine Marktanalyse durchgeführt, um zu untersuchen, welche Technologien bereits auf dem Markt verwendet werden. Dabei zeigte sich, dass es bereits grundlegende Ansätze für innovative Identifizierungstechnologien gibt, die sich jedoch nur auf den Fahrer beziehen. Anschließend wurden drei mögliche Identifizierungstechnologien recherchiert: Identifizierung mittels einer Handy-App, Gesichtserkennung und Stimmerkennung. Ein Brainstorming für einen Konzeptentwurf wurde für jede Technologie durchgeführt, um ihre Vor- und Nachteile zu bewerten. Für eine prototypische Implementierung im Fahrzeugkontext wurde die Stimmerkennung als Identifizierungstechnologie ausgewählt. Dazu wurde eine vierstufige Architektur entworfen und entwickelt, die für weitere Untersuchungen im Rahmen von Versuchen verwendet wurde. Die Ergebnisse der Experimente wurden evaluiert. Unter idealen Bedingungen funktionierte die Stimmerkennung präzise. Allerdings stellte sich heraus, dass der Einrichtungsprozess für die Sprecher zu langwierig ist und die Ergebnisse der Stimmerkennung zu viel Zeit in Anspruch nehmen. Somit eignet sich die Stimmerkennung nicht ohne weitere Aufwände als Technologie für die Identifizierung von Insassen. Durch die Implementierung und Auswertung konnte die Forschungsfrage beantwortet werden. Zum Schluss wurde eine Diskussion der Ergebnisse durchgeführt und ein Ausblick auf weitere Forschungsthemen gegeben.

Inhaltsverzeichnis

Abstract	i
Zusammenfassung	ii
Abbildungsverzeichnis	v
Tabellenverzeichnis	vi
Abkürzungsverzeichnis	vii
1 Einleitung	1
1.1 Motivation	1
1.2 Umfang der Arbeit	2
1.3 Problemstellung und Zielsetzung	3
1.4 Forschungsfrage	4
1.5 Aufbau der Arbeit	4
2 Grundlagen	6
2.1 Definitionen und Begriffserklärungen	6
2.1.1 Identifizierung	6
2.1.2 Authentisierung, Authentifizierung und Autorisierung . . .	7
2.1.3 Personenbezogene Daten	7
2.1.4 Fahrerprofil	8
2.1.5 Biometrische Identifizierung	8
2.1.6 Schlüsselbasierte Identifizierung	8
2.1.7 Maschinelles Lernen	9
2.1.8 Containerisierung	9

INHALTSVERZEICHNIS

2.2	Identifizierungstechnologien in BMW-Fahrzeugen	10
2.2.1	Identifizierungsmethoden für den Fahrer	11
2.2.2	Identifizierungsmethoden für die Insassen	11
2.2.3	Bewertung der vorhandenen Identifizierungsmethoden . . .	12
2.2.4	Fazit	13
2.3	Identifizierungstechnologien anderer Automobilhersteller	13
2.3.1	Auswertung der Recherche	14
2.3.2	Bewertung und Vergleich	15
3	Literaturrecherche	17
3.1	Auswahl der Identifizierungstechnologien	17
3.2	Evaluierung der Identifizierungstechnologien	18
3.2.1	Zuverlässigkeit und Genauigkeit	19
3.2.2	Sicherheit und Datenschutz	19
3.2.3	Integration im Fahrzeugkontext	20
3.3	Ergebnisse der Recherche	20
3.3.1	Mobile App & Authentifizierungsportal	21
3.3.2	Gesichtserkennung	24
3.3.3	Stimmerkennung	27
3.3.4	Kombination der Technologien	31
4	Implementierung	32
4.1	Auswahl der Identifizierungstechnologie	32
4.2	Anforderungen	33
4.3	Entwurf der Architektur	33
4.4	Auswahl der Hardware	36
4.5	Softwarekomponenten	37
4.5.1	Frontend	37
4.5.2	Backend	41
4.5.2.1	Zustand des Backends	42
4.5.2.2	Trainingsprozess	44
4.5.2.3	Klassifikationsverfahren	44
4.5.3	GPU-Backend	45

INHALTSVERZEICHNIS

4.5.3.1	Vorbereitung der Trainingsdaten	45
4.5.3.2	Aufbau des Modells	46
4.5.3.3	Trainingsprozess	48
4.5.4	Incar Komponente	49
4.5.4.1	Kalibrierung der Mikrofone	49
4.5.4.2	Aufnahme und Verarbeitung der Audiodaten . .	50
4.6	Zusammensetzung und Schnittstellen	51
5	Evaluierung der Implementierung	55
5.1	Versuchsübersicht und Ergebnisse	55
5.1.1	Aufbau	55
5.1.2	Bestimmung der Parameter	57
5.1.3	Bestimmung der Audiolänge	59
5.1.4	Funktionalität	60
5.1.5	Hinweis zur Validität	62
5.2	Bewertung der Implementierung	62
5.2.1	Datensicherheit	63
5.2.2	Benutzerfreundlichkeit	63
5.2.3	Funktionalität und Genauigkeit	64
5.2.4	Schlussfolgerung	65
6	Diskussion und Ausblick	66
6.1	Akzeptanz der Nutzer	66
6.2	Einsatzgebiete	67
6.3	Zusätzliche Einsatzgebiete	68
6.4	Beantwortung der Forschungsfrage	69
6.5	Fazit	70
6.6	Ausblick und Zukünftige Arbeiten	70
	Anhang 1: Bildschirmaufnahmen des Frontends	72
	Anhang 2: UML Sequenzdiagramm: Identifizierungsprozess	77
	Anhang 3: Modell Architektur	79

INHALTSVERZEICHNIS

Anhang 4: Incar-Komponente Kalibrierung	81
Anhang 5: Verlust- und Genauigkeitsgrafiken	83
Anhang 6: Versuchsaufbau	86
Literatur	88

Abbildungsverzeichnis

4.1	Architektur der Implementierung	35
4.2	Aufbau des ML-Modells	47
4.3	Schritte zum Hinzufügen von Sprechern und zum Trainieren des Modells	52
4.4	Dokumentation der Schnittstellen	53

Tabellenverzeichnis

3.1	Auflistung der ausgewählten Identifizierungstechnologien	18
4.1	Positionskürzel für die Kalibrierung der Mikrofone	50
4.2	Kommunikationsschnittstellen Beschreibung	53
5.1	Übersicht der Modellergebnisse	56
5.2	Versuchsergebnisse: Bestimmung der Parameter	58
5.3	Versuchsergebnisse: Bestimmung der Audiolänge	59
5.4	Übersicht der Funktionsversuche	61
5.5	Versuchsergebnisse: Funktionalität	61

Abkürzungsverzeichnis

5G	5th Generation (wireless communication technology)
ACK	Acknowledgement
AES	A dvanced E ncryption S tandard
API	A pplication P rogramming I nterface
ARM	A dvanced R ISC M achine (Prozessor Architektur)
AWS	A mazons W eb S ervices
BLE	B luetooth L ow E nergy
BMW	B ayerische M otoren W erke
CNN	C onvolutional N eural N etwork
CPU	C entral P rocessing U nit
CUDA	C ompute U nified D evice A rchitecture
DSGVO	Datenschutz-Grundverordnung
EC2	E lastic C ompute C loud
FFT	F ast F ourier T ransform
GiB	Gibibyte
GPU	G raphics P rocessing U nit
HTTP(S)	H ypertext T ransfer P rotocol (S ecure)
ID	Identität
ISO	International Organization for Standardization
IT	I nformation T echnology
JSON	J ava S cript O bject N otation
ML	M achine L earning
ODM	O n- D emand M obility

ABKÜRZUNGSVERZEICHNIS

OEM	O riginal E quipment M anufacturer (Hersteller)
OTA	O ver- T he- A ir
QR-Code	Q uick R esponse C ode
RAM	R andom A ccess M emory
RELU	R ectified L inear U nit
ResNet	R esidual N etwork
REST	R epresentational S tate T ransfer
RFID	R adio- F requency I Dentification
S3	S imple S torage S ervice
SSL	S ecure S ockets L ayer
TAN	Transaktionsnummer
UML	U nified M odeling L anguage
URL	U niform R esource L ocator
UX/UI	U ser E xperience/ U ser I nterface
WLAN	W ireless L ocal A rea N etwork

1 | Einleitung

*„**Personalisierung** ist sinnlos, ohne den Einzelnen zu kennen. Verstehe die Träume, Hoffnungen und Ängste, die deinen Kunden motivieren, treffe ihn dort, wo es darauf ankommt.“*

– Paul Gillin, Social-Media-Experte

Personalisierung ist ein zentrales Thema im Bereich der sozialen Medien, aber wie sieht es mit den Endgeräten aus, die wir täglich nutzen? Kennen diese uns genauso gut wie unsere sozialen Medien? Diese Arbeit beschäftigt sich mit der Personalisierung im Fahrzeug, insbesondere mit der Insassenidentifizierung.

1.1 Motivation

Die Integration von Insassenidentifizierungssystemen in Fahrzeugen eröffnet zahlreiche Anwendungsmöglichkeiten für ein personalisiertes Fahrerlebnis. Dieses ermöglicht die individuellen Bedürfnisse und Vorlieben der Fahrzeuginsassen zu erkennen und entsprechende Anpassungen vorzunehmen.

Ein praktisches Beispiel hierfür ist die gemeinsame Nutzung eines Fahrzeugs durch eine vierköpfige Familie. Angenommen, die Mutter nimmt auf dem Beifahrersitz Platz – die Insassenidentifizierung könnte automatisch die Klimaautomatik auf ihre bevorzugte Temperatur einstellen. Sobald der Sohn das Steuer übernimmt, passt das System die Fahrzeuleistung an, um den Fahrstil eines jungen und

weniger erfahrenen Fahrers zu berücksichtigen. Gleichzeitig kann das Unterhaltungssystem im Fond des Fahrzeugs auf die Interessen der Tochter auf der Rückbank abgestimmt werden. Bei Fahrten, bei denen nur die Eltern anwesend sind, könnte das System sogar automatisch die bevorzugte Musik auswählen, um eine angenehme Atmosphäre zu schaffen.

Ebenso bietet die Insassenidentifizierung einen praktischen Vorteil für häufige Nutzer von Mietfahrzeugen. Steigt ein regelmäßiger Nutzer in ein beliebiges Poolfahrzeug ein, wird sofort sein individuelles Fahrerprofil aktiviert. Dadurch werden persönliche Einstellungen wie Sitzposition, gespeicherte Reiseziele und bevorzugte Klimaeinstellungen automatisch angepasst.

Diese Kurzbeschreibung verdeutlicht, dass die Insassenidentifizierung nicht nur das Fahrerlebnis personalisiert, sondern auch den gesamten Fahrzeuginnenraum intelligent auf die Bedürfnisse der Insassen abstimmt. Dies bezieht sich nicht nur auf den Komfortaspekt, sondern trägt zur Verbesserung von Fahrsicherheit und Effizienz bei. In einer Zeit, in der die Umgestaltung der Mobilität in den Mittelpunkt rückt, könnten die beschriebenen innovativen Ansätze einen wesentlichen Beitrag zur zukünftigen Gestaltung der Fahrzeugnutzung leisten.

1.2 Umfang der Arbeit

Die vorliegende Arbeit befasst sich mit der Identifizierung von Insassen in Fahrzeugen, insbesondere im Kontext der BMW Group. Die Abteilung: On-Demand-Mobilität (ODM) der BMW Group entwickelt kontinuierlich neue Konzepte, um die Zukunft der Mobilität aktiv zu gestalten. Eine wichtige Beobachtung ist der rückläufige Trend des Fahrzeugbesitzes in Großstädten [1], wobei alternative Lösungen wie Carsharing oder Fahrzeugpools zunehmend an Bedeutung gewinnen. [2] Das übergeordnete Ziel besteht darin, BMW optimal auf die Herausforderungen der zukünftigen Mobilität vorzubereiten, indem innovative Konzepte entworfen werden, die unter anderem das Teilen von Fahrzeugen erleichtern.

Im Rahmen dieser Arbeit werden verschiedene Identifizierungsmethoden inner-

und außerhalb der Automobilbranche untersucht. Ein Vergleich mit anderen Automobilherstellern (OEMs) dient der Kontextualisierung und ermöglicht die Ableitung von bewährten Methoden. Die Vorstellung von drei ausgewählten Identifizierungsmethoden sowie die prototypische Umsetzung und Erprobung einer dieser Methoden bilden den Schwerpunkt dieser Bachelorarbeit.

Dabei stellt diese Arbeit einen konzeptionellen Ansatz dar und verfolgt nicht das Ziel eines direkten Produktentwurfs. Die Ergebnisse der Arbeit sollen als Grundlage für weitere Forschung und Entwicklung dienen und die BMW Group bei der Gestaltung zukünftiger Mobilitätskonzepte unterstützen.

1.3 Problemstellung und Zielsetzung

Die Einführung von Identifizierungstechnologien im Fahrzeugkontext ist zwangsläufig mit verschiedenen Herausforderungen verbunden. Ein zentraler Aspekt ist der Datenschutz, da es sich bei der Verarbeitung von personenbezogenen Daten um sensible Informationen handelt. Die Gewährleistung eines sicheren und rechtskonformen Umgangs mit diesen Daten ist unerlässlich.

Die Zuverlässigkeit der Identifizierungsergebnisse ist ein weiteres Problemfeld, das sich insbesondere bei unterschiedlichen Umgebungs- und Nutzungsbedingungen ergeben kann. Die Benutzerfreundlichkeit ist ein entscheidender Faktor, da die Akzeptanz der Identifizierungsmethode weitgehend von ihrer Einfachheit und Effizienz abhängt. Ein übermäßiger Aufwand bei der Einrichtung der Methode kann die Benutzerfreundlichkeit beeinträchtigen und eine verbreitete Verwendung der Methode verhindern.

Wie im Folgenden erörtert wird, ist die Identifizierung von Insassen in Fahrzeugen ein komplexes Thema, das mit verschiedenen Herausforderungen verbunden ist. Ziel dieser Arbeit ist daher die Untersuchung verschiedener Identifizierungsmethoden und die prototypische Umsetzung einer ausgewählten Methode. Es ist nicht das Ziel, ein voll funktionsfähiges System zu entwickeln, sondern einen Proof of Concept (Wirksamkeitsnachweis) zu erstellen, der weitere Forschungen und

Entwicklungen nach sich zieht.

1.4 Forschungsfrage

Mit dem obigen Wissensstand ist es jetzt möglich eine Forschungsfrage zu definieren, welche mit der prototypischen Umsetzung beantwortet werden kann.

Wie kann die Insassenidentifizierung unter Berücksichtigung von Datenschutz, Nutzbarkeit und Funktionalität umgesetzt werden?

Die Forschungsmethodik umfasst die Implementierung eines Prototyps für die Insassenidentifizierung unter Berücksichtigung der derzeit in Fahrzeugen eingesetzten Technologien. Die Auswertung dieser Implementierung erfolgt anhand von Versuchen, bei denen die Funktionalität, Datensicherheit und Benutzerfreundlichkeit der Implementierung bewertet werden. Die Ergebnisse der Versuche werden in einem umfassenden Kontext betrachtet und diskutiert.

1.5 Aufbau der Arbeit

Diese Bachelorarbeit ist zweigeteilt und widmet sich der innovativen Insassenerkennung in Fahrzeugen.

Der theoretische Teil der Arbeit konzentriert sich auf die Literaturrecherche verschiedener Technologien zur Insassenidentifizierung im Fahrzeug. Dabei wird eine Bestandsaufnahme der derzeit auf dem Markt verfügbaren Lösungen durchgeführt. Der Schwerpunkt liegt dabei auf der Bewertung ihrer Vor- und Nachteile sowie ihrer Anwendbarkeit im Fahrzeugkontext. Die umfassende Recherche gibt einen fundierten Überblick über den Stand der Forschung und die vorhandenen Lösungen in diesem Bereich.

Im zweiten Teil der Arbeit geht es um die prototypische Umsetzung einer ausgewählten Methode. Die Auswahl erfolgt auf der Grundlage der im theoretischen

Teil gewonnenen Erkenntnisse. Die Implementierung wird durchgeführt, um die praktische Anwendbarkeit und Funktionalität der ausgewählten Methode zu untersuchen, die im Anschluss anhand von Versuchen evaluiert werden soll.

Die Ergebnisse der theoretischen Forschung und der Implementierung werden in einem ganzheitlichen Kontext betrachtet. Mögliche Implikationen für die weitere Entwicklung und Anwendung der Insassenidentifizierung in Fahrzeugen werden diskutiert. Die Arbeit schließt mit einem Ausblick auf mögliche Ansätze für weitere Forschungsprojekte auf diesem komplexen Gebiet.

2 | Grundlagen

In diesem Kapitel werden die Grundlagen für die Identifizierung von Insassen in Fahrzeugen erklärt. Dabei wird eine aktuelle Marktanalyse durchgeführt um sowohl den Stand der Technologie bei Fahrzeugen der BMW Group zu ermitteln, als auch ein Vergleich zu anderen OEMs zu ziehen.

2.1 Definitionen und Begriffserklärungen

Zunächst werden einige zentrale Begriffe definiert, die im Kontext der Insassen-identifizierung von Bedeutung sind. Diese Begriffserklärungen dienen als Grundlage für das Verständnis der folgenden Kapitel.

2.1.1 Identifizierung

Die Identifizierung einer Person ist ein komplexer Prozess, der verschiedene Aspekte umfasst und insbesondere im Hinblick auf Datenschutz von entscheidender Bedeutung ist.

Identifizierung: „jemanden/etwas meist an bestimmten Merkmalen (wieder) erkennen.“ [3]

Hierbei spielen Begriffe wie Authentisierung, Authentifizierung und Autorisierung eine Schlüsselrolle, welche auch häufig verwechselt werden.

2.1.2 Authentisierung, Authentifizierung und Autorisierung

Im Rahmen einer Authentisierung erbringt eine Person einen Beweis dafür, dass sie ist, wer sie zu sein vorgibt. Im Alltag geschieht dies zum Beispiel durch die Vorlage des Personalausweises. In der IT wird hierfür häufig ein Passwort in Kombination mit einem Benutzernamen genutzt.

Authentifizierung ist der spezifischere Prozess der Überprüfung der Identität eines Subjekts, normalerweise durch die Verwendung von Anmeldeinformationen wie Benutzername und Passwort, Fingerabdruck, RFID-Karte oder andere biometrische Merkmale.

Autorisierung bezieht sich auf den Prozess der Festlegung von Berechtigungen oder Rechten für eine authentifizierte Entität. Nach erfolgreicher Authentifizierung legt die Autorisierung fest, welche Aktionen oder Ressourcen die authentifizierte Entität nutzen kann. [4]

2.1.3 Personenbezogene Daten

[...] "personenbezogene Daten" alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person [...] beziehen; als identifizierbar wird eine natürliche Person angesehen, die direkt oder indirekt, insbesondere mittels Zuordnung zu einer Kennung wie einem Namen, zu einer Kennnummer, zu Standortdaten, zu einer Online-Kennung oder zu einem oder mehreren besonderen Merkmalen, die Ausdruck der physischen, physiologischen, genetischen, psychischen, wirtschaftlichen, kulturellen oder sozialen Identität dieser natürlichen Person sind, identifiziert werden kann; (Art. 4 Nr. 1 DSGVO)

Bei der prototypischen Entwicklung der Insassenidentifizierung ist es von besonderer Bedeutung, personenbezogene Daten durch korrekte Authentisierung zu schützen und sicherzustellen, dass die Autorisierung angemessen erfolgt.

2.1.4 Fahrerprofil

Fahrerprofile sind personalisierte Einstellungen, die auf die individuellen Präferenzen und Bedürfnisse des Fahrers zugeschnitten sind. Diese Einstellungen können verschiedene Aspekte des Fahrzeugs umfassen, wie zum Beispiel die Sitzposition, die Klimaeinstellungen, die Radiosender oder die Navigationseinstellungen. Die Verwendung von Fahrerprofilen ermöglicht es, das Fahrerlebnis zu personalisieren und den Komfort zu erhöhen. [5]

In den folgenden Kapiteln wird erläutert, wie verschiedene OEMs Fahrerprofile einbinden.

2.1.5 Biometrische Identifizierung

Biometrische Identifizierung bezieht sich auf die Verwendung biometrischer Merkmale zur Überprüfung der Identität einer Person. Biometrische Merkmale können physische Eigenschaften wie Fingerabdrücke, Gesichtsmerkmale, Iris- oder Netzhautmuster, Handgeometrie oder Stimme umfassen. Die Verwendung biometrischer Merkmale zur Identifizierung bietet eine genaue und zuverlässige Methode zur Überprüfung der Identität einer Person.

2.1.6 Schlüsselbasierte Identifizierung

Bei der schlüsselbasierten Identifizierung wird die Identität einer Person anhand eines physischen Schlüssels oder einer Karte überprüft. Dieser Ansatz ist in der Automobilindustrie weit verbreitet und ermöglicht eine einfache und effektive Identifizierung des Fahrers. Die Verwendung von Schlüsseln zur Identifizierung ist ein bewährtes Verfahren, das sich durch Zuverlässigkeit und Benutzerfreundlichkeit auszeichnet.

2.1.7 Maschinelles Lernen

Maschinelles Lernen (ML) ist ein Teilgebiet der künstlichen Intelligenz, bei dem Computer anhand von Eingabedaten und einer Reihe von Regeln, die ihnen zur Verfügung gestellt werden, lernen und sich in einer bestimmten Aufgabe verbessern. Um dies zu ermöglichen, werden spezielle Algorithmen, die auf mathematischer Optimierung und Rechenstatistik basieren, in einem komplexen System kombiniert. [6]

Im Zusammenhang mit ML gibt es einige wichtige Begriffe. Grundsätzlich handelt es sich um die Entwicklung von Modellen, die aus Daten lernen, um Vorhersagen oder Entscheidungen zu treffen. Während des Trainingsprozesses werden diese Modelle mit Trainingsdaten trainiert, wobei die Modellparameter angepasst werden, um die Genauigkeit zu verbessern. Während des Trainings werden Metriken wie Verlust und Genauigkeit verwendet, um die Leistung des Modells zu bewerten. Der Verlust misst die Abweichung zwischen der Vorhersage und der gewünschten Ausgabe, während die Genauigkeit den Prozentsatz der korrekten Vorhersagen angibt. Ein weiterer Schlüsselbegriff ist die Klassifikation, bei der das Modell die Daten in vordefinierte Kategorien oder Klassen einteilt. Diese Konzepte bilden die Grundlage für viele Anwendungen des maschinellen Lernens, von der Bilderkennung bis zur Sprachverarbeitung.

Ein häufig verwendetes Werkzeug für die Entwicklung solcher Modelle ist TensorFlow welches von Google entwickelt wurde [7]. Im Rahmen der Implementierung in Kapitel 4 wird ML vertieft erläutert.

2.1.8 Containerisierung

Containerisierung ist eine Methode zur Bereitstellung und Ausführung von Anwendungen in isolierten Umgebungen, sogenannten Containern. Sie sind eine Art von Virtualisierung, die es ermöglicht, Anwendungen und ihre Abhängigkeiten in einer isolierten Umgebung auszuführen. Container bieten eine leichte und effiziente Möglichkeit, Anwendungen unabhängig von der unterliegenden Infrastruktur

zu verpacken, zu verteilen und auszuführen. Ein bekanntes Containerisierungstool ist Docker, mit dem Entwickler Anwendungen in Containern erstellen und verwalten können. [8]

2.2 Identifizierungstechnologien in BMW-Fahrzeugen

BMW-Kunden, die Fahrzeuge mit dem Betriebssystem "Operating System 7" oder einer neueren Version besitzen, haben Zugriff auf die Funktion der Fahrerprofile, die sogenannten BMW IDs. [9] Diese Funktion ermöglicht es dem Fahrer, verschiedene Einstellungen zu synchronisieren, um ein individuelles Fahrerlebnis zu schaffen. Zu den synchronisierbaren Einstellungen gehören:

- Profilbild
- Navigation (letzte Ziele, Heimatadresse, Karteneinstellungen)
- Medien (gespeicherte Radiosender)
- iDrive (Konfiguration des Hauptmenüs, Sprache, Einheiten)
- Sprachassistent (Vorschläge, Aktivierungswort)
- Außenbeleuchtung
- Sitzposition und Klimaeinstellungen
- Datenschutz Einstellungen. [5]

Diese Funktionalität bietet Fahrern die Möglichkeit, das Fahrzeug an die persönlichen Vorlieben anzupassen, selbst wenn mehrere Nutzer das Fahrzeug teilen. Im „Operating System 7“ können bis zu drei Fahrerprofile gespeichert werden. Im neueren „Operating System 8“ kann ein Hauptnutzer und bis zu sechs zusätzliche Mitnutzer ihre individuellen Fahrerprofile speichern. [10]

Trotz der Vorteile dieser Funktion stellt sich die Frage nach der Aktivierung einer BMW ID im Fahrzeug, insbesondere bezüglich der Datensicherheit, da es sich dabei durchaus um personenbezogene Daten handeln kann. Es stellt sich auch

die Frage, ob die BMW ID auf den Fahrer beschränkt ist oder ob auch andere Insassen diese Funktion nutzen können.

2.2.1 Identifizierungsmethoden für den Fahrer

Die Aktivierung von Fahrerprofilen erfolgt durch eine Kopplung mit dem Fahrzeugschlüssel. Sobald die BMW ID mit einem spezifischen Fahrzeugschlüssel verknüpft ist, wird bei Verwendung dieses Schlüssels das entsprechende Fahrerprofil automatisch ausgewählt. Diese Verknüpfung ermöglicht ein personalisiertes Fahrerlebnis, ohne dass der Fahrer manuell eingreifen muss.

Eine weitere Möglichkeit Fahrerprofile zu aktivieren und zu ändern, während eine andere Person das Fahrzeug benutzt, bietet das iDrive-System. Dieser Vorgang erfordert jedoch eine zusätzliche Sicherheitsebene, da das iDrive-System eine PIN-Eingabe für Profiländerungen anfordert. Diese Sicherheitsmaßnahme stellt sicher, dass nur autorisierte Benutzer Fahrerprofile ändern und auf persönliche Einstellungen zugreifen können. [5]

Trotz der sorgfältigen Konzeption dieses Systems gibt es diverse Herausforderungen, insbesondere wenn das Fahrzeug im Rahmen von Carsharing oder in familiärer Umgebung gemeinsam genutzt wird. In diesen Fällen kann die nahtlose Identifizierung und automatische Auswahl des Fahrerprofils beeinträchtigt werden.

2.2.2 Identifizierungsmethoden für die Insassen

Die BMW ID bietet ein personalisiertes Fahrerlebnis für den Fahrer. Allerdings gibt es eine Einschränkung bei der Authentifizierung von weiteren Insassen. Im aktuellen System ist es nicht möglich, dass sich Beifahrer oder weitere Insassen mit ihren individuellen Identitäten authentifizieren.

Die Authentifizierung bleibt primär auf den Fahrer beschränkt, der das Fahrzeug mit seinem gekoppelten Schlüssel entsperrt. Dies bedeutet, dass die personali-

sierten Einstellungen und Präferenzen nur auf den Fahrer zugeschnitten sind, der das Fahrzeug aktiv nutzt. Für Mitfahrer, die in das Fahrzeug einsteigen, besteht gegenwärtig keine Möglichkeit, eigene Profile zu aktivieren oder spezifische Einstellungen vorzunehmen.

Diese Beschränkung kann in bestimmten Szenarien als herausfordernd wahrgenommen werden, besonders wenn mehrere Fahrzeuginsassen unterschiedliche Präferenzen haben. Die fortlaufende Entwicklung der Fahrzeugtechnologie könnte jedoch zukünftig neue Möglichkeiten für erweiterte Authentifizierungsfunktionen auch für Mitfahrer hervorbringen.

2.2.3 Bewertung der vorhandenen Identifizierungsmethoden

Das Konzept der Fahrerprofile in BMW-Fahrzeugen mit einem bekannten Fahrerkreis und deren Verknüpfung mit dem Fahrzeugschlüssel ist eine und effektive Lösung. Die Verknüpfung mit einem physischen Schlüssel sorgt für eine automatische Authentifizierung des Fahrers und damit für ein individuelles Fahrerlebnis. Dies ist besonders vorteilhaft für Familien oder begrenzte Gruppen von Nutzern, die regelmäßig dasselbe Fahrzeug verwenden.

Jedoch ergibt sich eine gewisse Einschränkung dieses Konzepts, wenn das Fahrzeug ausgeliehen wird, insbesondere bei Carsharing-Modellen, bei denen verschiedene Fahrer unterschiedliche Fahrzeuge nutzen wollen. Das vorliegende Konzept ist in solchen Szenarien weniger flexibel und kann zu Komplikationen führen.

Die nicht gegebene Möglichkeit für weitere Fahrzeuginsassen, sich zu authentifizieren, ist eine klare Limitierung. Bei der Nutzung durch verschiedene Personen bleibt das System primär auf den individuellen Fahrer beschränkt, der in das Fahrzeug mit seinem persönlichen Schlüssel einsteigt.

2.2.4 Fazit

Insgesamt zeigt sich, dass die BMW ID mit etablierten Fahrern und Schlüsselkopplung für bestimmte Nutzungsszenarien effektiv ist. Sie bietet ein personalisiertes und komfortables Fahrerlebnis. Allerdings können alternative Lösungen oder Anpassungen des Konzepts für die Insassenidentifizierung in Sonderfälle wie Carsharing erforderlich sein, wo Flexibilität gefragt ist. Die Möglichkeit, auch Mitfahrer zu authentifizieren und ihre individuellen Einstellungen zu berücksichtigen, könnte das Fahrerlebnis weiter verbessern und künftige Anforderungen an die Fahrzeugpersonalisierung erfüllen.

2.3 Identifizierungstechnologien anderer Automobilhersteller

Die Automobilindustrie hat in den letzten Jahren rasante Fortschritte bei innovativen Technologien gemacht. Diese haben das Ziel das Fahrerlebnis weiter zu verbessern. Neue Technologien spielen eine entscheidende Rolle bei der personalisierten Fahrzeugnutzung und tragen zu mehr Komfort und Sicherheit bei. Dieses Kapitel wirft einen Blick auf verschiedene Identifizierungstechnologien in Fahrzeugen anderer Hersteller und konzentriert sich dabei auf den Vergleich zwischen den OEMs und die zukünftige Bedeutung der Insassenidentifizierung.

Die Automobilindustrie ist von Natur aus wettbewerbsorientiert, und die Rivalität zwischen den OEMs hat zu ständigen Innovationen geführt. Ein wichtiger Einflussfaktor in der Automobilindustrie ist der wachsende Einfluss der asiatischen OEMs. Unternehmen aus Asien, insbesondere aus China, sind zu wichtigen Akteuren auf dem Weltmarkt geworden. Ihr Engagement für innovative Fahrzeugtechnologien hat die Dynamik der Branche erheblich beeinflusst und trägt dazu bei, neue Standards für fortschrittliche Fahrzeugtechnologien zu setzen. [11]

2.3.1 Auswertung der Recherche

Die Konkurrenz zwischen Automobilherstellern zwingt Entwickler neue Technologien voranzubringen, um an der Marktspitze mithalten zu können. In diesem Zusammenhang ist es von großem Interesse, die Ansätze verschiedener OEMs zu untersuchen und zu vergleichen. Diese Analyse verdeutlicht die Betrachtung von Insassenidentifizierungstechnologien und die wachsende Bedeutung dieses Bereichs für die Zukunft der Automobilbranche. Sie unterstreicht die Notwendigkeit weiterer Forschung zur Fahrzeugpersonalisierung.

Tesla, ein Pionier der Elektromobilität, bietet eine ähnliche Funktionalität wie BMW durch die Verknüpfung von Fahrerprofilen über einen physischen Schlüssel, in diesem Fall eine Schlüsselkarte. Allerdings bleibt die Identifizierung auf den Fahrer beschränkt, ohne zusätzliche Insassenerkennung. [12]

Volkswagen und dessen Tochterunternehmen setzten ebenfalls auf Fahrerprofile, um bestimmte Funktionen zu synchronisieren. Unter anderem bietet die Volkswagen ID Fahrern die Möglichkeit ihre Profile zu speichern und zu synchronisieren [13]. Die Tochtergesellschaft **Audi** hat mit verschiedenen Konzepten mittels einem Fingerabdruck basierten Schlüsselsystem experimentiert. “[...] setzt Audi künftig als erster Hersteller auf den Fingerabdruck”. Dieser Artikel aus der Zeitung “Die Welt am Sonntag” [14] aus dem Jahr 2001 zeigt, dass Audi schon damals mit dem Audi A8 ein neues Schlüsselsystem eingeführt hat, das auf der Erkennung von Fingerabdrücken basiert. Dieses System ermöglicht es dem Fahrer, das Fahrzeug ohne physischen Schlüssel zu entriegeln und zu starten. Diese Technologie wird allerdings in aktuellen Modellen nicht mehr verwendet und ist auf den offiziellen Audi-Webseiten nicht mehr zu finden. **Mercedes-Benz** bietet in einigen neueren Fahrzeugmodellen ebenfalls einen Fingerabdrucksensor an, mit dem Funktionen des Infotainmentsystems aktiviert werden können. Mit dieser Funktion kann das Fahrzeug jedoch nicht gestartet werden [15]. **Hyundai** hat ebenfalls in bestimmten Modellen mit Fingerabdrucksensoren experimentiert, um den Fahrer zu identifizieren. [16]

Die Luxusautomarke **Genesis**, die zum Hyundai Konzern gehört, setzt Gesichts-

erkennung ein, um das Fahrzeug von außen zu entriegeln, ohne dass ein Schlüssel benötigt wird. Die Kamera ist auf der Fahrerseite angebracht. Zusätzlich befindet sich im Fahrzeuginneren ein Fingerabdrucksensor, mit dem Fahrerprofile ausgewählt werden können. Ein Demonstrationsvideo zeigt den Einbau und die Funktionen der Gesichtserkennung. [17] **Subaru** setzt ebenfalls auf ein Gesichtserkennungssystem, mit welcher Fahrerprofile freigeschaltet werden können. Gleichzeitig werden auch Sicherheitssysteme wie Aufmerksamkeitswarner verwendet. Diese Funktionalität ist ebenfalls nur auf den Fahrer begrenzt. [18]

Die Tech-Giganten **Apple** und **Amazon** betreten die Automobilbranche mit innovativen Ansätzen zur Insassenidentifizierung. Apple patentierte "Comfort Profiles", das auf die Erkennung von Personen im Fahrzeug abzielt, auch mittels kamera-basierter Technologien. Die Absicht des Unternehmens, Fahrerprofile für das autonome Fahren zu schaffen, verdeutlicht die Beteiligung der Identifizierung an der künftigen Entwicklung autonomer Fahrzeuge. [19]

Auch **Amazon** verfolgt einen innovativen Ansatz mit einem Patent zur Erstellung von Fahrgastprofilen, das verschiedene Sensoren, einschließlich der Fingerabdruckerkennung, integriert. Dies unterstreicht das Engagement der Tech-Giganten, die traditionelle Automobilindustrie mit fortschrittlichen Technologien zu verändern. [20]

Insgesamt verdeutlichen diese unterschiedlichen Ansätze den intensiven Wettbewerb und die Vielfalt bei der Integration von Identifizierungstechnologien zwischen verschiedenen traditionellen Automobilherstellern und Tech-Giganten.

2.3.2 Bewertung und Vergleich

Die Evaluierung der verschiedenen Insassenidentifizierungstechnologien in der Automobilbranche verdeutlicht eine Vielzahl von Ansätze, um die Fahrzeugpersonalisierung zu verbessern. Die Verwendung unterschiedlicher Technologien bietet sowohl Chancen als auch Herausforderungen, die es zu berücksichtigen gilt.

Eine gemeinsame Charakteristik, die sich durch die untersuchten Insassenidentifi-

zierungstechnologien zieht, ist die Fokussierung auf den Fahrer als Hauptnutzer. Sowohl die schlüsselbasierte Identifizierung als auch biometrische Technologien wie Fingerabdrucksensoren und Gesichtserkennung konzentrieren sich primär auf die Authentifizierung des Fahrers. Diese Ausrichtung spiegelt die aktuelle Praxis wider, bei der die personalisierte Erfahrung hauptsächlich auf die Bedürfnisse und Präferenzen des Fahrers basiert.

- Die **schlüsselbasierte Identifizierung** stellt eine bewährte und effektive Methode dar, welche einfach und effektiv ist. Die Identifizierung ist auf den Fahrer beschränkt und kann eine Herausforderung darstellen, wenn das Fahrzeug mit mehreren Personen geteilt wird.
- **Biometrische Technologien** wie Fingerabdrucksensoren und Gesichtserkennung repräsentieren eine fortschrittlichere und vielseitigere Methode der Insassenidentifizierung. Im Gegensatz zur schlüsselbasierten Identifizierung können biometrische Technologien auf mehrere Insassen erweitert werden, was eine umfassendere Personalisierung ermöglicht. Trotz ihrer Effektivität und Präzision sind jedoch technische Herausforderungen zu überwinden. Unterschiedliche Lichtverhältnisse können die Leistung beeinträchtigen, und die nahtlose Integration solcher Technologien in Fahrzeuge erfordert sorgfältige Planung und Umsetzung.

Die aktuelle Ausrichtung auf den Fahrer als zentrale Person zur Identifizierung verdeutlicht die Notwendigkeit, zukünftige Entwicklungen in der Insassenidentifizierungstechnologie dahingehend zu erforschen, damit auch die Bedürfnisse und Präferenzen der weiteren Insassen effektiver berücksichtigt werden können. Eine Weiterentwicklung dieser Technologien könnte dazu beitragen, die Personalisierung im Fahrzeug auf ein neues Niveau zu heben.

3 | Literaturrecherche

Bei der Implementierung von Insassenidentifizierungstechnologien im Fahrzeugkontext sind nicht nur Zuverlässigkeit und Präzision entscheidend, sondern auch Aspekte der Sicherheit und Datenschutz spielen eine zentrale Rolle. Dieses Kapitel widmet sich der Forschungsmethodik, die darauf abzielt, die Anwendbarkeit von Identifizierungstechnologien im Fahrzeugumfeld zu evaluieren.

3.1 Auswahl der Identifizierungstechnologien

Die grundlegenden Erkenntnisse aus vorangegangenen Untersuchungen über Technologien im aktuellen Fahrzeugkontext wie schlüsselbasierte Identifizierung und biometrische Lösungen bilden die Basis für die hier vorgestellten Identifizierungstechnologien. Die Bewertung dieser Technologien ermöglicht es, den Einsatz im Fahrzeugbereich detailliert zu untersuchen und sowohl ihre Vorzüge als auch potenzielle Herausforderungen zu identifizieren.

Die ausgewählten Identifizierungstechnologien haben ihre Ursprünge in verschiedenen Branchen außerhalb des Automobilsektors. Die Querverbindung und Anwendbarkeit dieser Technologien in unterschiedlichen Kontexten, sei es in der Telekommunikationsbranche oder bei Banken, unterstreicht ihre Vielseitigkeit und Relevanz, um auch das mögliche Potenzial dieser Technologien im Fahrzeugbereich vollständig zu erfassen. Diese interdisziplinäre Perspektive ermöglicht es, bewährte Verfahren und Innovationen aus anderen Kontexten zu integrieren und somit einen umfassenden Blick auf die Möglichkeiten der Identifizierungstechno-

logien im Automobilsektor zu werfen.

Tabelle 3.1: Auflistung der ausgewählten Identifizierungstechnologien

Technologie	Kontext	Einsatz als Identifizierungstechnologie
Mobile App	Gerätebasiert	Mehrfaktor Authentifizierung, Online-Banking TAN Verfahren
Gesichtserkennung	Biometrie	Einsatz in Smartphones und Tablets zur Entsperrung
Stimmerkennung	Biometrie	Einsatz in intelligente Assistenten und Kundenservice Hotlines

Die Tabelle 3.1 zeigt die ausgewählten Identifizierungstechnologien und ihre Einsatzbereiche. Die Technologien wurden aufgrund ihrer Relevanz und ihres Potenzials für den Einsatz im Fahrzeugbereich ausgewählt. Die Bewertung dieser Technologien ermöglicht es, den Einsatz im Fahrzeugbereich detailliert zu untersuchen und sowohl ihre Vorzüge als auch potenzielle Herausforderungen zu identifizieren.

3.2 Evaluierung der Identifizierungstechnologien

Im Folgenden wird beschrieben, wie die Identifizierungstechnologien im Fahrzeugkontext evaluiert werden. Die Evaluierung erfolgt anhand der Kriterien Zuverlässigkeit und Genauigkeit, Sicherheit und Datenschutz sowie Integration im Fahrzeugkontext.

Zunächst wird die Technologie in einer möglichen Umsetzung im Fahrzeugkontext vorgestellt. Dabei werden mögliche Ansätze erdacht und konzipiert. Ziel der Konzeptionierung ist es, über eine mögliche Umsetzung zu diskutieren, um daraus weitere Kriterien für die Evaluation abzuleiten. Es soll verdeutlicht werden, dass es sich hierbei um ein Brainstorming handelt, welches nicht zwangsläufig die beste Lösung für die Umsetzung darstellt.

3.2.1 Zuverlässigkeit und Genauigkeit

Die methodische Evaluierung der Zuverlässigkeit von Identifizierungstechnologien im Fahrzeugkontext beruht auf mehreren Ansätzen. Primär erfolgt eine Bewertung auf Basis der Branchen, in denen die betreffende Technologie bereits erfolgreich eingesetzt wird. Hierbei werden Studien und Erfahrungsberichte aus diesen Branchen herangezogen, um Erkenntnisse über die praktische Anwendbarkeit und Zuverlässigkeit zu gewinnen. Die Untersuchung von Studien ermöglicht eine breite Diversität an Anwendungsszenarien und Nutzerprofilen, wodurch die Technologie in realen Fahrbedingungen auf ihre Verlässlichkeit hin bewertet wird.

Die Identifizierung von potenziellen Schwachstellen kann durch die Analyse wissenschaftlicher und sicherheitstechnischer Konferenzen erfolgen. Diese Quellen bieten Einblicke in aktuelle Forschungsergebnisse, Sicherheitsanalysen und mögliche Schwachstellen von Identifizierungstechnologien.

3.2.2 Sicherheit und Datenschutz

Die Evaluierung der Sicherheit und des Datenschutzes von Identifizierungstechnologien im Fahrzeugkontext erfolgt anhand einer systematischen Methodik, die verschiedene Aspekte berücksichtigt. Die Sicherheit der Identifizierungstechnologie wird auf Grundlage einer Konzeptionierung einer möglichen Architektur für den Fahrzeugkontext bewertet. Es wird geprüft, ob sensible Daten ausschließlich im Fahrzeug gespeichert und verarbeitet werden können. In bestimmten Anwendungsfällen kann jedoch ein Datenaustausch mit einem Backend erforderlich sein, wie zum Beispiel beim Carsharing, bei dem ein Fahrzeug mehreren Nutzern zugeordnet werden kann. Die Einhaltung von Datenschutzvorgaben wird gewährt, um klare Vorgaben für die Verarbeitung, Speicherung und Übertragung von Identifizierungsdaten sicherzustellen.

3.2.3 Integration im Fahrzeugkontext

Die Integration von Identifizierungstechnologien im Fahrzeugkontext erfordert eine andere Herangehensweise als in andere Branchen. Unter Berücksichtigung von verschiedenen Faktoren muss eine mögliche Implementierung der Identifizierungstechnologie die ISO 15005 erfüllen, eine Norm, die sich mit ergonomischen Aspekten von Informations- und Steuersystemen im Fahrzeug befasst. Unter anderem wird in dieser Norm gefordert, dass Informationssysteme den Fahrer während der Fahrt nicht ablenken dürfen. Daher muss sichergestellt werden, dass der Fahrer wenig oder gar nicht mit dem System interagieren muss, um während der Fahrt erkannt zu werden.

Um eine effiziente und verzögerungsfreie Identifizierung zu gewährleisten, ist der Rechenaufwand für die Identifizierung eines Insassen relevant. Die Implementierung leistungsfähiger Hardwarekomponenten kann im Fahrzeugkontext aufgrund des hohen Energiebedarfs, der Größe und der Kostenfaktoren eher unrentabel sein, sofern sie nicht bereits zum Beispiel für autonome Fahrsysteme integriert sind.

Auch die Nutzertauglichkeit der Identifizierungstechnologie wird sorgfältig geprüft. Lösungen, die eine unangenehme oder langwierige Interaktion erfordern, könnten die Akzeptanz und Benutzerfreundlichkeit beeinträchtigen. Es ist daher wichtig, Technologien zu wählen, die eine bequeme und benutzerfreundliche Identifizierung gewährleisten, ohne Fahrer oder Fahrgäste in unerwünschte Situationen zu bringen.

3.3 Ergebnisse der Recherche

Im Folgenden werden die Ausgewählten Identifizierungstechnologien vorgestellt und untersucht. Ziel ist es, ein Verständnis für die vielfältigen Ansätze zur Insassenidentifizierung zu schaffen und diese zu evaluieren.

3.3.1 Mobile App & Authentifizierungsportal

Mobile Apps haben in den letzten Jahren eine eindrucksvolle Entwicklung durchlaufen und sind zu einem festen Bestandteil des digitalen Alltags geworden. Allein in Deutschland lag der Anteil der Smartphone-Nutzer im Jahr 2022 bei 81,1 % und wird Prognosen zufolge bis 2027 auf 86,3 % steigen. [21] Insbesondere die Verknüpfung von mobilen Apps mit dem Unterhaltungssektor hat zu innovativen Nutzererlebnissen geführt. Dieses Potential könnte auch für die Nutzung der Insassenidentifizierung von großer Bedeutung sein.

Grundlagen

Im Kontext der Insassenidentifizierung eröffnen mobile Apps eine innovative Perspektive, gestützt durch ihre vielfältigen Schnittstellenfunktionen. Die heutigen Mobilgeräte sind mit leistungsfähiger 5G-Technologie ausgestattet, die eine wichtige Rolle bei der Anbindung an das Backend spielt. Darüber hinaus ermöglichen weitere Kommunikationstechnologien wie Bluetooth oder WLAN die direkte Vernetzung mit dem Fahrzeug in unmittelbarer Nähe.

Im BMW-Kontext ist die Headunit das Hauptbedienelement für das Infotainmentsystem. Sie ist mit einem eingebetteten System, Display und Bedienelementen ausgestattet. Die Kommunikation mit anderen Steuergeräten im Fahrzeug erfolgt über diverse Bussysteme. Eine Verbindung zu externen Geräten wie Smartphones wird über Bluetooth oder WLAN hergestellt und die Anbindung an BMW-Backend-Server erfolgt über das Mobilfunknetz.

BMW stellt seinen Kunden bereits heute bewährte Apps wie die MyBMW [22] oder MINI App [23] zur Verfügung, mit denen Nutzer Einblicke in den Zustand ihres Fahrzeugs sowie der Fahrdaten erhalten können. Die Nutzung dieser etablierten Apps als Plattform für die Insassenerkennung könnte eine überzeugende Lösung sein, da die Nutzer bereits mit der Benutzeroberfläche und der Umgebung vertraut sind.

Durch die Einbeziehung von Smartphones als Identifizierungsmittel entfällt die Notwendigkeit spezifisch verbauter Hardware. Moderne BMW-Fahrzeugmodelle

sind bereits mit den notwendigen Komponenten über die Headunit ausgestattet, um eine Verbindung zu Mobilgeräten herzustellen.

Ein weiterer Vorteil ist die Adaptivität der mobilen Apps, die durch regelmäßige Updates neue Funktionen und Verbesserungen nahtlos integrieren können. Diese kontinuierliche Optimierung ermöglicht eine flexible Anpassung an veränderte Anforderungen und technologische Fortschritte im Bereich der Insassenidentifizierung.

Konzeptionierung

Für die Umsetzung der Insassenidentifizierung mit Smartphones bieten sich verschiedene Ansätze. Wesentlich ist die Nutzung der bestehenden BMW ID, die jedoch so erweitert werden sollte, dass sie nicht nur fahrerspezifisch ist, sondern auch für andere Insassen des Fahrzeugs genutzt werden kann. Diese Erweiterung ermöglicht einen umfänglichen Datensatz über mögliche Insassen und schafft die Grundlage für personalisierte Einstellungen.

Ein möglicher Ansatz basiert auf einem effizienten Kommunikationsmechanismus mit Bluetooth Low Energy (BLE). Sobald die Insassen ihre Position im Fahrzeug einnehmen, wählen sie ihre individuelle Sitzposition über die Smartphone-App aus, bei der die Insassen bereits mit ihrer BMW ID angemeldet sind. Über die BLE-Schnittstelle erfährt das Fahrzeug, welcher Insasse sich auf welchem Platz befindet. Die Verbindung über BLE sorgt für eine energiesparende und zuverlässige Übertragung der Identifizierungsdaten zwischen der App und dem Fahrzeug und gewährleistet eine nahtlose Integration in das bestehende System.

Eine zweite Variante setzt auf die Identifizierung mit einem QR-Code, was für die Insassen eine intuitive Methode ist. Beim Einsteigen in das Fahrzeug wird ein Startbildschirm auf der Headunit angezeigt, der die Insassen auffordert, sich mit ihrem Konto anzumelden. Durch das Scannen des Codes öffnet sich ein Anmeldeportal, in dem sich die Insassen mit ihrer individuellen BMW ID anmelden können. Jeder Insasse kann dann seine Sitzposition auswählen. Die eingegebenen Informationen werden über ein Backend an das Fahrzeug übermittelt, so dass das Fahrzeug darüber informiert ist, welche Insassen auf welcher Position im Fahrzeug

sitzen. Dieser Ansatz ist vorteilhaft für Nutzer, die nur ungern eine spezifische Anwendung herunterladen, ist aber auch wesentlich vom Mobilfunknetz abhängig und daher nicht immer in der Lage, eine Verbindung herzustellen, beispielsweise in Tiefgaragen.

Eine aussichtsreiche Strategie könnte darin bestehen, beide Ansätze zu kombinieren.

Vorteile

Die Notwendigkeit, zusätzliche Hardware-Komponenten in das Fahrzeug einzubauen, wird durch die Implementierung des Smartphone-Ansatzes eliminiert. Dadurch entsteht eine kosteneffiziente Lösung, die den Umfang der Hardware im Fahrzeug minimiert und gleichzeitig eine zuverlässige Identifizierung der Insassen gewährleistet. Darüber hinaus ermöglichen Over-the-Air (OTA) Updates die Aktualisierung und Erweiterung der Identifizierungsfunktionalitäten. Diese Funktion ermöglicht die Aufrüstung der Insassenidentifizierung in bestehenden Fahrzeugen, ohne dass das Fahrzeug physisch verändert werden muss. Dies erhöht die Anpassungsfähigkeit des Systems erheblich und trägt zu langfristiger Effizienz und Modernität bei.

Durch die Integration in die App ist es auch möglich, sitzplatzabhängige Infotainment-Einstellungen über die App vorzunehmen. So können zum Beispiel Fahrzeitinformationen, Musiksteuerung und Klimaeinstellungen bequem und persönlich über das Mobilgerät gesteuert werden.

Auch im Hinblick auf den Datenschutz ist dieser Ansatz vielversprechend, da nur begrenzt Daten für die BMW ID benötigt werden. Im Vergleich zu anderen Identifizierungsmethoden, bei denen möglicherweise sensible persönliche Informationen, gesammelt werden können, beispielsweise Biometrie Daten, beschränkt sich die Identifizierung mittels Smartphone auf relevante und anonymisierte Daten.

Herausforderungen

Die oben genannten Vorteile beschreiben nicht das gesamte Konstrukt, da der komplexe Ansatz eine Reihe von Herausforderungen mit sich bringt.

Eine zentrale Herausforderung besteht darin, die Position der Insassen zu erfassen, ohne dass diese selbst aktiv eingreifen müssen. Die Bereitschaft der Nutzer, diesen Zwischenschritt zu machen, wirft Fragen zur Nutzerakzeptanz und Freiwilligkeit auf. Es ist zu klären, ob die Insassen bereit sind, ihre Position aktiv anzugeben, um eine Identifizierung der Insassen zu ermöglichen.

Die Identifizierung wird durch die Dynamik während der Fahrt beeinflusst. Kurze Ausstiege, Sitzwechsel oder das Hinzukommen neuer Insassen können die Genauigkeit der Erkennung beeinträchtigen. Eine mögliche Lösung besteht darin, die Erkennung zu aktivieren, wenn sich ein Insasse in das Fahrzeug setzt, mittels der Sensortechnologie, die bereits für Sicherheitsgurtwarnungen verwendet wird.

Die Abhängigkeit vom Mobiltelefon wirft ebenfalls Fragen auf, insbesondere wenn die Insassen in das Fahrzeug einsteigen, ihre mobilen Geräte aber ausgeschaltet sind oder keine ausreichende Batterieleistung haben. Außerdem können Kinder oder andere Personen, die möglicherweise kein eigenes Mobilgerät besitzen, von dieser Identifizierungsmethode ausgeschlossen werden.

Diese Szenarien erfordern alternative Lösungen, um eine zuverlässige Identifizierung der Insassen zu gewährleisten, unabhängig von der Verfügbarkeit und Funktionalität der mobilen Geräte.

3.3.2 Gesichtserkennung

Im Rahmen der Marktanalyse in Kapitel 3.2 wird deutlich, dass die Gesichtserkennung bereits erfolgreich bei Subaru und Genesis, eingesetzt wird – sowohl im Bereich der Sicherheit als auch zur Fahreridentifizierung. Angesichts dieser Anwendung stellt sich die Frage, inwiefern die Gesichtserkennungstechnologie auf die weiteren Insassen erweitert werden kann.

Grundlagen

Die Gesichtserkennungstechnologie hat in den letzten Jahren erhebliche Fortschritte gemacht und wird zunehmend in verschiedenen Bereichen eingesetzt. Durch den Einsatz von ML und diverser Sensortechnologien wird unterstrichen,

dass sehr präzise Ergebnisse erzielt werden können. [24] Bekannte Beispiele für die Integration dieser Technologie in den Alltag sind die Face ID Funktionen in Mobilgeräten von Apple wie dem iPhone, die die Zuverlässigkeit und Präzision der Gesichtserkennungstechnologie demonstrieren. [25] Diese Technologie ermöglicht es iPhone-Nutzern, ihr Gerät zu entsperren, automatisch Passwörter einzugeben und sogar Zahlungen vorzunehmen.

Die Implementierung der kamerabasierten Gesichtserkennung hat sich, bezüglich des Rechenaufwandes als sehr effizient erwiesen. *“Mit einer single core Intel CPU mit 2,2 GHz [...] läuft DeepFace mit 0,33 Sekunden pro Bild, wenn man die Bilddenkodierung, [...] und die endgültige Klassifizierungsausgabe berücksichtigt.”* Das oben erwähnte DeepFace ist ein von Facebook entwickelter Gesichtserkennungsalgorithmus, der Gesichter mit einer Genauigkeit von 97,35 % ($\pm 0,25$ %) erkennen kann. Das ist fast so genau wie die menschliche Gesichtserkennung, die dem Paper zufolge bei über 97,5 % liegt. [26]

Der Einsatz von ML beweist, dass die Erkennung von Gesichtern auf der Grundlage verschiedener Merkmale die Genauigkeit der Identifizierung verbessert. Dies kann auch angewendet werden, wenn mehrere Personen auf dem Bild zu sehen sind.

Konzeptionierung

Bei der Erkennung der Insassen mittels Gesichtserkennungstechnologie im Fahrzeug wird eine zentrale Kamera in den Dachhimmel des Fahrzeugs integriert. Je nach verwendetem Algorithmus können auch mehrere Sensoren im Fahrzeug angebracht werden, um eine eindeutige Erkennung aller Sitze zu ermöglichen. Diese Technologie ermöglicht die Identifizierung von Insassen, unabhängig davon, ob eine bestehende BMW ID vorhanden ist oder nicht, da die Klassifizierung einer bereits erfassten Person nicht unbedingt mit einem Kundenkonto verknüpft sein muss.

Es ist erwähnenswert, dass BMW-Fahrzeuge, die mit autonomem Fahren ausgestattet sind, bereits über eine Gesichtserkennungstechnologie verfügen, die unter anderem dazu dient, die Aufmerksamkeit des Fahrers beim autonomen Fahren zu

erkennen. Sie wird auch eingesetzt, um die Augenposition des Fahrers zu erfassen und einen Spurwechsel einzuleiten. [27] Diese Technologie könnte möglicherweise in eine zentrale Kamera integriert werden, um verschiedene Anwendungen in einem System zu kombinieren.

Durch die Verwendung von vordefinierten Segmenten im Bild kann die genaue Positionierung der Insassen und die Bestimmung der Sitzposition ermöglicht werden. Wichtig ist, dass die zentrale Kamera jede beliebige Position im Fahrzeug erfassen kann, wobei es bei verschiedenen Fahrzeugbauarten zu Variationen kommen kann.

Vorteile

Die Gesichtserkennungstechnologie hat gegenüber der Verwendung einer mobilen App eine Reihe von Vorteilen.

Die Insassen müssen keine manuellen Eingriffe vornehmen, sie müssen sich lediglich in das Fahrzeug setzen. Ebenso ist für die Gesichtserkennung keine Verknüpfung mit einem zuvor angelegten Insassenkonto wie der BMW ID erforderlich. Somit wird eine bedienerfreundliche Erfahrung ermöglicht, da die Insassen nach erfolgreicher Einrichtung ohne zusätzlichen Aufwand identifiziert werden können. Dieser Ansatz ist jedoch nur in einem Familienkontext mit einer begrenzten Anzahl an möglichen Insassen von Vorteil, da ansonsten die Fahrerprofile aus einem Backend, zum Beispiel im Fall von Carsharing, übernommen werden müssen.

Ein weiterer Vorzug ist die Einfachheit und Schnelligkeit des Einrichtungsprozesses. Ein anschauliches Beispiel für die Effizienz ist die Einrichtung von Face ID auf einem iPhone, die in der Regel weniger als zwei Minuten benötigt.

Dieses unkomplizierte und benutzerfreundliche Konzept trägt erheblich zur Attraktivität der Gesichtserkennungstechnologie für die Identifizierung von Fahrzeuginsassen bei.

Herausforderungen

Es gibt jedoch viele Herausforderungen und offene Fragen bei der Gestaltung dieser Technologie im Fahrzeugkontext. Ein wichtiger Aspekt ist der Datenschutz,

denn biometrische Daten gehören zu den besonderen Kategorien nach DSGVO Art. 9 Abs. 1. Die Verordnung schreibt vor, dass diese Daten prinzipiell nicht verarbeitet werden dürfen. Dies wirft Bedenken hinsichtlich der Sicherheit und des Umgangs mit diesen Daten auf, da die Verarbeitung dieser Daten nur unter Rechtmäßigkeitsvoraussetzungen möglich ist.

Ein weiterer Aspekt sind die Lichtverhältnisse im Fahrzeug, bei denen herkömmliche Bilderkennungssysteme oft Schwierigkeiten haben. Moderne Sensortechnologien, wie beispielsweise bei Face ID, funktionieren auch bei völliger Dunkelheit sehr gut, erfordern aber besondere Vorkehrungen bei der Implementierung im Fahrzeugkontext, um optimale Ergebnisse zu erzielen. Eine solche Vorkehrung ist beispielsweise die Einhaltung des Abstands zwischen Sensor und Gesicht.

Die Positionierung der Kamera und der Sensoren in verschiedenen Fahrzeugformen ist ebenfalls ein entscheidender Faktor. Die Vielfalt der Fahrzeugdesigns erfordert eine flexible Anpassung der Sensorplatzierung, um eine optimale Erkennung von Gesichtern auf allen Sitzen zu gewährleisten. Es kann zu Verzerrungen aufgrund von Kamerawinkeln kommen, was den Einsatz von Bildverarbeitungsfiltern erfordert, um diese Herausforderung zu bewältigen.

Hinzu kommt der Aspekt des wechselnden Erscheinungsbildes der Insassen. Die Gesichtserkennungstechnologie muss in der Lage sein, Personen unabhängig von ihrer Kleidung, Frisur oder Brille zu erkennen.

3.3.3 Stimmerkennung

Die Stimmerkennungstechnologie (oder auch Stimmenerkennung) ist zwar noch nicht weit verbreitet, findet aber bereits Anwendung bei Kundendienst-Hotlines und Sprachassistenten wie Alexa, wo sie zur Authentifizierung der Identität einer Person eingesetzt wird, bevor personenbezogene Daten freigegeben werden. [28] [29] Diese Technologie hat das Potenzial, eine innovative Lösung zur Identifizierung von Fahrzeuginsassen zu bieten, durch die Nutzung individueller Stimmmerkmale. Angesichts dieser Informationen stellt sich die Frage, ob die Stimmerkennungstechnologie ihren Weg in die Fahrzeug-Insassenidentifizierung finden

kann.

Grundlagen

Vergleichbar mit der Gesichtserkennung, nutzt die Stimmerkennung biometrische Daten der Stimme, um mithilfe von ML verschiedene charakteristische Merkmale zu identifizieren. Dabei werden Mikrofone als primäre Erfassungsgeräte verwendet, um die Stimme zu erfassen und zu klassifizieren und so ein individuelles Stimmprofil zu erstellen, das auf einzigartigen Sprach- und Aussprachemerkmalen basiert.

In BMW-Fahrzeugen werden bereits Mikrofone für Freisprecheinrichtungen und für die Nutzung von Sprachassistenten eingesetzt. Diese bestehende Infrastruktur bietet eine solide Basis für die Integration einer Stimmerkennungstechnologie in den Fahrzeugkontext. Darüber hinaus wird die Position der Stimme bereits erkannt.

Die Stimmerkennung funktioniert am besten, wenn nur eine kleine Datenmenge klassifiziert werden muss. Eine Studie zeigt, dass bei zwei Sprechern eine Genauigkeit von 98 % erreicht wurde. Allerdings ist zu vermerken, dass die Erkennung bei einer Datenbasis von 20 Sprechern eine Genauigkeit von 90,2 % erreicht. [30] Auch wenn diese Ergebnisse zunächst nicht sehr präzise erscheinen, wäre es durchaus aufschlussreich, eine mögliche Integration in das Fahrzeug zu untersuchen.

Der Einsatz der Stimmerkennungstechnologie zur Insassenidentifizierung eröffnet somit die Möglichkeit, ein alternatives biometrisches Identifizierungsverfahren im Fahrzeug zu etablieren.

Konzeptionierung

Auch hier sind Parallelen zur Gesichtserkennung zu erkennen, wobei die Kamera durch ein- oder mehrere Mikrofone ersetzt wird. ML wird genutzt, um eine Stimmerkennung durchzuführen, die den Insassen identifiziert, der gerade spricht. Dazu muss das ML-Modell zunächst trainiert werden, was voraussetzt, dass die Insassen identifizierbare Stimmprofile erstellen, indem sie eine Reihe von Sätzen

sprechen, die vom Modell aufgenommen und klassifiziert werden. Beim Carsharing kann das Training über eine mobile App mit Anbindung an ein Backend erfolgen, welches nach dem Training die Stimmprofile über alle Poolfahrzeuge synchronisiert. Im familiären Kontext kann das Training lokal im Fahrzeug stattfinden, was auch ohne die Verwendung einer BMW ID Synchronisierung erfolgen kann und somit keine Backend-Anbindung benötigt.

Zur Lokalisierung des Insassen, während der Stimmerkennung können verschiedene Methoden eingesetzt werden. Durch die Verwendung eines Mikrofons pro Sitzplatz kann beispielsweise ein Dezibel-Schwellenwert verwendet werden, um die lauteste Stimme zu isolieren, die gerade an diesem Platz gesprochen wird. Alternativ kann mit Hilfe der Mikrofontriangulation, bei der drei Mikrofone zum Einsatz kommen, die genaue Position der Stimme bestimmt werden. [31]

Diese innovative Herangehensweise eröffnet die Möglichkeit, die Stimmerkennung effektiv für die Insassenidentifizierung zu nutzen und dabei auf die bereits vorhandenen Infrastrukturen in modernen Fahrzeugen zurückzugreifen.

Vorteile

Diese Technologie weist eine Reihe von Vorteilen auf, die ihre Attraktivität für die Fahrzeugintegration unterstreichen. Einer der herausragenden Vorteile ist, dass die Stimmerkennung, wie auch die Gesichtserkennung, keine aktive Beteiligung der Insassen erfordert. Die Insassen müssen keine manuellen Tätigkeiten unternehmen, sondern können sich ganz natürlich im Fahrzeug unterhalten, ohne ihre Identität aktiv bestätigen zu müssen.

Ein weiterer Vorteil liegt in der kostengünstigeren Hardware im Vergleich zur Gesichtserkennung. Mikrofone sind in der Regel wesentlich kosteneffizienter als Kameras und Sensoren, die bei der Gesichtserkennung zum Einsatz kommen. Dieses Kostenersparnis könnte ein entscheidender Faktor bei der Implementierung dieser Identifizierungsmethode sein. Die Flexibilität bei der Platzierung von Mikrofonen ist ein weiterer Vorteil. Töne sind leichter zu erfassen als Bilder, da Schallwellen Hindernisse umgehen können, während Licht und visuelle Signale im Allgemeinen einer geraden Linie folgen. Im Gegensatz zu Kameras lassen sich Mikrofone leicht

und unauffällig im Fahrzeuginnenraum verstecken. Dies ermöglicht eine diskrete Integration, ohne das Design oder den Komfort zu beeinträchtigen.

Zusammenfassend bieten die genannten Vorteile der Stimmerkennungstechnologie eine vielversprechende Grundlage für ihre Anwendung als effektive und benutzerfreundliche Methode zur Insassenidentifizierung in Fahrzeugen.

Herausforderungen

Dennoch ergeben sich bei der Integration dieser Technologie in das Fahrzeug mehrere Komplikationen. Eines der Hauptprobleme betrifft den Schutz der Privatsphäre. Die Stimme eines Menschen wird als biometrisches Mittel auch zu den besonderen Kategorien eingestuft und ist unter Rechtmäßigkeitsvoraussetzungen zu verarbeiten.

Ein weiteres Szenario, das die Stimmerkennung vor eine Herausforderung stellt, kommt zur Geltung, wenn die Insassen oder der Einzelfahrer schweigen. Konträr dazu Situationen, in denen mehrere Insassen gleichzeitig sprechen. Hintergrundgeräusche, sei es durch Fahrgeräusche, Musik oder andere akustische Elemente, können die Stimmerkennung beeinträchtigen. Diese Situationen erfordern fortschrittliche Algorithmen, um Zuverlässigkeit und Präzision zu gewährleisten.

Verschiedene Stimmlagen, die durch Stimmungsschwankungen oder gesundheitsbedingte Veränderungen entstehen, sind ebenfalls zu berücksichtigen, sowie die Ähnlichkeit von Stimmen, insbesondere bei Geschwistern oder Personen mit ähnlichen stimmlichen Merkmalen.

Insgesamt erfordert die erfolgreiche Integration eine ganzheitliche Berücksichtigung dieser Herausforderungen und eine kontinuierliche Weiterentwicklung der Technologie, um ihre Wirksamkeit in unterschiedlichen Szenarien zu gewährleisten.

3.3.4 Kombination der Technologien

Eine Möglichkeit besteht darin, verschiedene Technologien zu kombinieren, um eine umfassende und zuverlässige Identifizierung der Insassen zu ermöglichen. Durch die Kombination dieser Technologien können die Vorteile verschiedener Ansätze genutzt und gleichzeitig Herausforderungen minimiert werden.

Die Kombination aus Gesichts- und Stimmerkennungstechnologie könnte eine effektive Methode zur Identifizierung von Insassen darstellen. Die Gesichtserkennung könnte eine erste Prognose zur Identifizierung der Insassen liefern, während die Stimmerkennung die Ergebnisse bestätigt. Die Identifizierung des Fahrers könnte weiterhin durch ein schlüsselbasiertes System erfolgen, während die weiteren Insassen durch Gesichts- und Stimmerkennung identifiziert werden.

Die Verknüpfung dieser Technologien würde eine Triangulation der Identifikation ermöglichen, wodurch die Genauigkeit und Zuverlässigkeit der Insassenidentifikation im Fahrzeug erhöht werden könnte.

4 | Implementierung

Das folgende Kapitel liefert einen Einblick in die implementierte Identifizierungslösung. Es ist zu beachten, dass die Umsetzung im Fahrzeugkontext einen anderen Ansatz als bei herkömmlichen Lösungen erfordert. Die Konzeptionierung der vorgestellten Lösungsansätze in der Recherche dient als Leitfaden für die anschließende Implementierung. Aufgrund der Vielschichtigkeit erfolgt die Ausarbeitung des Konzeptes im Verlauf der Implementierung.

4.1 Auswahl der Identifizierungstechnologie

Als Implementierungsmethode wurde die Stimmerkennung gewählt, die speziell von der ODM-Abteilung von BMW gefordert wurde. Diese Entscheidung resultierte aus der Tatsache, dass dieser Ansatz bisher noch nicht im Fahrzeugkontext betrachtet wurde. Diese Technologie besitzt ein ungenutztes Potenzial für einen innovativen Ansatz im Fahrzeugbereich.

Die Stimmerkennungstechnologie, wie in Kapitel 3.3.3 beschrieben, präsentiert klare Vorzüge im Vergleich zu bestehenden Lösungen wie der Gesichtserkennung. Dies ist insbesondere auf die bereits im Fahrzeug vorhandenen Mikrofone zurückzuführen, was zu weniger Veränderungen in der Fahrzeugarchitektur führt.

Es ist von Interesse, die Zuverlässigkeit dieser Technologie in verschiedenen Situationen zu bewerten. Die beschriebenen Herausforderungen werden im Anschluss sorgfältig ausgewertet und analysiert.

4.2 Anforderungen

Die Anforderungen an die Implementierung beruhen auf den ersten Forschungsergebnissen und sollen als Leitfaden für die weitere Entwicklung dienen. Die Implementierung ist als Proof of Concept ausgelegt und zielt nicht auf eine endgültige Lösung ab. Sie soll die Umsetzbarkeit der Stimmerkennung in einem Fahrzeugkontext demonstrieren und als Grundlage für weitere Entwicklungen dienen.

Einerseits sollte die Anwendung für den Einsatz im Fahrzeug optimiert sein, andererseits aber auch übertragbar sein, um in verschiedenen Umgebungen getestet werden zu können. Diese Vielseitigkeit ermöglicht eine ausführliche Evaluierung der Technologie in verschiedenen Szenarien.

Ein Hauptkriterium ist die Erkennung von bis zu fünf Personen und deren Sitzposition. Dieser Aspekt ist entscheidend für den praktischen Einsatz der Stimmerkennung im Fahrzeugumfeld. Darüber hinaus ist es wichtig auf Benutzerfreundlichkeit zu achten, um die Durchführung von Versuchen zu erleichtern.

Effiziente Ressourcennutzung ist ein weiteres Anforderungskriterium. Die Implementierung sollte in der Lage sein, vorhandene Ressourcen optimal zu nutzen, um eine Funktionalität bei begrenzter Rechenleistung im Fahrzeug sicherzustellen.

4.3 Entwurf der Architektur

Es wurde eine vierstufige Architektur gewählt, bei der jede Komponente eine bestimmte Funktion erfüllt. Dieses Design ähnelt einer Microservices-Struktur, da es ihre Aufgaben auf verschiedene Systeme verteilt. Im Gegensatz zu einem Monolithen ist dies einfacher zu testen und einzusetzen. Auch die Skalierbarkeit einer Microservice-Architektur ist wesentlich simpler, da sie bei Bedarf angepasst werden kann. [8]

Im Fahrzeug selbst sind fünf Mikrofone installiert. Als Ersatz für die Headunit wird eine Incar-Komponente eingesetzt, die der begrenzten Rechenleistung des

Systems entgegenkommt. Aus diesem Grund leitet die Incar-Komponente die Sprachdaten nur weiter und führt selbst keine Klassifizierung durch.

Ein Backend-Server ist für die umfassende Verarbeitung aller Informationen zuständig. Die übrigen Komponenten kommunizieren ausschließlich mit dem Backend, wo sie dann verarbeitet und weitergeleitet werden.

Die direkte Nutzerschnittstelle ist eine Frontend Webapp, über die der Anwender sämtliche Funktionen im Backend ausführen kann. Hierzu zählt die Verwaltung von Sprechern, das Training des Modells und das Einsehen der Ergebnisse der Klassifizierung.

Außerdem gibt es eine separate Backendkomponente, das so genannte GPU-Backend, das für das Training des Modells verwendet wird. Die Trennung ergibt sich aus den Kosten für eine grafikartenbeschleunigte Instanz, die nur dann aktiviert wird, wenn das Training erforderlich ist. Die Klassifizierungen laufen über das normale Backend, da diese weniger rechenintensiv sind. Diese Struktur gewährleistet eine effiziente Nutzung der Ressourcen und eine Kostenoptimierung.

Abbildung 4.1 veranschaulicht die beschriebene Architektur, die einer Sternarchitektur entspricht, was für die Abkapselung einzelner Komponenten vorteilhaft ist. So können Komponenten wie das GPU-Backend abgeschaltet werden, um Kosten zu sparen. In der Praxis würde das System auch dann weiterlaufen, wenn die Incar-Komponente nicht angeschlossen ist oder aufgrund eines schwachen Mobilfunknetzes ausfällt.

Für die Implementierung im Rahmen dieser Bachelorarbeit werden diese Komponenten aus verschiedenen Gründen mittels Docker containerisiert. Unter anderem überzeugt eine Abstraktionsschicht, bei der Code und Abhängigkeiten in einem Container zusammengeführt werden. Dies erleichtert die Entwicklung auf einer Plattform und die Bereitstellung in der Cloud, was die Produktivität erhöht. [8]

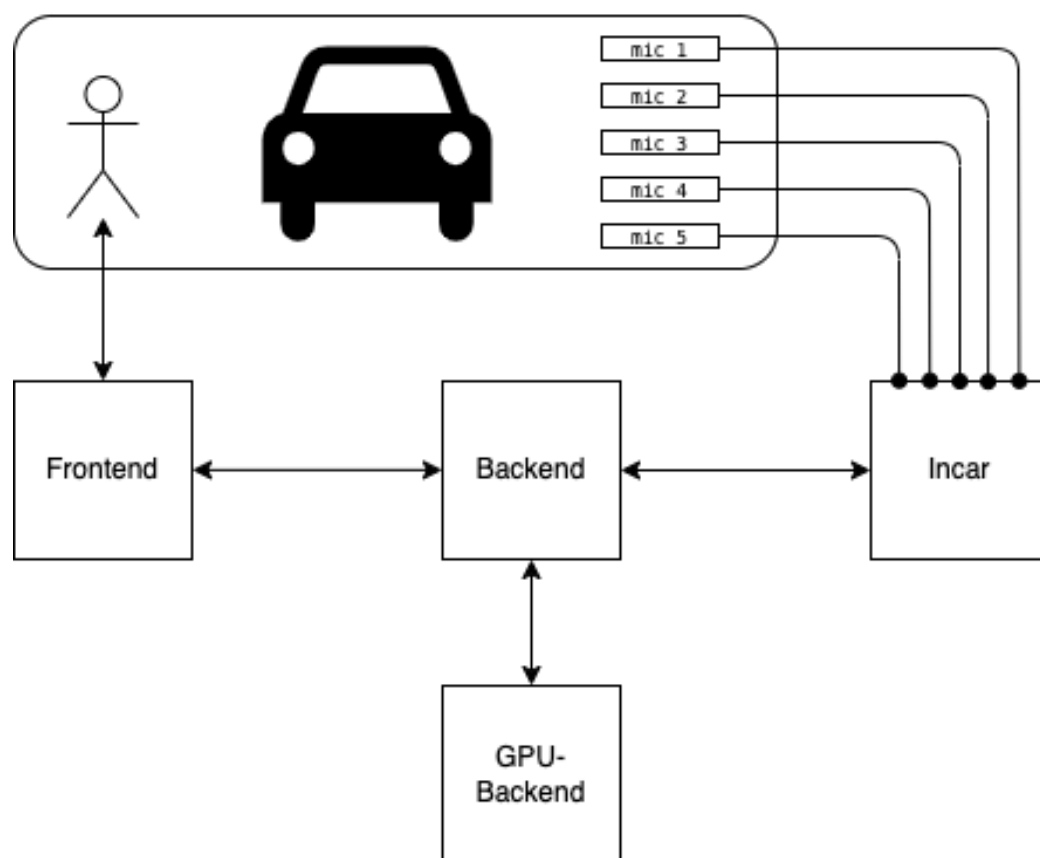


Abbildung 4.1: Architektur der Implementierung

4.4 Auswahl der Hardware

Die gewählte Hardware ermöglicht eine fahrzeugnahe Umsetzung im BMW-Kontext. Zum Einsatz kommen fünf Lavaliermikrofone der Firma AGPTEK mit einem 3,5mm Klinkenanschluss. Es sollten fünf identische Mikrofone verwendet werden, um mögliche Abweichungen bei den Ergebnissen zu vermeiden.

Der Einfachheit halber wurden externe Mikrofone gewählt, da der Anschluss an die bereits im Fahrzeug verbauten Mikrofone in Entwicklungsfahrzeugen nur unter bestimmten Bedingungen möglich ist. Die gewählte Lösung ist daher auch für Tests unter wechselnden Bedingungen besser geeignet, nämlich für Tests außerhalb des Fahrzeugs.

Die Incar-Komponente kann auf unterschiedlichen Systemen laufen. Für diese Implementierung wird ein Apple MacBook Pro 13 Zoll mit vier Thunderbolt 3 Anschlüssen und einem Klinkenanschluss der 2020er Generation verwendet. Ein Mikrofon wird direkt über eine 3,5-mm-Klinkenbuchse angeschlossen und die restlichen Mikrofone mittels eines Apple USB-C auf Klinkenadapters angeschlossen. Nach mehreren Fehlversuchen war dies die am besten funktionierende Lösung. Der ursprüngliche Ansatz mit fünf USB-A-Mikrofonen, die über einen USB-Hub angeschlossen waren, erwies sich als nicht erfolgreich, da die Mikrofone eines Herstellers die gleiche Geräte-ID im USB-Bus besitzen. Das bedeutet, dass nur die zuletzt angeschlossenen Mikrofone angesprochen werden konnten.

Eine Amazon Web Services (AWS) Elastic Compute Cloud (EC2) Instanz wird für die Backend- und Frontend-Komponenten eingesetzt. Diese wird durch BMW bereitgestellt, um eine flexible und skalierbare Lösung für die Durchführung von Tests im Fahrzeug zu bieten. Es wurde eine 'EC2-t4g.large' Instanz gewählt, die durch ihr Preis-Leistungs-Verhältnis bei allgemeinen Arbeitslasten überzeugt. Sie ist mit zwei virtuellen ARM-basierten Prozessorkernen und 8GiB Arbeitsspeicher ausgestattet. Diese Konfiguration ist für die Entwicklung und den mäßigen Einsatz zu Test- und Evaluierungszwecken gedacht und kann bei Bedarf skaliert werden. [32]

Das GPU-Backend läuft zunächst lokal auf einem GPU-beschleunigten Computer. Dieser verfügt über einen Intel i7 4790K, 16GiB RAM und eine Nvidia GTX 970 mit 4GiB Grafikspeicher. Durch die Verwendung der 1664 CUDA-Recheneinheiten [33] konnten die Trainingszeiten im Vergleich zur t4g-Instanz deutlich reduziert werden.

In Zukunft soll das ML-Backend auch auf einer AWS EC2-Instanz laufen, die GPU-beschleunigt ist. Das gesamte System wäre dann lediglich über einen Amazon Simple Storage Server (S3) Bucket [34] mit dem Backend verknüpft, was die Speicherung von, in diesem Fall ML-Modelle, zwischen verschiedenen Komponenten vereinfacht.

4.5 Softwarekomponenten

Die einzelnen Komponenten werden im Folgenden erläutert. Es wird unter anderem gezeigt, welche Anforderungen sie zu erfüllen haben und wie verschiedene Problemstellungen gelöst wurden. Durch die Verwendung von UML-Diagrammen und einzelnen Codeblöcken wird die Implementierung und bestimmte entscheidende Faktoren erläutert.

Der Vollständigkeit halber ist es wichtig zu erwähnen, dass Sockets für die Kommunikation verwendet werden. Die genauen Schnittstellen und Kommunikationsprozesse werden im Verlauf der Implementierung erläutert.

4.5.1 Frontend

Die Frontend-Komponente ist zweigeteilt: Ein Nginx-Reverse-Proxy und eine React-App, die mit JavaScript gebaut wurde.

Für die Entwicklung der React-App wurden zunächst Mocks erstellt. Die Benutzeroberfläche wurde auf der Grundlage dieser Mocks entworfen. Es sei darauf hingewiesen, dass bei diesem Prozess keine besondere Rücksicht auf UX/UI ge-

nommen wurde, da das Hauptziel darin besteht, den Benutzer, in diesem Fall den Entwickler, beim Testen zu unterstützen und keine endgültige Lösung zu liefern.

Der Startbildschirm der Web-App besteht aus fünf Feldern, die die identifizierten Insassen und ihre Position anzeigen. Außerdem gibt es eine Navigationsleiste am oberen Rand des Bildschirms, die einen Menüreiter enthält, über den der Benutzer verschiedene Funktionen ausführen kann. Darüber hinaus gibt es eine Start-/Stop-Taste, mit der die Insassenerkennung im Fahrzeug aktiviert bzw. deaktiviert wird. Die Navigationsleiste ist persistent und kann auf jeder Seite der Web-App gefunden werden. Am unteren Rand des Bildschirms befindet sich eine Statusleiste, die den aktuellen Verbindungsstatus mit dem Backend, dem GPU-Backend und der Incar-Komponente anzeigt. Das aktuell ausgewählte ML-Modell wird ebenfalls angezeigt. Die Statusleiste ist auch persistent.

Über das Menü kann der Benutzer auf die Seite "View Speakers" zugreifen, auf der alle aktuell aufgezeichneten Sprecher oder, in Bezug auf ML, die Identifizierungsklassen für die Klassifizierung aufgelistet sind. Hier ist es möglich, Sprecher zu entfernen.

Ein weiterer Menüpunkt ist "Add Speaker", wo ein Sprecher hinzugefügt werden kann. Der Nutzer hat die Möglichkeit, einen beliebigen Artikel als Lesematerial zu wählen. Über eine Taste werden die Mikrofonberechtigungen vom Browser erteilt und die Aufnahme kann erfolgen. Im Anschluss kann die Aufzeichnung nochmals angehört und ein Sprechername eingegeben werden. Die Daten können nun an das Backend gesendet werden.

Der Ablauf der Aufzeichnung des Mikrofons ist in den folgenden Codeabschnitten dargestellt. Die MediaStream Recording API wird verwendet, um auf das Mikrofon des Geräts zuzugreifen.

Die Methode `getUserMedia()` gibt ein Promise-Objekt zurück, das ein `MediaStream`-Objekt enthält, welches die angeforderten Mediengeräte darstellt, in diesem Fall den Mikrofon-Datenstrom. [35]

```
const stream = await navigator.mediaDevices.getUserMedia({
```

```
    audio: true,  
    video: false,  
  });
```

Die Aufzeichnung wird mit folgendem Codeausschnitt gestartet. Es wird ein MediaRecorder-Objekt erzeugt, das den Datenstrom des Mikrofons aufzeichnet. Die Audiodaten werden in einem Array gespeichert.

```
const startRecording = async () => {  
  // Erstellt eine neue Media-Recorder-Instanz mit dem Stream  
  const media = new MediaRecorder(stream, { type: "audio/wav" });  
  // Setzt die MediaRecorder-Instanz auf die Referenz  
  mediaRecorder.current = media;  
  // Startet den Aufnahmeprozess  
  mediaRecorder.current.start();  
  let localAudioChunks = [];  
  // Fügt die Audiodaten in ein Array ein  
  mediaRecorder.current.ondataavailable = (event) => {  
    if (typeof event.data === "undefined") return;  
    if (event.data.size === 0) return;  
    localAudioChunks.push(event.data);  
  };  
  // Speichert das Array für die Audiodaten  
  setAudioChunks(localAudioChunks);  
};
```

Mit dem folgenden Code-Snippet wird die Aufnahme gestoppt und zur Verfügung gestellt. Die Audiodaten werden in ein Blob-Objekt konvertiert, das versendet werden kann, sowie in eine URL, die im Frontend abgespielt werden kann.

```
const stopRecording = () => {  
  setRecordingStatus("inactive");  
}
```

```
// Stoppt die Aufnahmeinstanz
mediaRecorder.current.stop();
mediaRecorder.current.onstop = () => {
    // Erzeugt ein Blob-Objekt aus den Audiodaten
    const audioBlob = new Blob(audioChunks, { type: "audio/wav" });
    setAudioBlob(audioBlob); // speichert das Blob-Objekt zum versenden
    // Erzeugt eine URL aus den Audiodaten zum abspielen
    const audioUrl = URL.createObjectURL(audioBlob);
    // Leert das Array für die nächsten Aufnahmen
    setAudioChunks([]);
};
};
```

Die Seite "Train Model" ermöglicht es dem Benutzer, das Training auf dem GPU-Backend zu starten. Hier kann auch die Anzahl der Epochen für das Training festgelegt werden, wobei 100 Epochen voreingestellt sind. Ein Icon zeigt den Status des Trainingsprozesses an, bei erfolgreichem Training wird ein grüner Haken angezeigt und das neue Modell in der Statusleiste am unteren Rand angezeigt.

Einige Bildschirmaufnahmen des beschriebenen Frontends sind in Anhang 1 zu finden. Um das Frontend funktional im Browser darzustellen, wurde ein Reverse Proxy mit Nginx verwendet. Somit können ebenfalls Anfragen direkt an das Backend weitergeleitet werden. Außerdem wird der Datenverkehr über HTTPS umgeleitet, da einige Browser eine Mikrofonberechtigung nur über eine SSL-Terminierung zulassen.

Im Folgenden sind die wichtigsten Konfigurationseinstellungen für den Nginx-Reverse-Proxy aufgeführt:

```
http {
    # Datenverkehr über HTTPS umleiten
    server {
        listen 80;
```

```
        server_name [...];
        return 301 https://[...]$request_uri;
    }
}

server {
    # HTTPS-Konfiguration
    listen 443 ssl;
    server_name [...];

    ssl_certificate [...];
    ssl_certificate_key [...];

    # Für die Kommunikation mit dem Backend
    client_max_body_size 40MB; # maximale Größe der Anfrage
                             # kann bis zu 40MB sein

    location / {
        [...]
    }

    # API-Endpunkt für die Kommunikation mit dem Backend
    location /api {
        proxy_pass http://backend:5005;
        [...]
    }
}
```

4.5.2 Backend

Das Backend wurde in Python 3.11.6 implementiert. Wie bereits erwähnt, ist das Backend für die Steuerung und Verarbeitung aller Anfragen der Komponenten zuständig. Python wurde als Sprache gewählt, da es ein schnelles prototypisches

Erstellen und Entwickeln ermöglicht, was sich für ein solches 'Proof of Concept' anbietet. Außerdem bietet die Sprache umfangreiche Bibliotheken und wird häufig im Zusammenhang mit ML verwendet.

4.5.2.1 Zustand des Backends

Globale Variablen werden definiert, um den aktuellen Zustand des Systems zu verfolgen und um Fehler abzufangen, zum Beispiel wenn eine Komponente nicht läuft. Insbesondere zeigen die Booleans `recognition_status` an, ob die Identifizierung im Fahrzeug aktiv ist, `connected_incar` signalisiert eine bestehende Verbindung zur Incar-Komponente und `connected_ML` gibt an, ob eine Verbindung zum GPU-Backend besteht.

Für die Verwaltung der Zustände der Sprecher, der gespeicherten Modelle und der identifizierten Sprecher, die im Frontend angezeigt werden, werden JSON-Objekte verwendet. Der Vorteil von JSON-Objekten ist, dass dieses Format plattformunabhängig ist und von jeder Komponente gelesen werden kann. Die Unterstützung von Arrays in JSON-Objekten ist ebenfalls nützlich, um pro Objekt unterschiedliche Daten zu speichern. Der folgende Auszug aus der Datei `speaker.json` zeigt ein Beispiel:

```
[
  {
    "id": 0,
    "name": "Peter",
    "storing_date": "24-02-2024_13:21:34",
    "success": "true"
  },
  {
    "id": 1,
    "name": "Manfred",
    "storing_date": "24-02-2024_13:24:11",
    "success": "true"
  }
]
```

```
},  
[...]  
]
```

Die Verwendung von JSON-Objekten ermöglicht eine standardisierte und plattformübergreifende Datenrepräsentation, die die Kompatibilität zwischen verschiedenen Systemkomponenten sicherstellt.

Eine Flask-Anwendung aus der Flask-Bibliothek wird zur Verarbeitung von REST-Anfragen verwendet, wobei Sockets für die Kommunikation zwischen den Komponenten genutzt werden. Die Socket-Kommunikation erfolgt über Namespaces, die in separaten Klassen definiert sind, um die Anfragen der verschiedenen Komponenten getrennt zu behandeln.

```
class Webapp(Namespace):  
    # Wird aufgerufen, wenn die Webapp eine Verbindung herstellt  
    def on_connect(self):  
        print('Webapp connected')  
        get_status()  
        get_speakers()  
  
    # Wird aufgerufen, wenn die Webapp die Verbindung trennt  
    def on_disconnect(self):  
        print('Webapp disconnected')
```

Der obige Codeausschnitt verdeutlicht, wie eine Verbindung zur Webanwendung hergestellt wird. Diese Verbindung wird im Normalfall nicht unterbrochen. Um sicherzustellen, dass die Daten immer aktuell sind, wird sofort der aktuelle Status mit der Methode `get_status()` und der aktuelle Status der Sprecher mit `get_speakers()` übertragen. Diese Methoden werden auch aufgerufen, wenn sich die Daten ändern, zum Beispiel wenn neue Sprecher identifiziert, hinzugefügt oder entfernt werden oder wenn eine Verbindung zum GPU-Backend hergestellt wird.

4.5.2.2 Trainingsprozess

Nachdem alle Sprecherdaten erfasst und die Sprecherliste gepflegt wurde, kann das Modell trainiert werden.

Alle im `speakers.json` vermerkten Sprecher werden zusammen mit der Anzahl der Epochen über den Socket an das GPU-Backend gesendet und dort verarbeitet. Das Frontend wird informiert, dass die Daten gesendet wurden. Wenn das GPU-Backend das Training beendet hat, wird das Modell an das Backend zurückgesendet, wobei zu beachten ist, dass ein Modell in der Regel ca. 40 MB groß ist und daher in kleinere Pakete aufgeteilt und gesendet wird. Diese werden vom Backend mit einer ACK bzw. Nachricht bestätigt. Nachdem das letzte Paket gesendet wurde, erhält das Backend alle Informationen zu diesem Modell, die in einem JSON-Objekt gespeichert sind.

Wenn das Modell in der Validierung eine Genauigkeit von mehr als 95% und einen Verlust von weniger als 10% hat, wird es als neues Modell verwendet. In diesem Fall wird dem Frontend ebenfalls mitgeteilt, dass das Modell erfolgreich trainiert wurde und nun das neue Modell zur Verfügung steht.

4.5.2.3 Klassifikationsverfahren

Die Klassifizierung wird vom Backend angesteuert. Dies erfordert einen gut durchdachten Algorithmus, um die Sprecher korrekt zu identifizieren. Wenn im Frontend die "Start Recognition" Taste betätigt wird, sendet das Backend eine Nachricht an die Incar-Komponente und das aktuelle Modell wird mittels Keras geladen. Vorausgesetzt ist dass die Incar-Komponente verbunden ist. Ab diesem Zeitpunkt sendet die Incar-Komponente die Audiodaten an das Backend. Diese Audiodaten werden in einem temporären Verzeichnis gespeichert und indiziert. In Anhang 2 befindet sich ein UML-Sequenzdiagramm der Schnittstelle zwischen dem Backend und der Incar-Komponente.

4.5.3 GPU-Backend

Die GPU-beschleunigte Backend-Infrastruktur ist ebenfalls in Python implementiert und wird innerhalb der Implementierung als ML-Backend bezeichnet. Auf eine Containerisierung mittels Docker wurde bewusst verzichtet, um einen direkten Zugriff auf die GPU zu gewährleisten. Es sei jedoch darauf hingewiesen, dass die Nutzung der GPU in Verbindung mit Docker unter Umständen ebenfalls realisierbar wäre.

4.5.3.1 Vorbereitung der Trainingsdaten

Das Backend sendet alle Audiodaten, die trainiert werden sollen, und erstellt eine neue Verzeichnisstruktur für den Datensatz, in der alle gesendeten Audiodaten abgelegt werden. Mit Hilfe der "Pydub"-Bibliothek [36] und des Werkzeugs "FFmpeg" [37] werden aus den Audiodaten AudioSegment-Objekte erzeugt, die eine nachträgliche Anpassung der Samplerate auf 16000 Hz ermöglicht. FFmpeg ist eine leistungsfähige Softwarelösung zur Bearbeitung von Multimediadaten. Es ermöglicht die Konvertierung von Dateiformaten, die Anpassung von Sampleraten und bietet eine Vielzahl von Funktionen zur Audio- und Videobearbeitung. In diesem Fall werden die Audioaufnahmen in Segmente mit einer Länge von jeweils einer Sekunde aufgeteilt. In der Praxis bedeutet dies, dass eine Audiodatei mit einer Länge von einer Minute in 60 einzelne Audiosegmente zerlegt wird. Diese einzelnen Segmente werden als WAV-Dateien in einem eigenen Verzeichnis pro Sprecher gespeichert.

Im nächsten Schritt werden Hintergrundgeräusche für das Training präpariert, indem sie mit den Trainingsdaten überblendet werden, um mögliche Störgeräusche bei der Erkennung zu minimieren. Bei diesen Hintergrundgeräuschen handelt es sich sowohl um weißes Rauschen als auch um Fahrgeräusche, die beispielsweise durch das Überfahren von Bodenwellen oder das Abspielen leiser Musik im Hintergrund entstehen. Mit der Methode `add_noise()` werden die Geräusche verhältnismäßig zu den Audiosegmenten hinzugefügt. Die Integration von Hintergrundgeräuschen erhöht die Robustheit des Modells, da es dadurch besser in der

Lage ist, relevante Sprachmerkmale auch unter schwierigen akustischen Bedingungen zu erkennen. Dieser Ansatz erweitert nicht nur den Trainingsdatensatz, sondern trägt auch dazu bei, die Varianz der Daten zu erhöhen und Overfitting zu vermeiden. Overfitting bedeutet, dass ein Modell während des Trainings zu stark an die spezifischen Merkmale und Muster der Trainingsdaten angepasst wird. Das Modell passt sich so genau an die Trainingsdaten an, dass es Schwierigkeiten hat, verallgemeinerbare und zuverlässige Vorhersagen für neue, noch nicht gesehene Daten zu machen. [38]

Der Datensatz wird nun für das Training vorbereitet, wobei 10% der Audiodaten für die Validierung reserviert werden, während der Rest für das eigentliche Training verwendet wird. Die Sprecher-IDs, die vom Backend zur Verfügung gestellt werden, werden für die Erstellung der Klassifikations-Labels verwendet.

Nach diesen Vorbereitungen ist der Datensatz einsatzbereit und kann für das Training verwendet werden.

4.5.3.2 Aufbau des Modells

Die Modellentwicklung erfolgt mit Hilfe der Keras-Bibliothek, einer Open-Source Deep-Learning-Bibliothek, die als High-Level-API offiziell in TensorFlow eingebunden ist. Keras ermöglicht es Entwicklern, neuronale Netze effizient zu entwerfen, zu trainieren und zu evaluieren. Das hier verwendete Modell wurde von Fadi Badine in der Keras-Dokumentation als Beispiel vorgestellt. [39]

Dieses spezifische Modell unterstützt die Klassifizierung von Sprechern auf der Grundlage der Frequenzbereichsdarstellung von Sprachaufnahmen, die durch die Verwendung der Fast Fourier Transform (FFT) gewonnen wird. Das Modell selbst wird durch ein eindimensionales Konvolutionsnetzwerk mit Residualverbindungen erzeugt, welches speziell auf die Anforderungen der Audio-Klassifikation zugeschnitten ist.

Die Abbildung 4.2 wurde mit Hilfe der Visualisierungsbibliothek "visualkeras" erstellt und illustriert die Struktur des Modells, insbesondere den Aufbau der

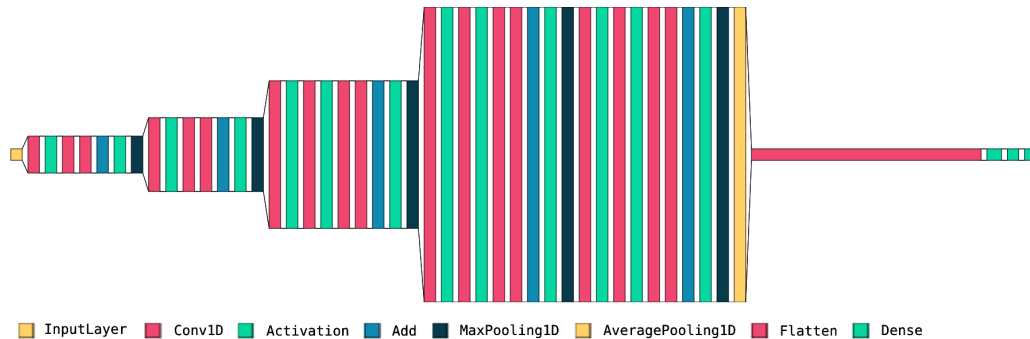


Abbildung 4.2: Aufbau des ML-Modells

verschiedenen Schichten. Eine detailliertere Version des Modells ist in Anhang 3 zu finden.

Bei der Implementierung kommt ein Convolutional Neural Network (CNN) zum Einsatz, das auf den so genannten “residual blocks” basiert. Diese speziellen Blöcke wurden in Residual Networks (ResNet) eingeführt, um das Problem des verschwindenden Gradienten während des Trainings zu lösen. [38] Jeder Residualblock besteht aus einer Shortcut-Verbindung und mehreren Faltungsebenen. Die Shortcut-Verbindung ermöglicht die direkte Ausbreitung des Gradienten und erleichtert somit das Training tiefer Netzwerke.

Die Eingabeschicht besteht aus einem 1×8000 Tensor, da für das Training die positiven Frequenzen der FFT verwendet werden, also $16000/2$. Die Architektur des Modells besteht aus mehreren Aufrufen der Methode `residual_block()`. Anfangs werden Residualblöcke mit verschiedenen Filtergrößen und Faltungsschichten verwendet. Nach den Residualblöcken wird eine Average Pooling Schicht mit einer Poolinggröße von 3 und einem Stride von 3 verwendet. Diese Schicht reduziert die räumliche Dimension der Daten, indem die Werte in jedem Pooling-Fenster gemittelt werden. Darauf folgt eine Flattening-Schicht, die die Ausgabe in einen eindimensionalen Vektor umwandelt, um die Daten für die nachfolgenden Dense-Schichten vorzubereiten.

Durch die Dense-Schichten ist es dem Modell möglich, komplexe Merkmale zu extrahieren, wo bis zu 256 Neuronen zum Einsatz kommen. Auf die Ausgabe dieser Schicht wird die Aktivierungsfunktion “relu” (Rectified Linear Unit) angewen-

det, um nichtlineare Aspekte in das Modell zu integrieren und die Fähigkeit des Netzes zur Mustererkennung zu verbessern. Schließlich wird eine DenseSchicht mit einer Softmax-Aktivierungsfunktion verwendet, um die Klassifikation durchzuführen. Diese wandelt die Ausgabe in eine Wahrscheinlichkeitsverteilung über die Klassen um, wodurch die Klasse mit der höchsten Wahrscheinlichkeit als Sprechervorhersage bestimmt wird. Diese Schichten arbeiten zusammen, um eine robuste und leistungsfähige Klassifikationsarchitektur zu bilden.

Das Modell verwendet den Adam-Optimierer, einen adaptiven Algorithmus, der die Effizienz des Trainings tiefer neuronaler Netze verbessert. Dieser passt die Lernraten auf der Grundlage vergangener Gradienten an. Die begleitende "Sparse Categorical Crossentropy"-Verlustfunktion (Loss) misst die Differenz zwischen den vorhergesagten und den tatsächlichen Klassenlabels. Zusammen ermöglichen diese Komponenten eine effiziente Optimierung und Training für Klassifikationsaufgaben.

Zum besseren Verständnis des oben beschriebenen Modells wurde das Buch "Deep Learning Projects Using TensorFlow 2" von Vinita Silaparasetty als Literaturhilfe herangezogen. [38]

4.5.3.3 Trainingsprozess

Für das Training des Datensatzes wird die Funktion `model.fit()` verwendet. Dabei wird ein "Early Stopping Callback" verwendet, um Rechenressourcen zu sparen und das Training frühzeitig zu beenden, wenn keine weiteren Verbesserungen zu erwarten sind. Gleichzeitig wird ein "Model Checkpoint"-Mechanismus eingesetzt, um das beste Modell während des Trainings zu speichern und so sicherzustellen, dass die leistungsfähigste Version erhalten bleibt. Die Anzahl der Epochen wird vom Benutzer im Frontend festgelegt und hier eingestellt, aber durch den "Early Stopping Callback" ist es möglich, dass weniger Epochen durchlaufen werden, wenn das Modell bereits konvergiert ist.

[...]

```
earlystopping_cb = keras.callbacks.EarlyStopping(  

```

```
    patience=10, restore_best_weights=True)
mdlcheckpoint_cb = keras.callbacks.ModelCheckpoint(
    f"models/[...]", monitor="val_acc", save_best_only=True)

history = model.fit(
    training,
    validation_data=validation,
    epochs=epochs,
    callbacks=[earlystopping_cb, mdlcheckpoint_cb],
)
[...]
```

Nach Abschluss des Trainings wird das Modell mit der Dateiendung `.keras` gespeichert und zusätzlich werden alle Informationen zum Modell im JSON-Format angehängt. Anschließend wird das Modell zusammen mit den Informationen an das Backend gesendet. Dieser Vorgang wurde bereits in Kapitel 4.5.2.2 beschrieben.

4.5.4 Incar Komponente

Die Incar-Komponente, die sich im Fahrzeug befindet, ist ebenfalls in Python geschrieben. Hier soll der Ressourcenverbrauch geringgehalten werden.

4.5.4.1 Kalibrierung der Mikrofone

Beim Start der Komponente wird zunächst eine Liste aller angeschlossenen Ein- und Ausgabegeräte angezeigt. Diese werden mit Hilfe der „Sounddevice“-Bibliothek aufgelistet, die auch für die Aufzeichnung der Audiodaten verwendet wird. Aus dieser Liste wählt der Benutzer die Mikrofone aus, die für die Erkennung im Fahrzeug verwendet werden sollen. Im nächsten Schritt werden die Mikrofone auf ihre Position im Fahrzeug kalibriert. Dazu wird für jedes Mikrofon ein Dezibel-Graph angezeigt, der zur Identifizierung der Mikrofone

verwendet werden kann. In der Praxis kann der Benutzer die Mikrofone berühren und den Pegel während des Kalibrierungsprozesses beobachten. Die Positionen setzen aus den in Tabelle 4.1 aufgeführten Kürzeln zusammen.

Tabelle 4.1: Positionskürzel für die Kalibrierung der Mikrofone

Positionskürzel	Position im Fahrzeug
d	Fahrer
p	Beifahrer
bl	hinten links
br	hinten rechts
bm	hinten mittig

Es ist möglich, eine beliebige Anzahl an Mikrofonen zu kalibrieren, da dies beim Testen nützlich ist, jedoch werden die abweichenden Positionskürzeln vom Backend nicht berücksichtigt. Bei erfolgreicher Kalibrierung wird eine sogenannte „device_map“ ausgegeben, die eine Geräte-ID mit ihrer Position verknüpft. Diese kann angegeben werden, wenn die Komponente startet, um den Kalibrierungsprozess zu überspringen, falls keine Änderungen vorliegen.

In Anhang 4 ist der Kalibrierungsprozess mit drei Mikrofone dargestellt.

4.5.4.2 Aufnahme und Verarbeitung der Audiodaten

Nach der Kalibrierung wird eine Verbindung zum Backend hergestellt und ein Thread erzeugt, der für die Audioaufnahme verantwortlich ist. Dieser Thread startet eine Endlosschleife, die darauf wartet, dass ein „recognition_event“ als Ereignis ausgelöst wird, welches im unteren Codeabschnitt zu sehen ist. Dieses Ereignis wird vom Backend ausgelöst, wenn der Anwender die Sprecher Identifizierung aktiviert.

```
def recognition():  
    while True:
```

```
recognition_event.wait()
record(duration=2)
global audio_file_index # laufende Nummer der Aufnahmen
audio_file_index += 1
sleep(3)
```

Bei der Aufnahme der Sprachdaten ist zu beachten, dass alle kalibrierten Mikrofone gleichzeitig aufgenommen werden. Dies ermöglicht die individuelle Isolierung von Stimmen anhand des Vergleichs der Dezibel Werte.

Für die Aufnahme der einzelnen Mikrofone werden weitere Aufnahme-Threads verwendet, die parallele Aufnahmen ermöglichen. Die Bibliothek "Sounddevice" stellt einen Datenstrom vom Mikrofon zur Verfügung und ein weiterer Thread speichert den Datenstrom in eine WAV-Datei. Nach der Aufnahme aller Mikrofone werden die gesicherten Daten per Socket an das Backend gesendet, zusammen mit einer fortlaufenden Nummer, die es ermöglicht, den Stand im Fahrzeug mit dem Stand im Backend zu vergleichen.

Die Aufzeichnung der Mikrofone erfolgt in festgelegten Zeitintervallen, welche in Kapitel 5.1.2 untersucht werden. Diese gestaffelte Aufzeichnung stellt sicher, dass das Backend genügend Zeit hat, die Audiodaten zu verarbeiten, bevor der nächste Zyklus von der Incar-Komponente gesendet wird. Eine Erweiterung besteht darin, den Zyklus dynamisch an die Verbindungsgeschwindigkeit des Mobilfunknetzes anzupassen. Dadurch kann eine Überlastung des Netzes vermieden werden.

4.6 Zusammensetzung und Schnittstellen

Die implementierte Lösung besteht aus verschiedenen Komponenten, die mit dem Backend kommunizieren und teilweise gleichzeitig Daten austauschen. Die Abbildung 4.3 liefert einen Überblick über die Schritte, die notwendig sind, um Sprecher hinzuzufügen und das Modell zu trainieren. Sie veranschaulicht, wie die Komponenten miteinander interagieren und welche Schritte notwendig sind,

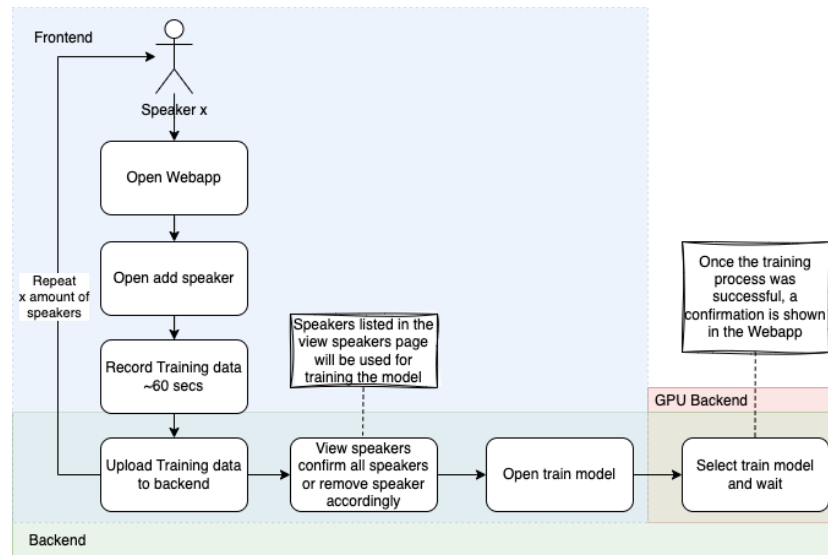


Abbildung 4.3: Schritte zum Hinzufügen von Sprechern und zum Trainieren des Modells

um das System zu nutzen. Wie bereits erläutert, erfolgt die Kommunikation hauptsächlich über Sockets.

Sockets ermöglichen eine Echtzeitkommunikation zwischen den einzelnen Komponenten, was insbesondere in Anwendungsfällen wie der Incar-Komponente von entscheidender Bedeutung ist. Dort ist eine schnelle und effiziente Datenübertragung mit geringer Latenz erforderlich. Durch die Möglichkeit der Full-Duplex-Kommunikation können Nachrichten in beide Richtungen gleichzeitig ausgetauscht werden, was die Leistungsfähigkeit und Reaktionsschnelligkeit der Gesamtsystemarchitektur erhöht. Ein weiterer Vorteil von Sockets ist, dass sie plattformunabhängig sind und somit zuverlässig in verschiedenen Komponenten eingesetzt werden können. [40]

Die Abbildung 4.4 zeigt die Schnittstellen zwischen den einzelnen Komponenten. In der Tabelle 4.2 wird die Kommunikation zwischen den Komponenten detailliert beschrieben.

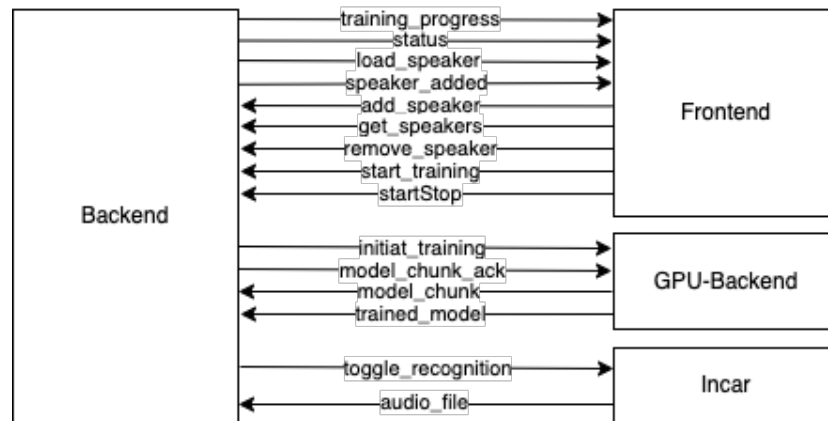


Abbildung 4.4: Dokumentation der Schnittstellen

Tabelle 4.2: Kommunikationsschnittstellen Beschreibung

Von	Nach	Nachricht	Payload
Frontend	Backend	get_speakers	-
Frontend	Backend	add_speaker	{speaker_name, audio_blob}
Frontend	Backend	remove_speaker	speaker_ID
Frontend	Backend	start_training	epochs
Frontend	Backend	startStop	-
Frontend	Backend	get_status	-
Backend	Frontend	status	{conn_incar, conn_ML, rec_status, rec_speakers, current_model}
Backend	Frontend	speakers	speakers.json
Backend	Frontend	training_progress	progress (in %)
Backend	Frontend	speaker_added	speaker_name
GPU-Backend	Backend	model_chunk	{chuncks[], filename}
GPU-Backend	Backend	trained_model	modelinfo.json

Von	Nach	Nachricht	Payload
Backend	GPU-Backend	initiat_training	{epochs, audio_files[]}
Backend	GPU-Backend	model_chunk_ack	-
Incar	Backend	audio_file	{index, audio: {d, p, bl, bm, br}}
Backend	Incar	toggle_recognition	recognition_status

5 | Evaluierung der Implementierung

Die Implementierung der Identifizierungstechnologie im Fahrzeugkontext erweist sich als komplexer Prozess, bei dem verschiedene Aspekte zu berücksichtigen sind. Die Bewertung der Implementierung erfolgt anhand der Kriterien Funktionalität und Genauigkeit, Datensicherheit sowie Benutzerfreundlichkeit. Die Ergebnisse der Implementierung werden in einem Gesamtkontext betrachtet und im Anschluss diskutiert.

5.1 Versuchsübersicht und Ergebnisse

Um die Funktionalität der Implementierung zu bewerten, werden verschiedene Versuche durchgeführt. Zunächst wird untersucht, wie das ML-Modell am besten trainiert werden kann, um eine hohe Genauigkeit zu erreichen, ohne den Insassen zu überfordern. Zusätzlich werden verschiedene Szenarien für die Aufnahme und Klassifizierung der Stimmdaten geprobt, um die Funktionalität des Systems unter verschiedenen Bedingungen zu testen.

5.1.1 Aufbau

Der Versuchsaufbau gliedert sich in zwei Abschnitte. Zuerst wurden Experimente zur Verbesserung des Erkennungsprozesses durchgeführt. Anschließend wurden

Experimente zur Bestimmung der Trainingslänge sowie zum Testen der Funktionalität durchgeführt.

Das Modelltraining wurde für alle Experimente im Vorfeld durchgeführt. Dazu wurden sechs Modelle mit fünf Klassen bzw. Sprechern trainiert, wobei pro Sprecher verschiedene Audioaufnahmen verwendet wurden, die sich insgesamt auf mindestens 15 Minuten summierten. Diese Audioaufnahmen wurden in verschiedenen Längen für das Training verwendet, um eine optimale Audiolänge für das Training zu bestimmen. Insgesamt wurden sechs Audiolängen ausgewertet. Die Tabelle 5.1 veranschaulicht den Trainingsprozess, wobei die Audiolänge der Trainingsdaten T_{train} , der Loss (Verlust) für die Trainingsdaten L_{train} und für die Validierungsdaten L_{val} , die Accuracy (Genauigkeit) für die Trainingsdaten A_{train} und für die Validierungsdaten A_{val} als auch die Anzahl der Epochen bis zur Konvergenz E und die Zeit bis zur Konvergenz T_{conv} aufgeführt sind. Alle Zeiten sind in Sekunden angegeben.

Tabelle 5.1: Übersicht der Modellergebnisse

T_{train}	L_{train}	L_{val}	A_{train}	A_{val}	E	T_{conv}
1	0.1455	0.3337	0.9352	0.8750	26	32
3	0.1344	0.2410	0.9506	0.9167	34	62
5	0.0421	0.1620	0.9806	0.9333	42	118
8	0.1217	0.3440	0.9595	0.9062	24	104
10	0.1630	0.3295	0.9394	0.9167	22	119
15	0.1813	0.2647	0.9302	0.9306	26	193

In Anhang 5 sind die detaillierten Ergebnisse der Modelltrainings anhand von Accuracy- und Loss-Graphen dargestellt. Die Graphen zeigen, dass die Modelle nach einer gewissen Anzahl von Epochen konvergieren und eine hohe Genauigkeit erreichen. Das Modell mit einer Trainingslänge von fünf Minuten erreichte die höchste Genauigkeit und konvergierte am schnellsten. Das liegt daran dass die längeren Trainingsdaten eine höhere Datenvielfalt bieten, was zu einer besseren Generalisierung des Modells führt. Dafür könnte es jedoch notwendig sein, die patience (Geduld) des “Early Stopping Callback” zu erhöhen um somit das Modell

länger trainieren zu lassen.

Für die Bestimmung der Parameter für den Identifizierungsprozess wurde das Modell mit 15-minütigen Trainingsdaten verwendet. Die Stimmerkennung wurde mit zwei Personen erprobt und mit zwei Mikrofonen an das Backend verbunden. Es ist zu beachten, dass diese Experimente in einer leeren, büroähnlichen Umgebung mit wenig Hintergrundgeräuschen durchgeführt wurden. Obwohl dies keine realistische Fahrzeugumgebung darstellt, hat sich die Wahl dieser Umgebung als optimal für schnelle Anpassungen der Parameter und wiederholte Tests erwiesen.

Weitere Experimente zur Bestimmung der optimalen Audiolänge für das Training wurden simuliert. Dazu wurde im Vorfeld ein Dialog der fünf trainierten Sprecher aufgenommen, der zur Evaluierung der verschiedenen Modelle verwendet wurde. Jeder Sprecher verfügte über ein eigenes Mikrofon, so dass für die Auswertung fünf Tonspuren entstanden sind. Die identischen Daten für die Klassifizierung gewährleisten einen homogenen Vergleich. Bei der Aufnahme des Dialogs wurde darauf geachtet, dass dieser mindestens drei Minuten lang ist und jeder Sprecher den gleichen Redeanteil hat.

Die funktionellen Tests wurden im Fahrzeug durchgeführt. Dort wurden fünf Mikrofone installiert, die mit der Incar-Komponente verbunden waren. Für die Durchführung der Tests wurde ein Fahrzeug der Kompaktklasse mit fünf Sitzplätzen gewählt. Die Lavaliermikrofone wurden am Dachhimmel des Fahrzeugs befestigt und jeweils auf den Kopf des sitzenden Insassen in der jeweiligen Sitzposition ausgerichtet. In Anhang 6 wird dieser Aufbau anhand von Bildern gezeigt. Die Incar-Komponente wurde über Mobilfunk mit dem Backend verbunden.

5.1.2 Bestimmung der Parameter

Es wurden verschiedene Kombinationen aus der Aufnahmedauer der Incar-Komponente und der Intervallzeit zwischen dem Versenden der Nachricht getestet. Die Parameter wurden mittels dynamischer Tests angepasst. Dabei konnten verschiedene Durchläufe mit angepassten Parametern zur Laufzeit bewertet und verbessert werden. Die Dokumentation der unterschiedlichen

Ergebnisse war für das Erkennen und Nachvollziehen von Zusammenhängen von großer Bedeutung.

Im Laufe der Tests konnten bereits Schwachstellen in der Fahrzeugerkennung erkannt werden, darunter die Verarbeitung von Situationen, in denen es still ist, das heißt in denen keiner der Insassen spricht. Dies führte häufig zu falschen Klassifizierungen. Aus dieser Erkenntnis wurde ein 'leerer Sprecher' trainiert, der ausschließlich aus Audiodaten besteht, die während einer ruhigen Leerfahrt aufgenommen wurden, ohne Stimmen. Dieser leere Sprecher wird bei der Bewertung nicht berücksichtigt und fließt nicht in die Berechnung des Medians zur Bestimmung des Sprechers ein.

Im Folgenden, eine kurze Dokumentation der beim Testen verwendeten Parameter, welche einen Zusammenhang erkennen lassen. Dabei steht t_D für die Aufnahmedauer, t_I für das Aufnahmeintervall, c für die Klassifizierungs-Confidence (auf zwei Dezimalstellen gerundet) und T für die Zeit, die erforderlich ist, um alle fünf Insassen richtig zu identifizieren. Alle Zeiten sind in Sekunden angegeben.

Tabelle 5.2: Versuchsergebnisse: Bestimmung der Parameter

t_D	t_I	c	T
1	1	0.82	51
1	2	0.79	58
5	1	0.93	34
5	3	0.91	41
10	5	0.94	65

Eine weitere Aufteilung der Audiodaten in Sekundenabschnitte hatte keinen Einfluss auf die Klassifikationsergebnisse und daher wurde immer die volle Länge von t_D als Modelleingabe verwendet. Aus Tabelle 5.2 ergab sich, dass eine kurze t_D keine guten Ergebnisse lieferte und T daher angepasst werden musste. Zu lange t_D erwiesen sich als unpraktisch, da sie eine längere Verarbeitungszeit durch das Socket verursachten und somit eine höhere t_I verwendet werden musste. Die Bestimmung von t_I hängt auch von der Geschwindigkeit des Verarbeitens im

Backend ab. Obwohl keine Audiodaten verloren gehen, da diese in einer Queue landen, sollten die Daten zügig behandelt werden, um keine Inkonsistenzen zu haben. Das bedeutet auch, dass das Backend pro Intervall umso länger arbeiten muss, je mehr Insassen erkannt werden.

Final wurde ein t_D von fünf Sekunden und ein t_I von drei Sekunden gewählt, da dies die besten Ergebnisse lieferte ohne die Verarbeitungszeit zu sehr zu erhöhen.

5.1.3 Bestimmung der Audiolänge

Die Audiolänge für den Trainingsprozess stellt einen entscheidenden Faktor für die Nutzerakzeptanz dar. Eine zu lange Audioaufnahme kann die Benutzer negativ beeinflussen, daher ist es wichtig, die Audiolänge auf ein Minimum zu reduzieren. Die im Versuchsaufbau trainierten Modelle wurden hierfür einzeln untersucht. Um die zuvor aufgezeichneten Dialoge an das Backend übertragen zu können, wurde eine Anpassung der Incar-Komponente vorgenommen, die anstelle der Mikrofondaten die Aufzeichnungen im gleichen Sendeformat überträgt.

Die Ergebnisse dieser Untersuchung wurden direkt aus dem Frontend extrahiert. Es wurde evaluiert, ob alle Sprecher korrekt erkannt wurden und wie lange es dauerte, bis alle Sprecher korrekt erkannt wurden.

Tabelle 5.3: Versuchsergebnisse: Bestimmung der Audiolänge

T_{train}	Erkannte		Bemerkungen
	Spre- cher	Erkennungs- dauer	
1	0	-	Abgebrochen nach 3 Minuten. Keine schlüssigen Ergebnisse. Die Klassifizierung erfolgte sehr sporadisch.
3	2	-	Abgebrochen nach 3 Minuten. Klassifizierung zwar besser aber nicht schlüssig.
5	5	~160s	Nimmt zu viel Zeit in Anspruch.
8	5	~80s	Nimmt weniger Zeit in Anspruch.

T_{train}	Erkannte		Bemerkungen
	Spre- cher	Erkennungs- dauer	
10	5	~65s	Nimmt weniger Zeit in Anspruch.
15	5	~60s	Nimmt minimal weniger Zeit in Anspruch.

Wie aus der Tabelle 5.3 hervorgeht, werden Sprecher erst ab einer Trainingsdatenlänge von fünf Minuten zuverlässig erkannt. Außerdem nimmt die Zeit bis zur vollständigen Erkennung mit zunehmender T_{train} ab. Die Erkennungsdauer ist jedoch nicht linear proportional zur T_{train} . Sie nimmt ab, bis sie bei einer T_{train} von acht Minuten konstant bleibt. Dies deutet darauf hin, dass eine T_{train} von circa acht Minuten ausreicht, um eine schnelle und zuverlässige Erkennung zu gewährleisten.

Interessanterweise wurde festgestellt, dass obwohl das Modell mit einer T_{train} von fünf Minuten die beste Genauigkeit beim Training erreichte, die Erkennungsdauer bei einer höheren Trainingsdauer deutlich kürzer war. Dies könnte darauf hindeuten, dass die Modelle mit einer höheren Trainingsdauer von acht Minuten besser generalisiert und somit schneller und zuverlässiger arbeitet.

5.1.4 Funktionalität

Basierend auf den vorherigen Erkenntnissen wurde ein neues Modell entwickelt und trainiert, welches drei Sprecher umfasst. Die jeweiligen Sprecher haben für das Training eine Audiolänge von etwa acht Minuten gesprochen. Verschiedene Szenarien wurden eingeführt, um die Leistungsfähigkeit der Implementierung zu bewerten. Die Szenarien wurden in fünf Versuchsvarianten unterteilt, die in Tabelle 5.4 aufgeführt sind.

Tabelle 5.4: Übersicht der Funktionsversuche

Versuch	Beschreibung
V_1	Keine Hintergrundmusik, Insassen sprechen in normaler Tonlage und kontinuierlich.
V_2	Keine Hintergrundmusik, Insassen sprechen in normaler Tonlage, jedoch nicht kontinuierlich.
V_3	Hintergrundmusik vorhanden, jedoch weiterhin normale Unterhaltung möglich, Insassen sprechen in normaler Tonlage und kontinuierlich.
V_4	Hintergrundmusik vorhanden, jedoch weiterhin normale Unterhaltung möglich, Insassen sprechen in erhöhter und aufgeregter Tonlage und kontinuierlich.
V_5	Laute Hintergrundmusik vorhanden, normale Unterhaltung erschwert, Insassen sprechen in erhöhter Tonlage und kontinuierlich.

Diese Versuche wurden während der Fahrt durchgeführt, so dass auch die Fahrgeräusche berücksichtigt wurden. Es ist zu beachten, dass kein Dialog vorgegeben wurde und die Insassen frei gesprochen haben. Diese Versuche sind daher nicht vollständig reproduzierbar, bieten aber einen realitätsnahen Kontext. Für jede Versuchsvariante wurde die Ausgabe im Frontend nach jeder vergangenen Minute aufgezeichnet und in Tabelle 5.5 festgehalten.

Tabelle 5.5: Versuchsergebnisse: Funktionalität

Versuch	$t > 1$	$t > 2$	$t > 3$	$t > 4$	$t > 5$
V_1	4	5	5	5	5
V_2	1	1	4	5	5
V_3	2	3	2	4	4
V_4	0	0	1	1	2
V_5	0	1	0	0	0

Die Ergebnisse des Versuchs ergaben, dass die Insassenidentifizierung in den gewöhnlichen Szenarien zuverlässig funktioniert. In Szenario V_1 und V_2 wurden

nach spätestens drei Minuten alle Insassen korrekt erkannt. In V_3 wurden die Insassen ebenfalls fasst vollständig korrekt erkannt, jedoch dauerte dies wegen der Hintergrundmusik etwas länger. In Szenario V_4 und V_5 wurden die Insassen nicht erkannt. Dies ist auf die laute Hintergrundmusik zurückzuführen, die die Stimmerkennung beeinträchtigt.

Aus diesen Erkenntnissen lässt sich schließen, dass die Implementierung in der Lage ist, Insassen in einem normalen Fahrzeugumfeld zuverlässig zu identifizieren. Jedoch sollte das Erstellen des Modells so angepasst werden, dass Störfaktoren wie Hintergrundmusik besser verarbeitet werden können.

5.1.5 Hinweis zur Validität

Unter den gegebenen Umständen ist es wichtig, darauf hinzuweisen, dass die durchgeführten Versuche nicht den strengen Kriterien wissenschaftlicher Validität entsprechen. Dies ist vor allem auf die zeitlichen Beschränkungen und die Schwierigkeit zurückzuführen, genügend Teilnehmer zu rekrutieren, die in der Lage wären, die für die Durchführung der Experimente erforderliche Zeit aufzubringen. Trotz dieser Einschränkungen ermöglichen die durchgeführten Versuche eine grundlegende Bewertung der Implementierung mit geringer Anzahl an Insassen. Diese ähneln einem Familienfahrzeugkontext und bilden ebenso eine solide Grundlage für zukünftige Forschungsarbeiten.

5.2 Bewertung der Implementierung

Aus den durchgeführten Versuchen konnten eine Reihe von Erkenntnissen gewonnen werden. Im Anschluss werden diese Erkenntnisse bewertet und Schlussfolgerungen sowie Verbesserungsmöglichkeiten für zukünftige Forschung im Bereich der Stimmerkennung und ihrer Umsetzung im Fahrzeugkontext diskutiert.

5.2.1 Datensicherheit

Die aktuelle Implementierung weist erhebliche Schwächen in Bezug auf den Schutz der personenbezogenen Daten auf. Die Übertragung aller Audiodaten für die Insassenerkennung erfolgt unverschlüsselt über ein Websocket zwischen der Incar-Komponente und dem Backend. Eine mögliche Alternative für eine datensichere Lösung wäre die lokale Ausführung der Insassenerkennung im Fahrzeug. Diese Herangehensweise wurde bereits in der Recherche als auch in der Implementierung diskutiert.

Eine sinnvolle Verbesserung der Datensicherheit ist möglich, wenn eine Verschlüsselung der Daten implementiert wird. Durch die Verwendung eines Verschlüsselungsalgorithmus, wie AES (Advanced Encryption Standard) kann sichergestellt werden, dass nur autorisierte Empfänger, in diesem Fall das Backend, die Daten entschlüsseln können. Die Verwendung von digitalen Zertifikaten kann ebenfalls gewährleisten, dass die Daten ausschließlich an den vorgesehenen Empfänger gesendet werden.

Darüber hinaus empfiehlt sich die Implementierung einer Integritätsprüfung, um sicherzustellen, dass die Daten während der Übertragung nicht manipuliert wurden.

Insgesamt ist es essenziell, diese Verbesserungen in einem realen Szenario zu implementieren, insbesondere wenn personenbezogene Daten verarbeitet werden, um die Einhaltung aller Datenschutzrichtlinien zu gewährleisten. Darüber hinaus ist zu beachten, dass der Einsatz von ML in zunehmendem Maße einer rechtlichen Überprüfung unterzogen wird.

5.2.2 Benutzerfreundlichkeit

Aus den durchgeführten Versuchen konnten erste Eindrücke der Versuchsteilnehmer gewonnen werden. Diese Eindrücke sind zwar wissenschaftlich nicht verwendbar, jedoch für eine erste Einschätzung des Systems aufschlussreich.

Während des Trainingsprozesses wurde festgestellt, dass Audioaufnahmen, die länger als drei Minuten andauern, als ziemlich unattraktiv und unkomfortabel empfunden werden. Eine sinnvollere Alternative könnte das Hochladen von voraufgezeichneten Audioquellen sein. Außerdem wurde festgestellt, dass das erzwungene Sprechen während des Versuchs als störend empfunden wird. Auch wenn dies in der Praxis nicht gefordert wird, ist es für das Funktionieren der Stimmerkennung notwendig, dass die Insassen sprechen.

Es zeigte sich, dass das Konzept der Insassenerkennung mittels Stimmerkennung von den Insassen als eher belastend empfunden wird. Interessanterweise haben die Teilnehmer jedoch das Potenzial der Insassenerkennung erkannt, was als Motivation für weitere Forschungsbemühungen dienen sollte.

5.2.3 Funktionalität und Genauigkeit

Die Ergebnisse hinsichtlich der Funktionalität sind nicht eindeutig. Obwohl die Genauigkeit mit verlängerter Laufzeit der Tests zunimmt, bleiben die Ergebnisse inkonsistent. Dies erschwert einen zuverlässigen Ersatz für die konventionelle Fahreridentifizierung und beeinträchtigt den Einsatz für die Insassen, da die Insassenerkennung nur während der Fahrt zuverlässige Ergebnisse liefert.

Um den Erkennungsprozess zu verbessern, sollten verschiedene Maßnahmen ergriffen werden. Unter anderem könnte mehr Aufwand in die Optimierung des ML-Modells investiert werden wie die Verfeinerung der Klassifizierungsalgorithmen durch den Einsatz von mehr Datenquellen oder erweiterte ML-Verfahren.

Darüber hinaus könnte die Aufnahme der Insassen in der Incar-Komponente verbessert werden. Wie aus den Versuchen hervorgeht, wird die Stimmerkennung durch laute Hintergrundgeräusche beeinträchtigt. Daher könnten Mechanismen zur Stimmisolierung implementiert werden, um irrelevante Frequenzen herauszufiltern und so die Erkennungsgenauigkeit zu verbessern.

Es ist wichtig zu beachten, dass die Bewertung und Verbesserung einer solchen Implementierung ein iterativer Prozess ist, der weitere Forschung und experimen-

telle Validierung erfordert, um die Leistungsfähigkeit des Systems zu maximieren.

5.2.4 Schlussfolgerung

Bei oberflächlicher Betrachtung erscheinen die Versuchsergebnisse hinsichtlich der Anwendbarkeit der Stimmerkennung zur Identifizierung von Fahrzeuginsassen negativ. Das Gesamtkonzept ist aufgrund seiner Komplexität und bezüglich der Genauigkeitsprobleme für die Umsetzung nicht geeignet. Zudem ist eine erhebliche Rechenleistung für das Training des Modells erforderlich und die Klassifikation dauert zu lange.

Dennoch ist festzuhalten, dass die Spracherkennung auch für andere Anwendungsfälle im Fahrzeug eingesetzt werden kann. Die vorliegende Implementierung kann bereits erfolgreich Personen identifizieren und die zugehörige Architektur ist für den Fahrzeugkontext geeignet. Um bessere Ergebnisse zu erzielen oder andere Anwendungsgebiete zu finden, ist jedoch weiterer Zeit- und Forschungsaufwand erforderlich.

6 | Diskussion und Ausblick

Im Rahmen der folgenden Diskussion soll erörtert werden, ob die Stimmerkennung als geeignete Methode zur Identifizierung von Fahrzeuginsassen eingesetzt werden kann. Dabei wird die Forschungsfrage beantwortet und ein Ausblick gegeben.

6.1 Akzeptanz der Nutzer

Die Akzeptanz der Stimmerkennung wird stark von Datenschutzbedenken beeinflusst. Die Nutzer sind zunehmend besorgt über die mögliche Verletzung ihrer Privatsphäre und die unsachgemäße Verarbeitung persönlicher Daten durch solche Systeme. Ein wichtiger Punkt ist das Bewusstsein der Nutzer, dass das System permanent aktiv ist und ihre Stimme aufzeichnet. Dies kann zu einem verstärkten Unsicherheitsgefühl führen, da das Gefühl einer ständigen Überwachung entsteht. Im Gegensatz dazu bieten Spracherkennungssysteme wie Alexa oder Siri ein gewisses Maß an Kontrolle, da sie nur aktiviert werden, wenn ein bestimmtes Schlüsselwort verwendet wird.

Darüber hinaus ist es wichtig zu betonen, dass Nutzer nur bereit sind, eine Stimmerkennung zu verwenden, wenn es einen erheblichen Mehrwert bietet. Wenn der Trainingsprozess zu langwierig ist oder der resultierende Nutzen zu gering erscheint, werden die meisten Nutzer voraussichtlich auf die Einrichtung des Systems verzichten. Der Aufwand, den die Nutzer in den Trainingsprozess investieren müssen, muss in einem angemessenen Verhältnis zu dem entstehenden Mehrwert liegen. Wird dieses Gleichgewicht nicht erreicht, kann dies zu einer geringen Ak-

zeptanz und Nutzung des Systems führen, selbst wenn Datenschutzbedenken ausgeräumt sind.

In der Zusammenfassung ist zu erkennen, dass die Akzeptanz und Nutzung von Stimmerkennung stark von verschiedenen Faktoren abhängig ist. Einerseits sind Datenschutzbedenken ein zentrales Anliegen der Nutzer und erfordern klare Richtlinien und Transparenz. Andererseits ist es entscheidend, dass sich die weitere Forschung auf die Schaffung eines signifikanten Mehrwerts für die Nutzer konzentriert, um die Akzeptanz zu fördern. Die derzeitige Umsetzung der Insassenidentifizierung mittels Stimmerkennung mag nicht die optimale Lösung darstellen, bietet aber als Proof of Concept Möglichkeiten für weitere Untersuchungen.

6.2 Einsatzgebiete

Im Folgenden werden die beiden wichtigsten Einsatzgebiete, die in der Einleitung beschrieben wurden, sowie mögliche neue Einsatzgebiete, die sich im Laufe der Arbeiten ergeben haben, hinsichtlich ihrer Umsetzung und ihrer Ergebnisse bewertet.

Wie die Versuche gezeigt haben, ist die Umsetzung der Insassenidentifizierung mittels Stimmerkennung in einem Familienfahrzeug vorteilhaft. Insbesondere in einem Szenario, in dem die Anzahl der möglichen Insassen überschaubar ist, erscheint diese Implementierung attraktiv. Eine lokale Lösung, bei der die Klassifizierung auf der im Fahrzeug verbauten Hardware stattfindet, könnte vorteilhaft sein, insbesondere wenn das Fahrzeug über eine leistungsfähige Hardware verfügt, wie sie beispielsweise bei autonomen Fahrsystemen zum Einsatz kommt.

Andererseits ist zu berücksichtigen, dass bei einer begrenzten Anzahl von Fahrgästen möglicherweise einfachere Lösungen zur Verfügung stehen. Eine manuelle Auswahl über ein Menü in der Headunit des Fahrzeugs oder über eine App könnte eine Alternative darstellen, die die Entwicklungskosten für die Stimmerkennung oder andere Identifizierungstechnologien möglicherweise nicht rechtfertigt.

Die Anwendung in Carsharing und Fahrzeugpools konnte im Rahmen der durchge-

fürten Versuche nicht umfassend behandelt werden. Es ist jedoch offensichtlich, dass der Nutzen hier besonders groß ist, vor allem für den Fahrer, der sich die Zeit sparen könnte, sich jedes Mal in ein geliehenes Fahrzeug einzurichten. In diesen Szenarien wäre die Verwendung eines Backends zur Synchronisierung der Sprecher zwischen mehreren Pool-Fahrzeugen unerlässlich.

Ein Nachteil der Insassenidentifizierung mittels Stimmerkennung ist, dass, wenn sich nur ein Fahrer im Fahrzeug befindet, die für diese Implementierung notwendigen natürlichen Gespräche fehlen. Dies gilt beispielsweise für Firmenfahrzeugpools, in denen sich häufig nur ein Insasse, der Fahrer, sich befindet.

Insgesamt ist es wichtig, die Implementierung sorgfältig auf die Einsatzgebiete abzustimmen. Während die Stimmerkennung in bestimmten Szenarien wie Familienfahrzeugen durchaus sinnvoll sein kann, müssen auch einfachere Alternativen in Betracht gezogen werden, insbesondere wenn sie kosteneffektiver und benutzerfreundlicher sind. Eine umfassende Bewertung der Technologien und ihrer Anwendbarkeit in realen Anwendungsszenarien ist daher unerlässlich.

6.3 Zusätzliche Einsatzgebiete

Die Identifizierung von Fahrzeuginsassen mit Hilfe von Stimmerkennung bietet auch Potenzial für weitere Einsatzgebiete. Ein wichtiges Anwendungsgebiet ist das Führen eines Fahrtenbuches, welches für den Fahrzeugbesitzer bzw. Fahrzeugvermieter nützlich sein kann. Auch für Versicherungsunternehmen könnte es von Bedeutung sein, zu verstehen, wer sich wo im Fahrzeug befand, wer gefahren ist und ob sich Kinder im Fahrzeug befanden. Diese Informationen, in Kombination mit anderen Datenquellen, könnten dazu beitragen, Konflikte im Straßenverkehr zu lösen. Die Implementierung ist dafür gut geeignet, da die Genauigkeit der Ergebnisse mit der Fahrzeit zunimmt.

Ein weiteres Einsatzgebiet könnte für OEMs von Interesse sein. In einer Welt, in der immer mehr Daten gesammelt werden, könnten Informationen darüber, wie die Kunden ihre Fahrzeuge nutzen, von Vorteil sein, um die Bedürfnisse

der Kunden besser zu verstehen und gezielte Anpassungen an den Fahrzeugen vornehmen zu können.

Selbstverständlich müssen diese Anwendungsbereiche im Hinblick auf Datenschutz und Sicherheit sorgfältig geprüft werden. Dennoch könnten sie von den Vorteilen der Stimmerkennung profitieren und neue Möglichkeiten zur Verbesserung von Fahrzeugen und Dienstleistungen eröffnen.

6.4 Beantwortung der Forschungsfrage

Die in Kapitel 1.4 formulierte Forschungsfrage kann nun mit Hilfe der prototypischen Umsetzung und der diskutierten Ergebnisse beantwortet werden. Die Implementierung der Stimmerkennung als Proof of Concept liefert zwar einen ersten Ansatz zur Insassenidentifizierung, stellt aber keine eindeutige Antwort auf die Forschungsfrage dar.

Datenschutzrechtliche Aspekte konnten im Rahmen der prototypischen Umsetzung nicht abschließend untersucht werden, jedoch wurden in Kapitel 5.2.1 Methoden diskutiert, wie datenschutzrechtlich relevante Themen ausgearbeitet werden können. Hinsichtlich der Benutzerfreundlichkeit konnte festgestellt werden, dass die Stimmerkennung als Mittel zur Insassenidentifizierung im Fahrzeug nicht ohne Einschränkungen geeignet ist, wie in der Auswertung der Versuche in Kapitel 5.2.2 beschrieben.

Es ist anzumerken, dass diese Arbeit einen breiten Themenbereich der Insassenidentifizierung abdeckt. Die Beantwortung der Forschungsfrage erfolgt durch die Analyse der implementierten Lösung und die darauf aufbauende Diskussion der Ergebnisse. Es wurden auch Rückschlüsse für die weitere Entwicklung der Technologie gezogen, wobei eine sinnvolle Eingrenzung der Themenbereiche möglicherweise weitere Entwicklungen bringen kann. Eine stärkere Fokussierung auf einzelne Komponenten der Insassenidentifizierung könnte für zukünftige Forschungsarbeiten von Vorteil sein, um detailliertere Einblicke und mögliche Fortschritte zu erhalten.

6.5 Fazit

Die vorliegende Arbeit hat sich umfassend mit der Thematik der Insassenidentifizierung beschäftigt. Zunächst wurde eine Marktanalyse durchgeführt, um die vorhandenen Technologien zu erfassen. Anschließend wurden verschiedene Technologien untersucht, die sich für die Insassenidentifizierung eignen. Die Stimmerkennung wurde als vielversprechender Ansatz ausgewählt und prototypisch implementiert. Die Implementierung wurde evaluiert und durch verschiedene Versuche validiert, um die verschiedenen Einsatzgebiete zu untersuchen. Abschließend wurden die Ergebnisse ausführlich diskutiert und bewertet.

Die Evaluation der Implementierung zeigt, dass die Stimmerkennung unter optimalen Bedingungen funktional ist. Allerdings nimmt die Genauigkeit der Identifizierung in Situationen mit Hintergrundmusik oder anderen Tonlagen der Sprecher deutlich ab. Dies führt zu relativ ungenauen Ergebnissen, die für eine zuverlässige Authentifizierung und insbesondere für die Freigabe personenbezogener Daten nicht ausreichend sind.

Die Ergebnisse dieser Arbeit betonen, dass trotz der nicht eindeutig positiven Implementierungsergebnisse, ein Potenzial für die Stimmerkennung als Methode zur Identifizierung von Insassen im Fahrzeugkontext besteht. Es bedarf jedoch weiterer Optimierungen, wie die Berücksichtigung von Umwelteinflüssen, um eine zuverlässige und einsatzfähige Lösung zu gewährleisten.

6.6 Ausblick und Zukünftige Arbeiten

Diese Arbeit stellt einen grundlegenden Ansatz dar, um die Insassenidentifizierungstechnologie weiterzuentwickeln und zu verbessern. Obwohl die Stimmerkennung nicht als optimale Lösung für dieses spezifische Einsatzgebiet identifiziert wurde, eröffnet sie dennoch interessante Möglichkeiten für zukünftige Forschungsansätze in der Industrie. Die prototypische Implementierung bietet wertvolle Einblicke in die Realisierbarkeit und Herausforderungen der Stimmerkennung im Fahr-

zeugkontext. Durch eine Weiterentwicklung und Verbesserung dieser Technologie können potenziell präzisere und zuverlässigere Lösungen entwickelt werden, die einen signifikanten Beitrag zur Sicherheit und Effizienz im Fahrzeugbereich leisten können. Weitere Forschungsarbeiten zur Insassenidentifizierung im Fahrzeugkontext könnten folgendermaßen beschrieben werden.

Eine mögliche Forschungsrichtung bezieht sich auf die Weiterentwicklung des verwendeten Modells von Fadi Badine aus der Keras-Dokumentation [39]. Obwohl das Modell funktionell ist, liefert es nicht immer nachvollziehbare Ergebnisse. Angesichts dieser Herausforderung wäre es interessant zu untersuchen, wie das ML-Modell speziell für diese Aufgabe optimiert werden kann.

In der Diskussion wurden diverse Ansätze für alternative Einsatzgebiete der Insassenidentifizierung im Fahrzeugkontext erörtert. Es wäre interessant zu untersuchen, wie diese Technologie auch in anderen Anwendungsbereichen eingesetzt werden könnte und welche neuen Geschäftsmodelle sich daraus ergeben könnten. Insbesondere aus Sicht von Versicherungsunternehmen könnten sich, wie bereits diskutiert, neue Möglichkeiten ergeben.

Eine weitere interessante Herausforderung wäre die Implementierung der Stimmerkennung auf der im Fahrzeug vorhandenen Hardware. Eine Untersuchung, inwieweit die Headunit im Fahrzeug für eine ähnliche Implementierung verwendet werden könnte, wäre von Interesse. Aus Gründen des Produktschutzes ist eine solche Aufgabe nur im OEM-Rahmen möglich.

Diese potenziellen Forschungsrichtungen könnten dazu beitragen, die Insassenidentifizierung im Fahrzeugkontext voranzutreiben und neue Möglichkeiten für die Anwendung und Entwicklung dieser Technologie aufzuzeigen.

Anhang 1: Bildschirmaufnahmen des Frontends

Lorem ipsum dolor sit amet, consectetur adipiscing elit

Lorem ipsum dolor sit amet, consectetur adipiscing elit

Lorem ipsum dolor sit amet, consectetur adipiscing elit

Lorem ipsum dolor sit amet, consectetur adipiscing elit

Anhang 2: UML Sequenzdiagramm: Identifizierungsprozess

Lorem ipsum dolor sit amet, consectetur adipiscing elit

Anhang 3: Modell Architektur

Lorem ipsum dolor sit amet, consectetur adipiscing elit

Anhang 4: Incar-Komponente Kalibrierung

Lorem ipsum dolor sit amet, consectetur adipiscing elit

Anhang 5: Verlust- und Genauigkeitsgrafiken

Lorem ipsum dolor sit amet, consectetur adipiscing elit

Lorem ipsum dolor sit amet, consectetur adipiscing elit

Anhang 6: Versuchsaufbau

Lorem ipsum dolor sit amet, consectetur adipiscing elit

Literatur

- [1] Prof. Dr.-Ing. G.-A. Ahrens, "Sonderauswertung zum forschungsprojekt „mobilität in städten“,“ Technische Universität Dresden, Fakultät Verkehrswissenschaften "Friedrich ListInstitut für Verkehrsplanung und Straßenverkehr, 2013. Accessed: Feb. 14, 2024. [Online]. Available: https://tu-dresden.de/bu/verkehr/ivs/srv/ressourcen/dateien/SrV2013_Staedtevergleich.pdf?lang=de
- [2] "CarSharing entlastet umwelt und verkehr." Accessed: Feb. 14, 2024. [Online]. Available: <https://carsharing.de/alles-ueber-carsharing/umweltbilanz/carsharing-entlastet-umwelt-verkehr>
- [3] "Identifizierung," *PONS Großwörterbuch Deutsch als Fremdsprache*. Accessed: Feb. 14, 2024. [Online]. Available: <https://de.thefreedictionary.com/Identifizierung>
- [4] "Authentisierung, authentifizierung und autorisierung," Perseus Technologies GmbH. Accessed: Feb. 14, 2024. [Online]. Available: <https://www.perseus.de/2022/04/19/authentisierung-authentifizierung-und-autorisierung/>
- [5] *Owner's manual, BMW i4 gran coupé*. BMW AG, 2024. Available: <https://www.bedienungsanleitu.ng/bmw/i4-grand-coupe-2024/anleitung>
- [6] Z. Somogyi, *The application of artificial intelligence - step-by-step guide from beginner to expert*. Singapore: Springer Nature, 2021.
- [7] "TensorFlow." Accessed: Mar. 01, 2024. [Online]. Available: <https://www.tensorflow.org/>
- [8] K. Jangla, *Accelerating development velocity using docker - docker across microservices*. New York: Apress, 2018.
- [9] "Wie kann ich die fahrerposition aus einem gastprofil zu meinem fahrerprofil speichern?" Accessed: Feb. 15, 2024. [Online]. Available: <https://faq.bmw.de/s/article/Fahrzeug-Sitze-Fahrerprofil-Gastprofil-im-Fahrerprofil-speichern-25YDx?language=de>
- [10] "Wie viele fahrer können sich mit ihrer BMW ID im fahrzeug anmelden?" Accessed: Feb. 15, 2024. [Online]. Available: <https://faq.bmw.de/s/article/BMW-iDrive-Personalisierung-BMW-ID-Anzahl-Zqrzn?language=de>

-
- [11] "Strategien und perspektiven für den größten automobilmarkt der welt." PrincewaterhouseCoopers, 2011. Accessed: Feb. 16, 2024. [Online]. Available: <https://www.pwc.de/de/automobilindustrie/assets/automobilindustrie-und-mobilitaet-in-china.pdf>
- [12] *Tesla model 3 owner's manual*. Tesla Inc., 2024. Available: https://www.tesla.com/ownersmanual/model3/en_us/GUID-A2D0403E-3DAC-4695-A4E6-DC875F4DEDC3.html
- [13] "Einfach in verbindung bleiben. Mit VW connect und we connect." Accessed: Mar. 18, 2024. [Online]. Available: https://www.volkswagen.de/de/konnektivitaet-und-mobilitaetsdienste/konnektivitaet/we-connect.html?adchan=sem&campaign=00010218_NC_IMC_BON_CON_VWB_VWB_BRD_DEU_NU_SEAH_PSEAC_TEAD_NU_GAD_GADW_CPC_NU_NU_KEW_NU_CXD_DE_AL_BRAD_%5BVolkswagen%5D_Connectivity&adgroup=We+Connect_%5BAMT%5D&publisher=GOOGLE&adcr=volkswagen+we+connect&adpl=GOOGLE&adlid=71700000081399336&country=DE&language=DE&gad_source=1&gclid=Cj0KCQjw2PSvBhDjARIsAKc2cgPVLqkHT9pP2v1pcC_RWs6T_2Zrz2rEflgqCXLloaPFvgLi7_ar6ggaAtTqEALw_wcB&gclsrc=aw.ds
- [14] "Die hand als türöffner." Accessed: Mar. 18, 2024. [Online]. Available: <https://www.welt.de/print-wams/article613492/Die-Hand-als-Tueroeffner.html>
- [15] "Übersicht fingerabdrucksensor." Accessed: Feb. 16, 2024. [Online]. Available: https://mba.i.mercedes-benz.com/baix/cars/223.0_mbox_2021_a/de_DE/page/ID_82581f232f25d91eac190d2b76cc7fe3-e547de532f25d91eac190d2b5ef50c2a-de-DE.html#
- [16] "Hyundai reveals world's first smart fingerprint technology to vehicle." Accessed: Mar. 22, 2024. [Online]. Available: <https://www.hyundai.news/eu/articles/press-releases/hyundai-reveals-worlds-first-smart-fingerprint-technology-to-vehicles.html>
- [17] "How to use fingerprint reader & face connect | genesis GV60 | genesis USA." Accessed: Mar. 22, 2024. [Online]. Available: <https://www.youtube.com/watch?v=ILQMebZtEiw>
- [18] "Subaru forester." Accessed: Feb. 16, 2024. [Online]. Available: <https://www.subaru.de/modelle/forester>
- [19] "US 2018/0208209 A1," 26.07.2018 Available: <https://ppubs.uspto.gov/dirsearch-public/print/downloadPdf/20180208209>
- [20] "US 2018/0210446 A1," 26.07.2018 Available: <https://ppubs.uspto.gov/dirsearch-public/print/downloadPdf/20180210446>
- [21] "Smartphones - anteil der nutzer in deutschland bis 2027." Accessed: Feb. 21, 2024. [Online]. Available: <https://de.statista.com/statistik/daten/studie/585883/umfrage/anteil-der-smartphone-nutzer-in-deutschland/>
- [22] "MyBMW app." Accessed: Feb. 15, 2024. [Online]. Available: <https://www.bmw.de/de/my-bmw-app/my-bmw-app.html>
-

-
- [23] "MINI app." Accessed: Feb. 15, 2024. [Online]. Available: https://www.mini.com/en_MS/home/products/mini-app.html
- [24] Y. Nirkin, I. Masi, A. Tuan Tran, T. Hassner, and G. Medioni, "Face segmentation, face swapping, and how they impact face recognition," in *Advances in computer vision and pattern recognition*, in Advances in computer vision and pattern recognition., Springer International Publishing, 2021.
- [25] "US 2019/0044723 A1," 07.02.2019 Available: <https://ppubs.uspto.gov/dirsearch-public/print/downloadPdf/20190044723>
- [26] M. R. Yaniv Taigman Ming Yang, "DeepFace: Closing the gap to human-level performance in face verification," *IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1701–1708, 2014.
- [27] A. Wilkens, "BMW 5er darf teilautonom fahren: Hände vom steuer bei 130 km/h," *heise online*, 2023, Available: <https://www.heise.de/news/BMW-5er-darf-teilautonom-fahren-Haende-vom-Steuer-bei-130-km-h-9187891.html>
- [28] "Azure / AI services / speech service / what is speaker recognition?" Accessed: Feb. 22, 2024. [Online]. Available: <https://learn.microsoft.com/en-us/azure/ai-services/speech-service/speaker-recognition-overview>
- [29] "Request recognized speaker contact information for use in your skill." Accessed: Feb. 22, 2024. [Online]. Available: <https://developer.amazon.com/en-US/docs/alexa/custom-skills/request-recognized-speaker-contact-information.html>
- [30] S. Tripathi and S. Bhatnagar, "Speaker recognition," in *2012 third international conference on computer and communication technology*, 2012, pp. 283–287. doi: 10.1109/ICCCT.2012.64.
- [31] J. Wendeberg, "Laufzeitbasierte schallortung mit unbekannten sender- und empfangenpositionen," Master's thesis, Albert-Ludwigs-Universität Freiburg, November 2009 (Rev. October 2012).
- [32] "Amazon-EC2-T4g-instances." Accessed: Feb. 26, 2024. [Online]. Available: <https://aws.amazon.com/de/ec2/instance-types/t4/>
- [33] "Nvidia SERIE GEFORCE® GTX 900." Accessed: Feb. 26, 2024. [Online]. Available: <https://www.nvidia.com/de-de/geforce/900-series/>
- [34] "Amazon S3." Accessed: Feb. 26, 2024. [Online]. Available: <https://www.nvidia.com/de-de/geforce/graphics-cards/30-series/>
- [35] "MediaStream recording API." Accessed: Mar. 03, 2024. [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/API/MediaStream_Recording_API
- [36] "Pydub." Accessed: Mar. 01, 2024. [Online]. Available: <https://github.com/jiaaro/pydub>
- [37] "FFmpeg." Accessed: Mar. 01, 2024. [Online]. Available: <https://ffmpeg.org/>
- [38] V. Silaparasetty, *Deep learning projects using TensorFlow 2 - neural network development with python and keras*. New York: Apress, 2020.
-

- [39] F. Badine, "Speaker recognition." Accessed: Mar. 01, 2024. [Online]. Available: https://keras.io/examples/audio/speaker_recognition_using_cnn/
 - [40] "Socket.io documentation." Accessed: Mar. 01, 2024. [Online]. Available: <https://socket.io/docs/v4/>
-