

# **KI-Unterstützung im Medikamentenmanagement mit Hilfe eines Augmented Reality Systems**

Zur Erlangung des akademischen Grades eines

## **DOKTORS DER INGENIEURWISSENSCHAFTEN (DR. -ING.)**

von der KIT-Fakultät für  
Elektrotechnik und Informationstechnik  
des Karlsruher Instituts für Technologie (KIT)

genehmigte

## **DISSERTATION**

von

**M.Sc. Jens Juhl**

Tag der mündlichen Prüfung: 24.03.2025

Hauptreferent:

Prof. Dr. rer. nat Wilhelm Stork

Korreferent:

Prof. Dr. rer. nat. Cornelius Neumann



# Kurzfassung

In der vorliegenden Arbeit wird die Entwicklung eines technischen Systems zur Unterstützung des manuellen Medikamentenrichtens vorgestellt. Vor dem Hintergrund des demografischen Wandels und der zunehmenden Herausforderungen im Pflegebereich wird ein System entwickelt, das mithilfe von Augmented Reality (AR) und Künstlicher Intelligenz (KI) den Medikamentenprozess sicherer und effizienter gestalten soll. Die zentrale Fragestellung dieser Arbeit lautet: Wie kann beim Medikamentenrichten die Fehleranzahl reduziert werden? Dabei wird die Hypothese untersucht, dass ein technisches System, das AR und KI integriert, die Fehlerquote signifikant verringern kann.

Das entwickelte System basiert auf einer Kombination aus einer Kamera, die den Arbeitsbereich erfasst, und einem Beamer, der visuelle Anweisungen auf die Arbeitsfläche projiziert. KI-Algorithmen, darunter YOLOv5 zur Objekterkennung und ResNet zur Klassifikation der Medikamente, ermöglichen die Echtzeitanalyse und Rückmeldung an die Pflegekraft. Ein lokaler Server stellt dabei sicher, dass alle Daten im eigenen Netzwerk verbleiben.

Die Ergebnisse der Evaluationen, die mit professionellen Pflegekräften und Laien durchgeführt werden, bestätigen die Wirksamkeit des Systems: Es zeigt sich eine signifikante Reduzierung der Fehlerquote und eine Steigerung der Effizienz. Die Anzahl korrekt sortierter Medikamente steigt bei professionellen Pflegekräften um etwa 32 %, während die Fehlerquote um fast 50 % reduziert werden kann. Die Benutzerfreundlichkeit und kognitive Entlastung werden von den Teilnehmern ebenfalls positiv bewertet.

Insgesamt zeigt diese Arbeit, dass der Einsatz moderner Technologien wie AR und KI das Potenzial hat, den Medikationsprozess sicherer und effizienter zu

gestalten. Die Forschungsergebnisse bestätigen die Hypothese und weisen darauf hin, dass das vorgestellte System einen bedeutenden Beitrag zur Verbesserung der Medikamentenverwaltung leisten kann.

## Abstract

This thesis focuses on the development of a technical system to support manual medication sorting. Against the backdrop of demographic change and increasing challenges in the healthcare sector, a system was developed that aims to make the medication process safer and more efficient through the use of Augmented Reality (AR) and Artificial Intelligence (AI). The central research question of this thesis is: How can the number of errors in medication sorting be reduced? The hypothesis being tested is that a technical system integrating AR and AI can significantly reduce the error rate.

The developed system is based on a combination of a camera, which captures the workspace, and a projector that displays visual instructions on the working surface. AI algorithms, including YOLOv5 for object detection and ResNet for medication classification, enable real-time analysis and feedback to the caregiver. The system was implemented in Python, and a local server ensures that all data remains within a known network.

The results of the evaluations conducted with professional caregivers and laypersons confirm the system's effectiveness: A significant reduction in the error rate and an increase in efficiency were observed. The number of correctly sorted medications increased by about 32% among professional caregivers, while the error rate was reduced by almost 50%. Participants also positively rated the system's user-friendliness and cognitive relief.

Overall, this thesis demonstrates that the use of modern technologies such as AR and AI has the potential to make the medication process safer and more efficient. The research findings confirm the hypothesis and suggest that the proposed system can make a significant contribution to improving medication management.



# Danksagung

Diese Dissertation entstand im Rahmen meiner Tätigkeit als wissenschaftlicher Mitarbeiter im Forschungsbereich Embedded Systems and Sensors Engineering (ESS) am FZI Forschungszentrum Informatik.

Meinen besonderen Dank möchte ich Prof. Dr. rer. nat. Wilhelm Stork für die Möglichkeit zur Promotion und die Betreuung dieser Arbeit aussprechen. Sein Vertrauen, seine Ermutigungen zur praktischen Anwendung der neuesten wissenschaftlichen Erkenntnisse und der Freiraum, den er mir während der Durchführung gewährte, haben nicht nur diese Arbeit, sondern auch mich persönlich nachhaltig geprägt. Ebenso bedanke ich mich bei Prof. Dr. rer. nat. Cornelius Neumann für die Übernahme des Korreferats und das damit verbundene Engagement.

Für fachliche Diskussionen und tatkräftige Hilfestellung danke ich insbesondere meinen Kollegen Christoph Zimmermann, Marc Schroth und Friedrich Gauger. Darüber hinaus bedanke ich mich bei meinen Kollegen Oliver Werthwein, Julia Hofmann, Lara Schweickert, Dennis Birkenmaier und Gergely Biri sowie allen von mir betreuten Studenten für die fantastischen Jahre der Zusammenarbeit.

Mein ganz besonderer Dank gilt Herrn Korbinian Rudolf. Er hat weit über seine Rolle als wissenschaftliche Hilfskraft hinaus Beiträge zu dieser Arbeit geleistet und ist dabei zu einem echten Freund geworden. Mein Dank gilt auch Herrn Prof. Dr. Günter Langhammer und Herrn Heiko Maier für ihren fachlichen und nicht-fachlichen Beistand. Zu guter Letzt danke ich Frau Patricia Langhammer für ihre Liebe, ihre Unterstützung und ihr Verständnis bei meinem Promotionsvorhaben.

Danke an alle, die zur Entstehung dieser Arbeit beigetragen haben.

Karlsruhe, im Oktober 2024

*Jens Juhl*



# Inhaltsverzeichnis

<b>Kurzfassung</b> . . . . .	<b>i</b>
<b>Abstract</b> . . . . .	<b>ii</b>
<b>Danksagung</b> . . . . .	<b>iii</b>
<b>Abkürzungen und Symbole</b> . . . . .	<b>ix</b>
<b>1 Einleitung</b> . . . . .	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Umfeld und Hintergrund . . . . .	3
1.3 Ziel der Arbeit . . . . .	4
1.4 Aufbau des Dokuments . . . . .	5
<b>2 Grundlagen</b> . . . . .	<b>7</b>
2.1 Extended Reality . . . . .	7
2.2 Künstliche Intelligenz . . . . .	8
2.2.1 Neuronale Netze . . . . .	8
2.2.2 Architekturen zur Klassifikation . . . . .	18
2.2.3 Architekturen zur Detektion . . . . .	22
2.3 Psychologische Grundlagen . . . . .	28
2.3.1 Kognitive Belastungstheorie . . . . .	28
2.3.2 Dual-Prozess-Theorie . . . . .	29
<b>3 Stand der Technik</b> . . . . .	<b>31</b>
3.1 Aktuelles Vorgehen beim Medikamentenmanagement . . . . .	31
3.1.1 Händisches Einsortieren . . . . .	32

3.1.2	Maschinelles Verfahren . . . . .	33
3.2	Stand der Wissenschaft . . . . .	34
3.2.1	KI im Medikamentenmanagement . . . . .	34
3.2.2	Konzepte zum Medikamentenmanagement . . . . .	36
3.3	Forschungsfrage . . . . .	38
<b>4</b>	<b>Konzept . . . . .</b>	<b>39</b>
4.1	Anforderungserhebung . . . . .	39
4.1.1	Interessengruppen . . . . .	39
4.1.2	Anwendungsfälle . . . . .	40
4.1.3	Formulierung der Anforderungen . . . . .	43
4.1.4	Anforderungen . . . . .	45
4.2	Arbeitsablauf . . . . .	47
4.3	Systementwurf . . . . .	49
4.3.1	Bestandteile und Einflussgrößen . . . . .	51
4.4	Repräsentation . . . . .	52
4.4.1	Informationsvermittlung . . . . .	53
4.4.2	Medikamentenauswahl . . . . .	54
<b>5</b>	<b>Prototyp . . . . .</b>	<b>57</b>
5.1	Serverhardware . . . . .	57
5.2	KI-Modellauswahl . . . . .	57
5.2.1	Vergleich verschiedener Architekturen . . . . .	58
5.2.2	Test im Medikamentenmanagement . . . . .	59
5.2.3	Training von YOLOv5 . . . . .	61
5.3	AR-Brille . . . . .	63
5.3.1	Technische Vorteile der HoloLens . . . . .	63
5.3.2	Lokale Verarbeitung der KI-Modelle . . . . .	64
5.3.3	Entwicklungsprozess der Nutzer-Anwendung . . . . .	64
5.4	Server . . . . .	66
5.4.1	Aufbau und Funktion des Servers . . . . .	66
5.4.2	Datenfluss und -verarbeitung . . . . .	67
5.5	Technische Auswertung . . . . .	68
5.6	Zwischenfazit . . . . .	69
<b>6</b>	<b>Technische Entwicklung . . . . .</b>	<b>71</b>

6.1	Hardwareumbau . . . . .	71
6.1.1	Kameraplatzierung . . . . .	72
6.2	Koordinatentransformation . . . . .	73
6.2.1	Homographie Berechnen . . . . .	74
6.2.2	Tischgrößenbestimmung . . . . .	76
6.3	KI Weiterentwicklung . . . . .	77
6.3.1	Training von ResNet . . . . .	78
6.3.2	t-SNE Auswertung . . . . .	80
6.3.3	Training von YOLOv5 . . . . .	82
6.4	Handerkennung . . . . .	83
6.5	Nutzerschnittstelle . . . . .	84
6.6	Gesamtübersicht . . . . .	86
<b>7</b>	<b>Nutzerevaluation . . . . .</b>	<b>89</b>
7.1	Methodik der Evaluation . . . . .	89
7.1.1	Versuchsablauf . . . . .	89
7.1.2	Fragebögen Auswahl . . . . .	91
7.2	Zwischenevaluation . . . . .	94
7.3	Endevaluation . . . . .	95
7.3.1	Professionelle Pflegekräfte . . . . .	95
7.3.2	Laien in der Pflege . . . . .	97
7.3.3	Gesamtergebnisse . . . . .	98
7.4	2-Stunden-Evaluation . . . . .	98
7.5	Fragebögen Antworten . . . . .	100
7.5.1	System Usability Scale . . . . .	100
7.5.2	NASA Task Load Index . . . . .	102
7.5.3	Freitextrückmeldungen . . . . .	104
<b>8</b>	<b>Ergebnisse . . . . .</b>	<b>105</b>
8.1	Skalierbarkeit . . . . .	105
8.2	Überprüfen der Anforderungen . . . . .	106
8.3	Interpretation der Nutzerumfrage . . . . .	107
8.3.1	Zeitabhängigkeit der Probandenleistung . . . . .	108
8.3.2	Interpretation der Rückmeldungen . . . . .	110
8.4	Nächste Schritte . . . . .	111

8.4.1	Mögliche Weiterentwicklungen . . . . .	111
8.4.2	Verwertungsplan . . . . .	112
8.4.3	Wirtschaftlichkeit . . . . .	113
<b>9</b>	<b>Zusammenfassung . . . . .</b>	<b>115</b>
9.1	Ergebniszusammenfassung . . . . .	116
9.2	Fazit . . . . .	118
<b>A</b>	<b>Anhang . . . . .</b>	<b>119</b>
A.1	Freitext Rückmeldungen . . . . .	119
	<b>Abbildungsverzeichnis . . . . .</b>	<b>125</b>
	<b>Tabellenverzeichnis . . . . .</b>	<b>129</b>
	<b>Literaturverzeichnis . . . . .</b>	<b>131</b>
	Eigene Veröffentlichungen . . . . .	137

# Abkürzungen und Symbole

## Mathematische Symbole

$n$	Natürliche Anzahl
$\mathbf{h}$	Fette Variablen sind Vektoren
$H$	Großbuchstaben sind Matrizen
$a_n$	(n-ter) Eingang eines Systems
$b_n$	(n-ter) Ausgang eines Systems
$P(x,y)$	Punkt in einem Koordinatensystem
$mAP$	Mittlere Durchschnittliche Präzision

## Schichten in neuronalen Netzen

FcL	Vollverbundene Schicht
PooL	Pooling-Schicht
BnL	Normalisierungsschicht
CspL	Queerphasen-Teilschicht
ConvL	Faltungsschicht

## Etablierte Abkürzungen

NN	Neuronales Netz
fps	Frames pro Sekunde
OSPAN	Operation Span Task
SUS	System Usability Scale
NASA-TLX	NASA Task Load Index

## Eigene Abkürzungen

MMS	Medikamentenmanagement System
PMP	Papiermedikamentenplan
NOK	Nutzeroberflächen-Koordinatensystem
KAK	Kamera-Koordinatensystem



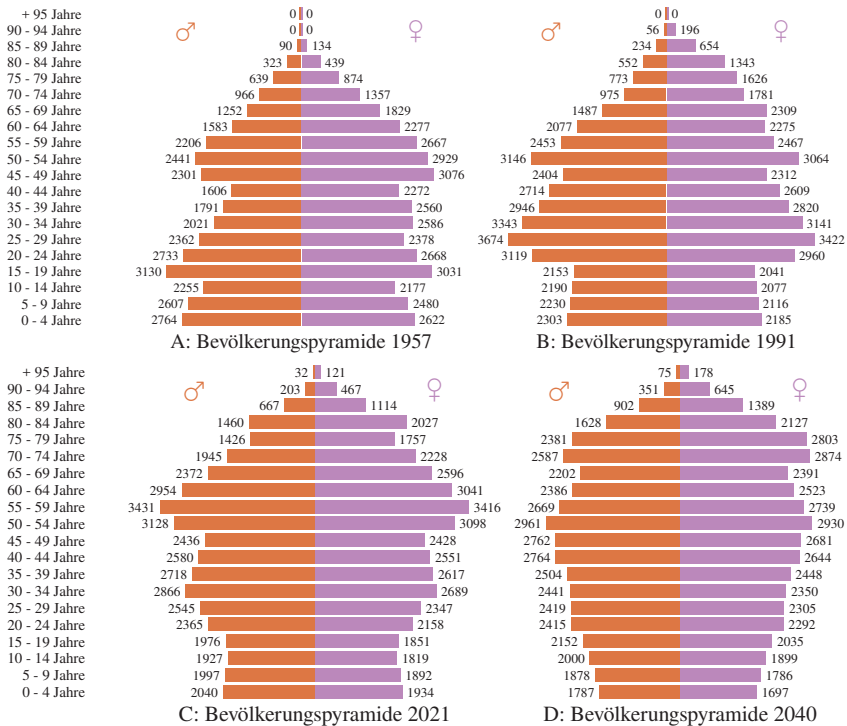


# 1 Einleitung

## 1.1 Motivation

Das Durchschnittsalter in Deutschland und den meisten anderen Industrieländern steigt jedes Jahr weiter an, da die Menschen immer älter werden und immer weniger junge Menschen geboren werden. Abbildung 1.1 zeigt die Entwicklung der Bevölkerung über die letzten Jahrzehnte, sowie die Extrapolation der Bevölkerungszahlen in der nahen Zukunft. Die durchschnittliche Kinderzahl pro Frau liegt seit den 80er Jahren deutlich unter der in der EU notwendigen Reproduktionsrate von 2,1 [1].

Diese demografische Veränderung hat nicht nur soziale, sondern auch finanzielle Auswirkungen auf das deutsche Sozialsystem. Ein prominentes Beispiel ist der deutsche soziale Generationenvertrag in Form der gesetzlichen Rentenversicherung. Dessen grundlegende Struktur ist seit der Rentenreform von 1957 gleich geblieben [2]. Zu Beginn im Jahr 1957 lag der sogenannte Altenquotient, also das Verhältnis zwischen Menschen über dem Renteneintrittsalter und jenen im erwerbsfähigen Alter, bei ca. 1:7 (siehe Abbildung 1.1 A). Im Jahr 1991 sank dieses Verhältnis auf 1:5 (siehe Abbildung 1.1 B). Bei den letzten veröffentlichten Daten im Jahr 2021 hatte es sich auf 1:3 fast halbiert (siehe Abbildung 1.1 C). Das bedeutet, dass drei Personen im erwerbsfähigen Alter für einen Rentenempfänger aufkommen müssen. Außerdem zeigen die Projektionen des statistischen Bundesamtes, dass das Verhältnis noch in den 40er Jahren dieses Jahrhunderts auf 1:2 sinken wird, unter der Annahme, dass jede Frau 1,44 Kinder bekommt und eine jährliche Netto-Zuwanderung von 290.000 Menschen nach Deutschland erfolgt (siehe Abbildung 1.1 D).



**Abbildung 1.1** Die deutsche Bevölkerung in Tausend gruppiert in jeweils fünf Jahre pro Balken. A-D zeigt die Bevölkerungsverteilung in den Jahren 1957 (A), 1991 (B), 2021 (C) und die vorhergesagte Bevölkerungsverteilung im Jahr 2040 (D) [3].

Jedoch bringt der demografische Wandel nicht nur finanzielle Herausforderungen mit sich, sondern beeinflusst auch die sozialen Strukturen. Die schrumpfende Anzahl an Arbeitsfähigen steht vor der Herausforderung, eine wachsende Anzahl von Senioren zu versorgen. Dies bedeutet auch eine personelle Herausforderung im Bereich der Pflege. Unabhängig davon, ob es sich um professionelle Pflegekräfte oder um Laien wie Angehörige handelt, stehen immer weniger Pflegenden pro Pflegebedürftigen zur Verfügung. Diese Entwicklung führt dazu, dass der Beruf

der professionellen Pflegekraft aufgrund der wachsenden Belastung an Attraktivität verliert. Dies erhöht in einem sich selbst verstärkenden Kreis wiederum die Belastung der verbleibenden Pflegekräfte.

Migranten, die oft auf der Suche nach Arbeitsmöglichkeiten sind, bieten eine wertvolle Ressource, um diese Lücke zu verkleinern oder ganz zu schließen. Allerdings bringt die Integration von Migranten in das Gesundheitswesen bei gleichzeitiger Gewährleistung einer hochwertigen Pflege für eine diversifizierte Bevölkerung auch zusätzliche Herausforderungen mit sich. Kulturelle Unterschiede, Sprachbarrieren und besondere Bedürfnisse der Pflegebedürftigen erfordern eine Anpassung von Pflegepraktiken im Pflegealltag.

Die Bewältigung dieser Herausforderungen erfordert nicht nur eine Anpassung der Arbeitsweise und Schulung der Pflegekräfte, sondern auch eine verstärkte Implementierung von Technologien, die mit solchen Schwierigkeiten umgehen können. Daher werden Pflegende im Allgemeinen und professionelle Pflegekräfte im Besonderen zunehmend technologische Unterstützung benötigen, um das steigende Arbeitspensum bewältigen zu können. Es ist von entscheidender Bedeutung, sicherzustellen, dass die Entwicklung solcher Technologien die realen Bedürfnisse der Pflegenden nicht aus den Augen verliert. Eine erfolgreiche Integration in den Pflegealltag sollte nicht nur auf Effizienz und Arbeitsentlastung abzielen, sondern auch die Anforderungen und Herausforderungen der Individuen im Pflegealltag berücksichtigen.

## **1.2 Umfeld und Hintergrund**

Die vorliegende Arbeit entstand im Kontext des Projekts »IDial - Interaktives Dialogsystem zur Verbesserung der Pflegedokumentation und Unterstützung bei Pflegeprozessen«. Dieses Projekt wurde von Juli 2020 bis Juni 2022 öffentlich durch das Ministerium für Soziales, Gesundheit und Integration des Landes Baden-Württemberg gefördert. Das übergeordnete Ziel des Projekts IDial bestand in der

Steigerung der Effizienz der Pflegedokumentation und der Optimierung des Pflegeprozesses für Pflegekräfte. Um dieses Ziel zu erreichen, wurden Technologien wie Sprachsteuerung und Augmented Reality eingesetzt, um die Tätigkeiten der Pflegekräfte interaktiv zu erfassen. Im Verlauf des Projekts konnte die Expertise und das Feedback von professionellen Pflegekräften in verschiedenen Phasen des Projekts in Anspruch genommen werden. Dies wurde durch den Besuch von Pflegekongressen und die aktive Mitarbeit von Projektpartnern im Pflegekontext ermöglicht. Im Zuge dieser Zusammenarbeit wurden im Projekt drei zentrale Themen herausgearbeitet. So wurden vor allem die Versorgung von chronischen Wunden, das Anrichten von Medikamenten und das Windelwechseln aufgrund von Inkontinenz herausgearbeitet. Innerhalb des Projekts wurden daraufhin das Medikamentenmanagement und das Wundmanagement weiter verfolgt. Nach Ablauf der Projektlaufzeit wurde das System ohne direkte Projektunterstützung weiterentwickelt.

### 1.3 Ziel der Arbeit

In Deutschland sind im Jahr 2023 3,4 Millionen Menschen pflegebedürftig, und jährlich sterben etwa 30.000 Menschen an Fehlmedikation [4]. Die Fehlerquote beim manuellen Richten der Medikamente liegt je nach Institution bei bis zu 10% der gestellten Medikamente [5]. Diese Situation unterstreicht die dringende Notwendigkeit für innovative Lösungen im Medikamentenmanagement. Angesichts dieser Tatsache und des demografischen Wandels und der wachsenden Herausforderungen im Pflegebereich, die in Kapitel 1.1 diskutiert wurden, zielt diese Arbeit darauf ab, die Medikamentenverwaltung effizienter, sicherer und benutzerfreundlicher zu gestalten. Dabei stehen vor allem technische Lösungen und deren Implementierung im Fokus. Außerdem soll nicht nur die Arbeitslast für Pflegekräfte reduziert, sondern auch die Sicherheit und Qualität der Versorgung für Pflegebedürftige erhöht werden. Dies wird durch eine intensive Kooperation mit professionellen Pflegefachkräften umgesetzt, um von Anfang an potenzielle Ansätze auf ihre Praktikabilität hin zu bewerten.

Grundlagenforschung und die Entwicklung komplett neuartiger Algorithmen werden nicht betrachtet, da eine solche Entwicklung den Zeithorizont dieser Arbeit übersteigen würde. Die Untersuchung wird sich auf die technischen Ansätze konzentrieren, deren Funktion bereits in der Grundlagenforschung gezeigt wurde. Deren Potential wird für einen Einsatz in der pflegerischen Arbeit bewertet und die vielversprechendsten werden nach festgelegten Kriterien ausgewählt. Aufbauend darauf werden eigene Konzepte entwickelt, die gezielt auf die Anforderungen im Medikamentenmanagement zugeschnitten sind.

Neben der technischen Machbarkeit, ist vor allem die Akzeptanz der Technologien durch Pflegekräfte von Bedeutung. Hierzu sollen Pflegekräfte in den Entwicklungsprozess integriert werden, um regelmäßige Rückmeldungen und praxisnahe Testläufe zu ermöglichen. In simulierten Szenarien soll außerdem die Benutzerfreundlichkeit und Effizienz der Technologie überprüft werden.

Des Weiteren sollen die entwickelten Lösungen im Vergleich zu herkömmlichen Methoden bewertet werden, um ihre potenziellen Vorteile hinsichtlich Zeitersparnis bei gleichbleibender oder verbesserter Qualität zu quantifizieren.

Abschließend werden die Ergebnisse der Tests und Evaluationen genutzt, um Empfehlungen für die Implementierung der Technologien in die Pflegepraxis abzuleiten, sodass eine reibungslose Integration der Technologie in den Pflegealltag gewährleistet werden kann.

## **1.4 Aufbau des Dokuments**

Diese Arbeit ist in neun Kapitel unterteilt. Nachdem hier in Kapitel 1 die Bedeutung dieser Arbeit in einen größeren Kontext gesetzt wurde, wird in Kapitel 2 zunächst auf die notwendigen technischen und psychologischen Grundlagen eingegangen, die zum Verständnis der Arbeit notwendig sind. Der aktuelle Stand der Wissenschaft und Technik wird im Kapitel 3 dargelegt. Dabei werden auch aktuelle im Pflegealltag übliche Verfahren im Medikamentenmanagement beschrieben

und es wird auf die Vor- und Nachteile der Nutzung von Augmented Reality eingegangen. Daraus leiten sich anschließend die zentralen Forschungsfragen ab, die im Zuge dieser Arbeit beantwortet werden. In Kapitel 4 werden dazu zunächst die konkreten Randbedingungen herausgearbeitet und unter Berücksichtigung dieser ein Konzept entwickelt. In Kapitel 5 wird die Entwicklung des ersten Prototyps vorgestellt, der auf einer Augmented-Reality-Brille basiert. Des Weiteren werden die Details, die zur Verwendung der spezifischen neuronalen Netz-Architekturen geführt haben, gezeigt. In Kapitel 6 wird dann die technische Entwicklung des Gesamt-Systems vorgestellt. Dieses basiert auf einer Kombination von Kamera und Beamer. In Kapitel 7 werden Prototyp und Gesamt-Systeme durch Nutzer evaluiert. Diese Evaluation wird zusammen mit der technischen Auswertung in Kapitel 8 weiter erörtert und mögliche Weiterentwicklungen werden dargelegt. Die Arbeit schließt in Kapitel 9 mit einer Zusammenfassung der wichtigsten Ergebnisse und einem Fazit zur Zukunft des Medikamentenmanagements.

## 2 Grundlagen

### 2.1 Extended Reality

Augmented Reality, Mixed Reality und Virtual Reality bilden ein Spektrum sich überschneidender Konzepte (siehe Abbildung 2.1). Alle drei Konzepte lassen sich gemeinsam unter dem Überbegriff der Extended Reality (XR) zusammenfassen.

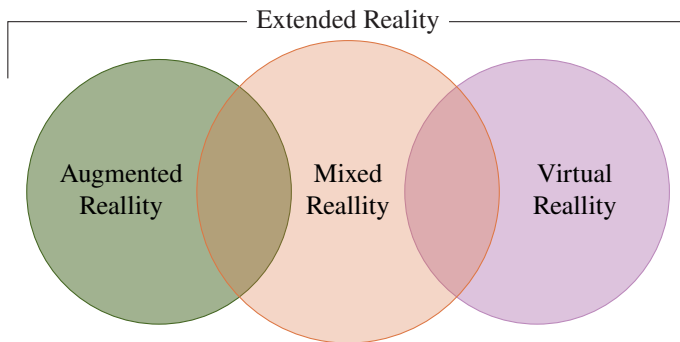


Abbildung 2.1 Übersicht über das Spektrum von Extended Reality

**Augmented Reality (AR)** erweitert die reale Umgebung um virtuelle Aspekte, indem es die physische Welt in Echtzeit mit digitalen Inhalten oder Bildern überlagert. Diese Technik bringt eine computergenerierte Überlagerung direkt ins Blickfeld des Benutzers und ermöglicht so eine nahtlose Integration zusätzlicher Informationen. Zum Beispiel kann ein AR-Headset dem Benutzer Informationen über die Umgebung anzeigen, wie Erklärungen zu konkreten Objekten im Sichtfeld.

**Virtual Reality (VR)** hingegen schafft vollständig immersive virtuelle Umgebungen, die den Benutzer von der physischen Realität isolieren und in eine vollständig digitale Welt eintauchen lassen. Bei der Nutzung von VR trägt der Benutzer ein Headset, das die reale Welt vollständig ausblendet und ihn stattdessen in eine computergenerierte Umgebung versetzt. Ein Beispiel hierfür wäre ein VR-Spiel, bei dem der Benutzer sich in einer komplett erfundenen Welt bewegt und interagiert.

**Mixed Reality (MR)** positioniert sich zwischen diesen beiden Extremen und stellt eine Verschmelzung dar, bei der digitale und reale Objekte unmittelbar miteinander in Wechselwirkung treten. Im Gegensatz zu AR, wo die reale Welt lediglich ergänzt wird, dient bei MR die Umgebung als Eingabe, auf die digitale Objekte dynamisch reagieren. Ein Beispiel für MR wäre ein System, bei dem ein Benutzer reale Objekte manipulieren kann, die dann in Echtzeit von digitalen Elementen ergänzt oder verändert werden, wie etwa eine Anwendung, die dem Nutzer in Echtzeit Rückmeldung zu seiner aktuellen Handlung gibt.

## 2.2 Künstliche Intelligenz

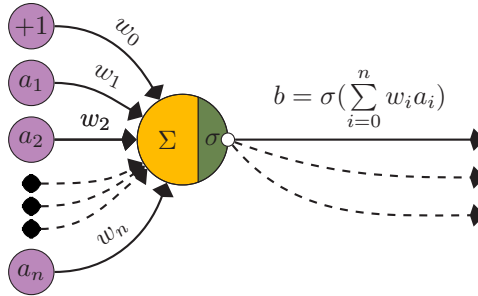
Künstliche Intelligenz (KI) ist ein Bereich der Informatik, der sich mit der Entwicklung und dem Einsatz von Maschinen und Software befasst, die in der Lage sind, Aufgaben zu erledigen, die normalerweise menschliche Intelligenz erfordern. Dazu gehören Lern- und Anpassungsfähigkeiten, Problemlösung, Sprachverstehen und Entscheidungsfindung. KI-Systeme bestehen oft aus Algorithmen, die durch Daten trainiert werden, um Muster zu erkennen und daraus Schlussfolgerungen zu ziehen.

### 2.2.1 Neuronale Netze

Künstliche Neuronen bilden den Grundbaustein für neuronale Netze (siehe Abbildung 2.2). Dabei sind  $a_1$  bis  $a_n$  die sogenannten Eingänge des Neurons. Die Eingänge können dabei die Ausgänge von anderen Neuronen sein, oder aus einer



anderen Quelle stammen. Jeder Eingang  $a_1$  bis  $a_n$  wird mit einem Gewicht  $w_1$  bis  $w_n$  versehen, wobei der Grundversatz (englisch Bias)  $w_0$  einen Eingangswert von  $+1$  bekommt. Jeder Eingang wird mit seinem entsprechenden Gewicht multipliziert und anschließend aufaddiert.



**Abbildung 2.2** Illustration eines künstlichen Neurons

Da es sich hierbei aber nur um lineare Operationen handelt, kann ein solches Neuron nur lineare Zusammenhänge abbilden, was sich auch auf ein ganzes Netzwerk aus solchen Neuronen überträgt. Darum wird diese Summe durch eine sogenannte Aktivierungsfunktion  $\sigma(a)$  nicht-linear abgebildet. Das Ergebnis der Aktivierungsfunktion erzeugt so das Ausgangssignal  $b$  des Neurons, welches dann wiederum an beliebig viele Neuronen weitergereicht werden kann.

Jede Aktivierungsfunktion hat spezifische Eigenschaften, die sie für bestimmte Arten von Anwendungsfällen geeignet machen. Außerdem ist wichtig, dass bei jeder Aktivierungsfunktion die Ableitung eindeutig definiert ist (mehr dazu in Kapitel 2.2.1.2):

$$\sigma(a) = \frac{e^a}{1 + e^a}, \quad \sigma'(a) = \frac{e^a}{1 + e^a} * \frac{1}{1 + e^a} \quad (2.1)$$

Die Ausgabe der Sigmoidfunktion (siehe Abbildung 2.3 A) hat einen Wertebereich zwischen 0 und 1. Daher kann die Ausgabe in vielen Anwendungsfällen als Wahrscheinlichkeit interpretiert werden.

$$\sigma(a) = \tanh(a), \quad \sigma'(a) = 1 - \tanh^2(a) \quad (2.2)$$

Der Tangens Hyperbolicus (siehe Abbildung 2.3 B) nimmt Werte zwischen -1 und 1 an und zentriert die Ausgaben um Null. Dies führt zu einem ähnlichen Verlauf wie die Sigmoid Funktion, aber erlaubt auch die Ausgabe von negativen Werten.

$$\sigma(a) = \max(0, a), \quad \sigma'(a) = 1 \quad (2.3)$$

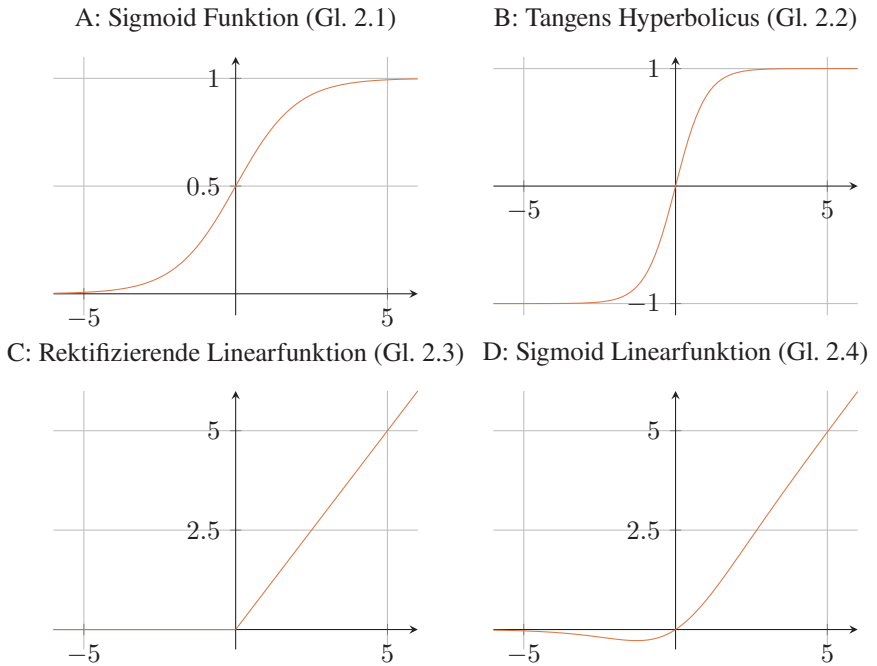
Die Rektifizierende Linearfunktion (englisch rectified linear unit) bietet eine sehr effiziente und effektive Art der nicht-linearen Aktivierung, indem sie alle negativen Eingänge auf null setzt und positive Eingänge unverändert lässt (siehe Abbildung 2.3 C). Diese Funktion wird in tiefen neuronalen Netzwerken verwendet, da sie sehr wenig Rechenleistung benötigt und gleichzeitig für große Eingaben in der Ableitung nicht gegen null geht.

$$\sigma(a) = \frac{a}{1 + e^{-a}}, \quad \sigma'(a) = \frac{1 + (1 + a) * e^{-a}}{(1 + e^{-a})^2} \quad (2.4)$$

Die Sigmoid Linearfunktion (englisch sigmoid linear unit) ist eine Kombination aus der Sigmoid Funktion und der Rektifizierenden Linearfunktion (siehe Abbildung 2.3 D) und erweitert die Vorteile der Rektifizierenden Linearfunktion durch einen vergrößerten Arbeitsbereich bei kleinen negativen Ausgaben, da die Ableitung nicht null ist.

### 2.2.1.1 Der Aufbau von neuronalen Netzen

Neuronale Netzwerke sind von der Struktur des menschlichen Gehirns inspirierte Algorithmen, die aus einer großen Anzahl von Neuronen bestehen. Dazu sind neuronale Netzwerke in Schichten organisiert, wobei jede Schicht spezifische



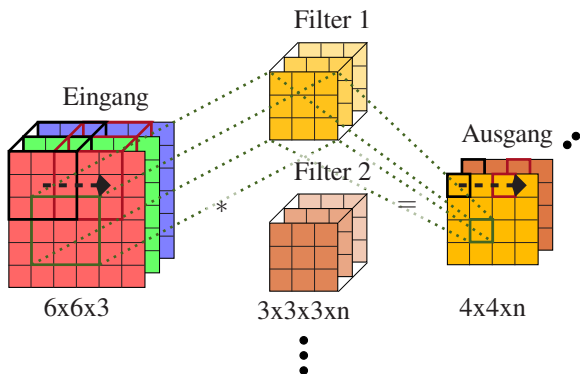
**Abbildung 2.3** Die häufigsten Aktivierungsfunktionen  $\sigma(a)$

Aufgaben erfüllt. Dabei sind Neuronen einer Schicht nicht direkt mit anderen Neuronen der gleichen Schicht verbunden, können sich jedoch die gleichen Eingänge und Ausgänge teilen. Grundsätzlich lässt sich ein Netzwerk in Eingabeschicht, verborgene Schichten und Ausgabeschicht unterteilen. Die Eingabeschicht nimmt die rohen Daten auf, die durch das Netzwerk fließen, während die verborgenen Schichten die Verarbeitung übernehmen und die Ausgabeschicht die Endvorhersage oder -klassifikation liefert. Neuronale Netzwerke sind typischerweise aus den folgenden Schichten aufgebaut:

**Voll-verbundene Schicht:** Voll-verbundene Schichten (englisch Fully connected Layer oder FcL) repräsentieren einen der grundlegenden Bausteine in einem neuronalen Netzwerk. In diesen Schichten sind alle Eingänge aus der vorherigen

Schicht mit jedem Neuron in der aktuellen Schicht verbunden. Diese umfassende Vernetzung ermöglicht es der Schicht, komplexe Muster und Beziehungen in den Daten zu lernen, indem sie auf den erlernten Merkmalen (englisch feature) aus den vorausgegangenen Schichten aufbaut. Die vollständig verbundenen Schichten sind besonders wichtig für die Klassifikations- und Regressionsaufgaben am Ende eines neuronalen Netzwerks, da ihre Ausgänge als Wahrscheinlichkeit interpretiert werden können.

**Faltungsschicht:** Die Faltungsschicht (englisch Convolutional Layer oder ConvL) ist das Herzstück eines faltungsneuronalen Netzwerks. Sie führt die Faltungsoperation durch, bei der mindestens ein Filter (englisch kernel) über das Eingabebild bewegt wird, um Merkmale wie Kanten, Ecken oder Texturen zu erkennen (siehe Abbildung 2.4). Dabei erzeugt der gelbe Filter (Filter 1) die gelbe Merkmalskarte (englisch feature map) am Ausgang, etc. Jeder Filter kann dabei auf das Erkennen spezifischer Merkmale im Bild spezialisiert werden. Durch die Anwendung mehrerer Filter können gleichzeitig verschiedene Merkmale eines Bildes erlernt werden, wobei sich die Anzahl an Merkmalskarten nach der Faltungsschicht auf die Anzahl ( $n$ ) der verwendeten Filter erhöht.



**Abbildung 2.4** Funktion einer Faltungsschicht: Jeder Filter wird schrittweise über den Eingang gezogen und erzeugt bei jedem Schritt einen Wert im Ausgang.

**Querphasen-Teilschicht:** Die Querphasen-Teilschicht (englisch Cross stage partial Layer oder CspL) ist eine moderne Weiterentwicklung der Faltungsschicht und verarbeitet die Merkmalskarten, die von einer vorgeordneten Schicht erzeugt werden, auf zwei Wegen. Dafür wird eine Kopie der Merkmalskarten direkt an nachfolgende Schichten weitergeleitet, während auf einem zweiten Weg eine Verarbeitung durch eine Reihe von Faltungsschichten durchgeführt wird (siehe Abbildung 2.5). Dabei ist sowohl die genaue Anzahl der ConvL als auch die Verwendung möglicher Aktivierungsfunktionen oder anderer Schichten flexibel. Diese Teilung ermöglicht es, dass sowohl originale als auch transformierte Merkmalskarten gemeinsam zur Informationsverarbeitung genutzt werden, was zum einen zu einer effizienteren Merkmalsnutzung führt und zum anderen dem Problem des verschwindenden Gradienten (vergleiche Kapitel 2.2.1.3) entgegenwirkt.

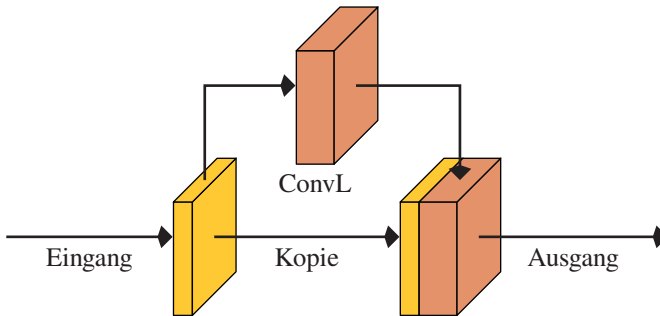


Abbildung 2.5 Aufbau der Querphasen-Teilschicht

**Pooling-Schicht:** Pooling-Schichten (englisch Pooling Layer oder PoolL) reduzieren die räumliche Dimension der Merkmalskarten, um die Anzahl der Parameter und die Rechenlast im Netzwerk zu verringern, wodurch auch die Gefahr von Überanpassung (siehe Kapitel 2.2.1.3) reduziert wird. Das am häufigsten verwendete Pooling-Verfahren ist das Max-Pooling, bei dem aus einem kleinen Bereich der Merkmalskarten der maximale Wert ausgewählt und die restlichen Werte verworfen werden. Aber es gibt auch andere Verfahren, wie z.B. Average-Pooling, das stattdessen den Durchschnittswert berechnet.

Eine spezielle Variante einer Pooling-Schicht ist das sogenannte räumliche Pyramiden Pooling. Dabei kann der Eingang verschiedene Größen und Dimensionen haben und die Merkmalskarten werden nach einer festen Pyramidenstruktur verarbeitet. In dieser Struktur wird jede Ebene der Pyramide mit einem unterschiedlichen Pooling-Bereich versehen, sodass aus denselben Merkmalskarten mehrere gebietsweise gepoolte Ausgaben erzeugt werden. Diese gepoolten Ausgaben werden dann zu einem einzigen Vektor fixer Länge zusammengefasst, unabhängig von der ursprünglichen Größe oder dem Seitenverhältnis des Eingabebildes.

**Normalisierungsschicht:** Normalisierungsschichten, wie die Stapel-Normalisierung (englisch Batch normalization Layer oder BnL), werden eingesetzt, um das Training zu stabilisieren und zu beschleunigen. Sie normalisieren die Ausgabe einer anderen Schicht, sodass die Aktivierung eine bestimmte Verteilung (in der Regel mit einem Mittelwert von 0 und einer Standardabweichung von 1) hat. Dies hilft, das Problem des internen kovarianten Verschiebens (siehe Kapitel 2.2.1.3) zu reduzieren, bei dem die Verteilung der Eingaben einer Schicht während des Trainings stark schwankt.

**Dropout-Schicht:** Dropout ist eine Technik zur Vermeidung von Überanpassung (siehe Kapitel 2.2.1.3). Während des Trainings werden zufällig Neuronen (und ihre Verbindungen) in der Dropout-Schicht mit einer bestimmten Wahrscheinlichkeit deaktiviert, was verhindert, dass das Netzwerk zu abhängig von bestimmten Neuronenverbindungen wird und somit eine bessere Generalisierung auf unbekannte Daten ermöglicht.

Diese Schichten arbeiten zusammen, um ein neuronales Netz zu bilden, das in der Lage ist, komplexe Aufgaben wie Objekterkennung, Bildklassifikation und sogar Bildsegmentierung durchzuführen. Die Fähigkeit, automatisch und effizient relevante Merkmale aus Bildern zu lernen, ohne dass eine manuelle Feature-Extraktion erforderlich ist, ist dabei eine essenzielle Stärke.

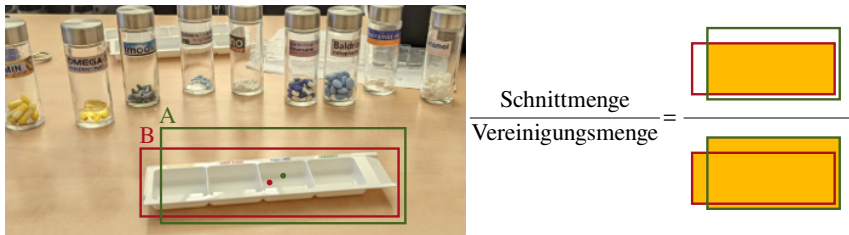
### 2.2.1.2 Der überwachte Lernprozess

Der Lernprozess in neuronalen Netzwerken ist ein iteratives Verfahren, bei dem das Netzwerk kontinuierlich angepasst wird, um die Diskrepanz zwischen den Vorhersagen und den tatsächlichen Daten zu minimieren. Das überwachte Lernen (englisch supervised learning) steht dabei im Gegensatz zu unüberwachtem Lernen, bei dem das neuronale Netzwerk ungesteuert versucht, relevante Informationen zu extrahieren. Die folgenden Aspekte spielen dabei eine wichtige Rolle in diesem Prozess:

**Trainingsdaten:** Im Rahmen des überwachten Lernens ist die Datenaufbereitung und -annotation entscheidend. Hierbei werden die Trainingsdaten gesammelt, vorverarbeitet und mit den entsprechenden Grundwahrheitswerten (englisch labels) versehen. Diese Grundwahrheitswerte werden vor dem Training manuell oder automatisiert festgelegt und zusammen mit den korrespondierenden Daten abgespeichert. Je nach gewünschter Modellausgabe können sie in Form von numerischen Werten korrespondierend zu einer Klasse oder Bildausschnitten wie in Abbildung 2.6 vorliegen. Sie dienen als Referenzpunkte, die das Netz während des Trainings verwendet, um die Korrektheit seiner Vorhersagen zu messen und seine Gewichte anzupassen. Die Qualität und Vielfalt der Daten beeinflussen maßgeblich die Leistungsfähigkeit des Modells, da gut annotierte und diverse Datensätze dem Modell erlauben, das Gelernte zu generalisieren und anschließend auf ungesehene Daten anzuwenden.

**Kostenfunktion:** Die Kostenfunktion quantifiziert den Fehler zwischen den vorhergesagten Werten des Netzwerks und den tatsächlichen Grundwahrheitswerten. Das Ziel des Trainings ist es, diesen Fehler zu minimieren. Es gibt verschiedene Arten von Kostenfunktionen, die je nach Art des Lernproblems ausgewählt werden. Für Klassifikationsaufgaben wird häufig die Kreuzentropie verwendet, die die Distanz zwischen den vorhergesagten Wahrscheinlichkeitsverteilungen und den tatsächlichen Verteilungen misst. Bei Regressionsaufgaben ist der mittlere quadratische Fehler eine gängige Wahl, der die durchschnittliche quadratische Differenz zwischen den Vorhersagen und den tatsächlichen Werten berechnet.

Bei Lokalisierungsaufgaben wird üblicherweise die Schnittmenge über der Vereinigungsmenge (englisch intersection over union) gewählt, also der Prozentsatz, mit der sich die Fläche der Modellausgabe mit dem Grundwahrheitswert deckt. Abbildung 2.6 zeigt, wie dieser Wert berechnet wird. Eine Weiterentwicklung davon ist die sogenannte Vollständige Schnittmenge über der Vereinigungsmenge, bei der zusätzlich der Abstand der Mittelpunkte und der Unterschied in den Seitenverhältnissen mit einfließt.



**Abbildung 2.6** Die Schnittmenge über der Vereinigungsmenge von Modellausgabe (A) und Grundwahrheitswert (B) auf einem Trainingsbild.

**Fehler-Rückpropagation:** Ein essenzieller Mechanismus im Lernprozess neuronaler Netzwerke ist die Fehler-Rückpropagation (englisch Backpropagation). Sie erlaubt es dem Netzwerk, den Fehler von der Ausgangsschicht rückwärts durch die verborgenen Schichten zu propagieren, um den Beitrag jedes Parameters (Gewicht und Bias) zum Gesamtfehler zu quantifizieren. Hierbei werden die Gradienten der Kostenfunktion hinsichtlich der Parameter berechnet, was dem Netzwerk ermöglicht, die Parameter in einer Weise zu optimieren, die den Fehler minimiert. Die Lernrate ist dabei der Faktor, mit dem der Fehler jedes Parameters multipliziert wird, bevor der Parameter angepasst wird. Dieser Prozess wird mehrfach in sogenannten Epochen wiederholt, wobei in jeder Epoche die gesamten Trainingsdaten durchgegangen werden. Die Lernrate wird dabei üblicherweise nach jeder Epoche reduziert, sodass die Parameter über die Zeit auf einen Wert konvergieren.

**Optimierungsmethoden:** Optimierungsmethoden sind Algorithmen, die die Parameter des neuronalen Netzwerks (wie Gewichte und Grundversatz) so anpassen,



dass die Kostenfunktion minimiert wird. Der wohl bekannteste Algorithmus ist der stochastische Gradientenabstieg (englisch Stochastic Gradient Descent), der die Parameter in die Richtung des negativen Gradienten der Kostenfunktion aktualisiert, basierend auf einer kleinen Stichprobe von Trainingsdaten. Es gibt viele Varianten des stochastischen Gradientenabstiegs, die zusätzliche Konzepte wie Momentum oder adaptive Lernraten einbeziehen, um die Konvergenzgeschwindigkeit zu erhöhen und Probleme wie lokale Minima oder das Verschwinden des Gradienten (siehe Kapitel 2.2.1.3) zu umgehen. Am häufigsten wird der sogenannte Adam-Algorithmus eingesetzt, der sowohl die Vorteile des Momentums als auch der adaptiven Lernraten miteinander kombiniert, um eine effiziente und effektive Optimierung zu gewährleisten.

Durch die Kombination dieser Elemente – Kostenfunktion, Optimierungsmethode und Fehler-Rückpropagation – wird ein neuronales Netz darauf trainiert, komplexe Muster in Daten zu erkennen und vorherzusagen. Der Erfolg dieses Lernprozesses hängt von der sorgfältigen Auswahl der Architektur, der Kostenfunktion und der Optimierungsmethode ab, sowie von der Qualität und Vielfalt der Trainingsdaten.

### 2.2.1.3 Mögliche Probleme

**Überanpassung:** Überanpassung (englisch Overfitting) tritt auf, wenn ein neuronales Netzwerk die Trainingsdaten einschließlich deren Rauschen und irrelevanten Details zu gut lernt. Dies führt dazu, dass das Modell auf den Trainingsdaten eine sehr hohe Genauigkeit erzielt, jedoch auf neuen, ungesehenen Daten schlecht generalisiert. Um Überanpassung zu verhindern, werden verschiedene Techniken eingesetzt, darunter die Verwendung von Regularisierungsmethoden wie Dropout-Schichten und die Erhöhung der Trainingsdatensatzgröße.

**Verschwindender Gradient:** Das Problem des verschwindenden Gradienten (englisch Vanishing Gradient) tritt auf, wenn die Gradienten der Kostenfunktion in den tieferen Schichten eines neuronalen Netzwerks extrem klein werden. Dies führt dazu, dass die Gewichte in diesen Schichten kaum aktualisiert werden, was das Lernen nahezu unmöglich macht. Um dieses Problem zu bewältigen, werden

verschiedene Methoden angewendet, wie der Verwendung von mitigierenden Aktivierungsfunktionen (z.B. die rektifizierende Linearfunktion), dem Einsatz von Stapel-Normalisierung oder der Querphasen-Teilschicht.

**Interne kovariante Verschiebung:** Interne kovariante Verschiebung (englisch Internal Covariate Shift) beschreibt die Änderung der Verteilung der Eingaben für jede Schicht während des Trainingsprozesses. Diese Verschiebung kann das Training instabil und ineffizient machen, da jede Schicht kontinuierlich ihre Eingaben anpassen muss. Stapel-Normalisierung ist eine weit verbreitete Technik zur Minderung dieses Problems. Dadurch bleibt deren Verteilung während des Trainings konstant, was die Trainingsgeschwindigkeit erhöht und die Konvergenz verbessert.

## 2.2.2 Architekturen zur Klassifikation

Die Klassifikation von Bildern ist ein grundlegendes Problem im Bereich des maschinellen Lernens. Sie zielt darauf ab, den Inhalt eines Bildes in eine oder mehrere Kategorien einzuordnen. Klassifikationsmodelle, insbesondere Faltungsneuronalen Netzwerke, sind in der Lage, Wahrscheinlichkeiten für die Zugehörigkeit zu jeder eintrainierten Klasse auszugeben.

### 2.2.2.1 Vorläufer: LeNet-5, AlexNet und VGG16

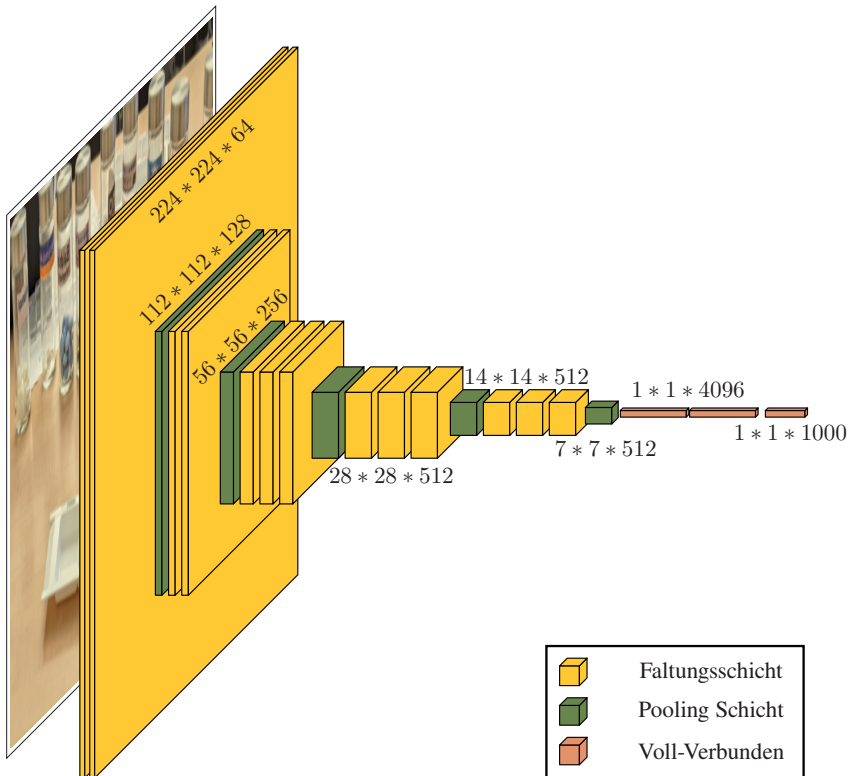
Die Entwicklung von LeNet-5 und AlexNet legte wesentliche Grundlagen für die Fortschritte im Bereich der Convolutional Neural Networks (CNNs). LeNet-5, veröffentlicht 1998, demonstriert erstmals die praktische Anwendbarkeit von CNNs zur automatisierten Handschriftenerkennung. Die Architektur von LeNet-5 besteht aus mehreren Faltungsschichten zur Merkmalsextraktion, Pooling-Schichten zur Dimensionsreduktion und voll verbundenen Schichten zur Klassifikation, und führte die Fehler-Rückpropagation erstmalig erfolgreich ein [6].

AlexNet, vorgestellt im Jahr 2012, baut auf diesen Konzepten auf und bringt entscheidende Innovationen, die die Leistungsfähigkeit von CNNs erheblich verbessern. Zu den wichtigsten Neuerungen gehören die Verwendung der ReLU-Aktivierungsfunktion zur Vermeidung des verschwindenden Gradientenproblems, die Implementierung von Dropout-Schichten zur Reduktion von Überanpassung sowie die Einführung von Normalisierungsschichten und überlappenden Pooling-Schichten zur Verbesserung der Merkmalsextraktion [7]. Diese Fortschritte führen dazu, dass AlexNet die ImageNet Large Scale Visual Recognition Challenge 2012 mit einer erheblich niedrigeren Fehlerrate von 15,3% gewinnt [8].

Die Einführung des VGG16-Modells, dargestellt in Abbildung 2.7, im Jahr 2014 durch die Visual Geometry Group der Universität Oxford stellt weitere Neuerungen wie die tiefe Architektur vor, die aus 16 gewichteten Schichten besteht und damit ein hohes Maß an interner Abstraktion ermöglicht. Zusätzlich nutzt dieses Modell ausschließlich sehr kleine Faltungsmasken von 3x3, was im Vergleich zu den vorherigen Modellen eine effizientere Extraktion von Bildmerkmalen erlaubt. VGG16 integriert außerdem mehrere vollständig verbundene Schichten am Ende des Netzwerks, die die erfassten Merkmale für komplexe Klassifikationsaufgaben interpretieren.

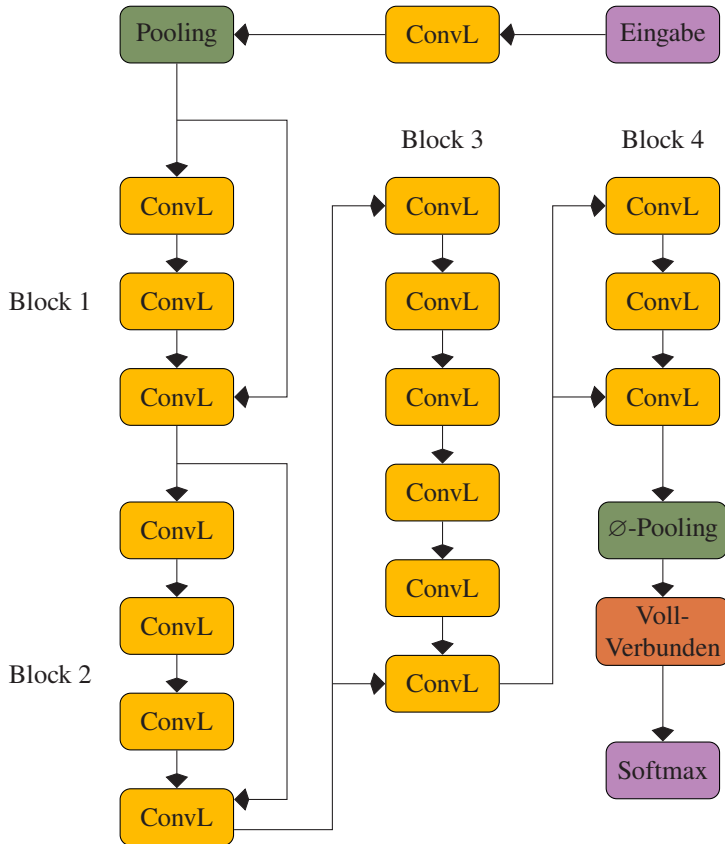
### **2.2.2.2 ResNet**

Die Einführung des ResNet-Modells im Jahr 2015 stellte einen bedeutenden Durchbruch im Bereich des Maschinenslernens dar [9]. ResNet zeichnet sich, wie in Abbildung 2.8 dargestellt, durch eine tiefe Architektur aus, die 50 gewichtete Schichten umfasst und damit im Vergleich zu VGG16 ein noch höheres Maß an interner Abstraktion ermöglicht. Das Modell adressiert das Problem des verschwindenden Gradienten durch die Einführung von Residual-Blöcken. Diese ähneln in ihrer Grundidee den Querphasen-Teilschichten aus Abbildung 2.5, wobei der Eingang eines Blocks über mehrere Faltungsschichten hinweg zum Ausgang des gesamten Blocks kopiert wird. Dies ermöglicht eine direkte Weiterleitung der Ausgabe einer Schicht zu einer tiefer gelegenen Schicht, wodurch das Training effizienter gestaltet wird.



**Abbildung 2.7** VGG16-Architektur mit individuellen Schichten

ResNet nutzt, wie zuvor VGG16, kleine Faltungsmasken von 3x3, erweitert diese jedoch durch die Residual-Verbindungen zu einer wesentlich tieferen Struktur. Zudem integriert es globale Mittelwertpooling-Schichten am Ende des Netzwerks, wodurch die Anzahl der Parameter reduziert und die Generalisierungsfähigkeit verbessert wird. Im Gegensatz zu VGG16 verwendet ResNet keine Pooling Schichten, die anfänglichen Dimensionen des Eingabebildes bleiben durch alle Schichten hinweg gleich. Das Modell erzielte Bestergebnisse bei der ImageNet Challenge 2015, indem es den ersten Platz mit einer Top-5-Fehlerquote von 3,57% belegte.



**Abbildung 2.8** ResNet-Architektur mit vier Blöcken und deren Beipassverbindungen.

Ein Schlüsselement des ResNet-Modells ist seine Skalierbarkeit und Flexibilität, die es ermöglicht, sehr tiefe Netzwerke ohne Überanpassung zu trainieren. Dies macht ResNet zu einer bevorzugten Wahl für viele Forschungsprojekte und industrielle Anwendungen. Trotz der größeren Tiefe bleibt das Modell dank der Residual-Strukturen effizient trainierbar.

### 2.2.2.3 Alternativen

InceptionNet, auch bekannt als GoogLeNet, kombiniert in seinen Inception-Modulen verschiedene Filtergrößen innerhalb derselben Ebene, wodurch Merkmale auf unterschiedlichen Skalen gleichzeitig extrahiert werden können [10]. Diese Architektur reduziert die Rechenkosten und erzielt dennoch hohe Genauigkeiten auf großen Datensätzen.

EfficientNet bietet ein skalierbares Modell für neuronale Netzwerke durch einen "Compound ScalingAnsatz [11]. Dabei werden Tiefe, Breite und Auflösung des Netzwerks systematisch skaliert, um eine hohe Genauigkeit bei geringer Rechenleistung zu erreichen. EfficientNet ist besonders geeignet für Anwendungen, bei denen Effizienz entscheidend ist.

## 2.2.3 Architekturen zur Detektion

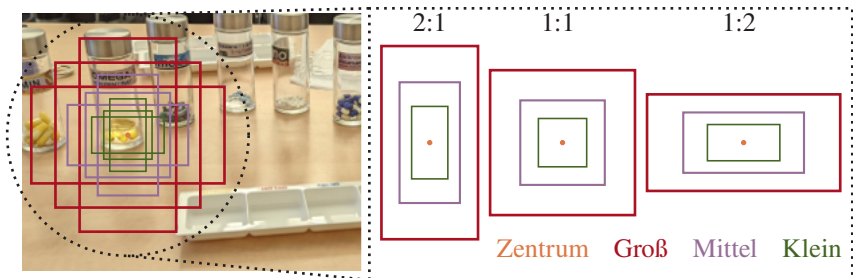
Bei der Objektdetektion geht es darum, nicht nur die Kategorien der Objekte in einem Bild zu identifizieren, sondern auch deren genaue Positionen innerhalb des Bildes zu bestimmen. Dies beinhaltet das Erkennen von Objekten verschiedener Größen und Formen in unterschiedlichen Szenarien und Hintergründen. Die Ausgabe der Modelle enthält die Begrenzungsrahmen und die Klassifikationswahrscheinlichkeiten.

### 2.2.3.1 Historie von YOLOv5

Die Evolution der YOLO (You Only Look Once) Detektoren beginnt mit YOLOv1, das 2015 von Joseph Redmon vorgestellt wird. Dieses revolutioniert die Objekterkennung, indem sie diese als ein einziges Regressionsproblem betrachtet und somit Echtzeit-Detektionen ermöglicht. So verarbeitet das Standardmodell 45 Bilder pro Sekunde auf einer handelsüblichen Grafikkarte [12]. Allerdings hat YOLOv1 Schwierigkeiten bei der genauen Lokalisierung kleiner Objekte und der Erkennung von Objekten, die dicht beieinander liegen.

YOLOv2, eingeführt 2016, verbessert die Genauigkeit und Geschwindigkeit durch die Einführung von Ankerrahmen zur Erkennung mehrerer Objekte in unmittelbarer Nähe, durch Batch-Normalisierung zur Stabilisierung des Trainings und Reduzierung von Überanpassung sowie durch die Darknet-19-Architektur, die aus 19 Faltungsschichten und 5 Max-Pooling-Schichten besteht und eine effizientere Merkmalsextraktion ermöglicht. Das Modell erreicht eine mittlere durchschnittliche Präzision bei einem Schnittmenge über der Vereinigungsmenge Schwellwert von 0,5 ( $mAP_{50}$ ) von 78,6% auf dem PASCAL VOC 2007 Datensatz im Gegensatz zu YOLOv1's 63,4% [13].

YOLOv3, veröffentlicht 2018, bringt weitere Optimierungen durch die Einführung von dimensionalen Clustern, also der Gruppierung von Ankerrahmen verschiedener Größen (vergleiche Abbildung 2.9). Zudem werden unabhängige, logistische Klassifikatoren zu den Ankerrahmen eingesetzt, die die Identifikation mehrerer Objekte ermöglichen und so eine mehrfache Klassifizierung desselben Objekts in unterschiedliche Kategorien zulässt (Ein »Dackel« kann zusätzlich als »Hund« klassifiziert werden). Neben den Verbesserungen an den Klassifi-



**Abbildung 2.9** Dimensionales Clusters von Ankerrahmen mit gemeinsamen Zentrum.

katoren wurde auch das Darknet-Modell weiterentwickelt. Durch die Integration zusätzlicher residualer Netzwerkstrukturen erhält das Modell Zugriff auf eine breitere Palette von Abstraktionsebenen, was die Klassifikationsgenauigkeit verbessert. Diese Strukturen erleichtern außerdem den Lernprozess, indem sie das Problem verschwindender Gradienten in tiefen Netzwerken verringern, wodurch

wiederum zusätzliche Schichten integriert werden können. Infolgedessen wurde das Darknet-Modell auf insgesamt 53 ConvL erweitert und trägt nun den Namen Darknet-53. [14].

YOLOv4, veröffentlicht 2020, setzt die Optimierung der Objekterkennung fort, indem sowohl die Geschwindigkeit als auch die Genauigkeit weiter gesteigert werden. Auf dem MS COCO (testdev 2017) Datensatz, der deutlich anspruchsvoller ist als der PASCAL VOC 2007, erreicht YOLOv4 eine  $mAP_{50}$  von 65,7%, im Gegensatz zu YOLOv3's 57,9% [15]. Das Modell verwendet die CSPDarknet-53-Architektur, die die Effizienz der Berechnungen durch die Einführung von CspL verbessert. Zusätzlich integriert YOLOv4 ein Pfadaggregationsnetzwerk, das bidirektionale Pfade zwischen verschiedenen Abstraktionsebenen einführt, um hochauflösende und semantisch reichhaltige Informationen effektiv zu kombinieren. Moderne Datenaugmentierungstechniken wie Mosaic, die vier Bilder zu einem einzigen kombinieren, und selbst-adversariales Training, das die Widerstandsfähigkeit des Modells durch gezielt generierte irreführende Bilder erhöht, tragen ebenfalls zur Leistungssteigerung bei [15].

### 2.2.3.2 YOLOv5

YOLOv5, das 2020 veröffentlicht wurde, stellt einen bedeutenden Fortschritt in der Familie der YOLO-Modelle dar. Diese Version führt mehrere technische Neuerungen ein, die sowohl die Automatisierung von Prozessen als auch die Leistungsfähigkeit des Modells erheblich verbessern. Eine der auffälligsten Änderungen ist die Implementierung in PyTorch, einem offenen maschinellen Lernframework, was die Integration in bestehende Projekte vereinfacht und gleichzeitig die Entwicklung und das Training beschleunigt. Die Architektur von YOLOv5 ist in Abbildung 2.10 gezeigt.

#### **Automatisierung:**

YOLOv5 nutzt eine automatisierte Anpassung der Ankerrahmengrößen durch K-Means-Clustering auf die Trainingsdaten. Dieser Prozess optimiert die Initialisierung der Ankerrahmen, was die Effizienz der Objekterkennung verbessert. Die



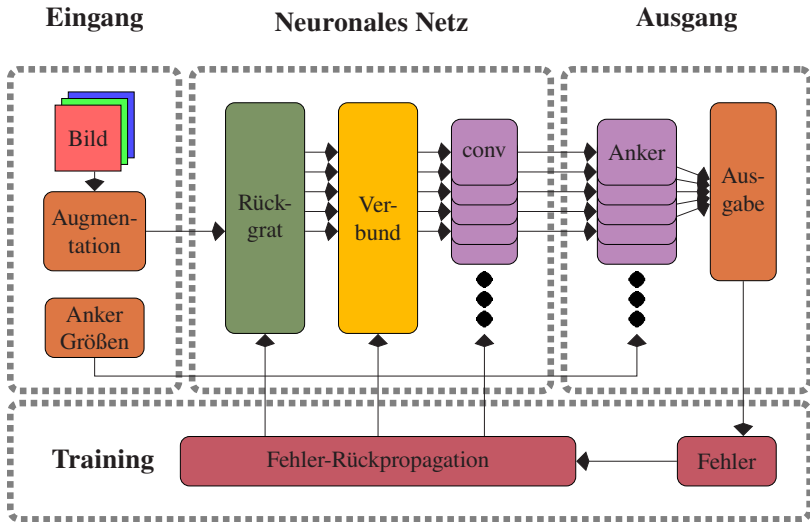


Abbildung 2.10 YOLOv5 Aufbau

dynamische Anpassung der Batch-Größe an die verfügbaren Systemressourcen ist eine weitere Automatisierungsmaßnahme, die eine effizientere Nutzung der Hardware ermöglicht und das Training auf unterschiedlichen Systemen vereinfacht. Diese Automatisierungen reduzieren den manuellen Aufwand und erlauben eine schnelle Anpassung des Modells an verschiedene Datensätze und Hardwareumgebungen

### Weiterentwicklung:

Zu den wichtigsten technischen Verbesserungen gehört die Weiterentwicklung der Querphasen-Teilschichten (CSPNet), die bereits in YOLOv4 eingeführt wurden (siehe Abbildung 2.11). Diese Schichten erhöhen die Modellgenauigkeit durch eine effizientere Repräsentation von Merkmalen, ohne die Verarbeitungsgeschwindigkeit zu beeinträchtigen. Darüber hinaus integriert YOLOv5 Techniken zur Datenanreicherung wie Mosaic und MixUp, bei denen aus vier Trainingsbildern ein neues generiert wird. So wird das Modell robuster gegenüber unterschiedlichen Bildbedingungen und Überanpassungen werden vermieden.

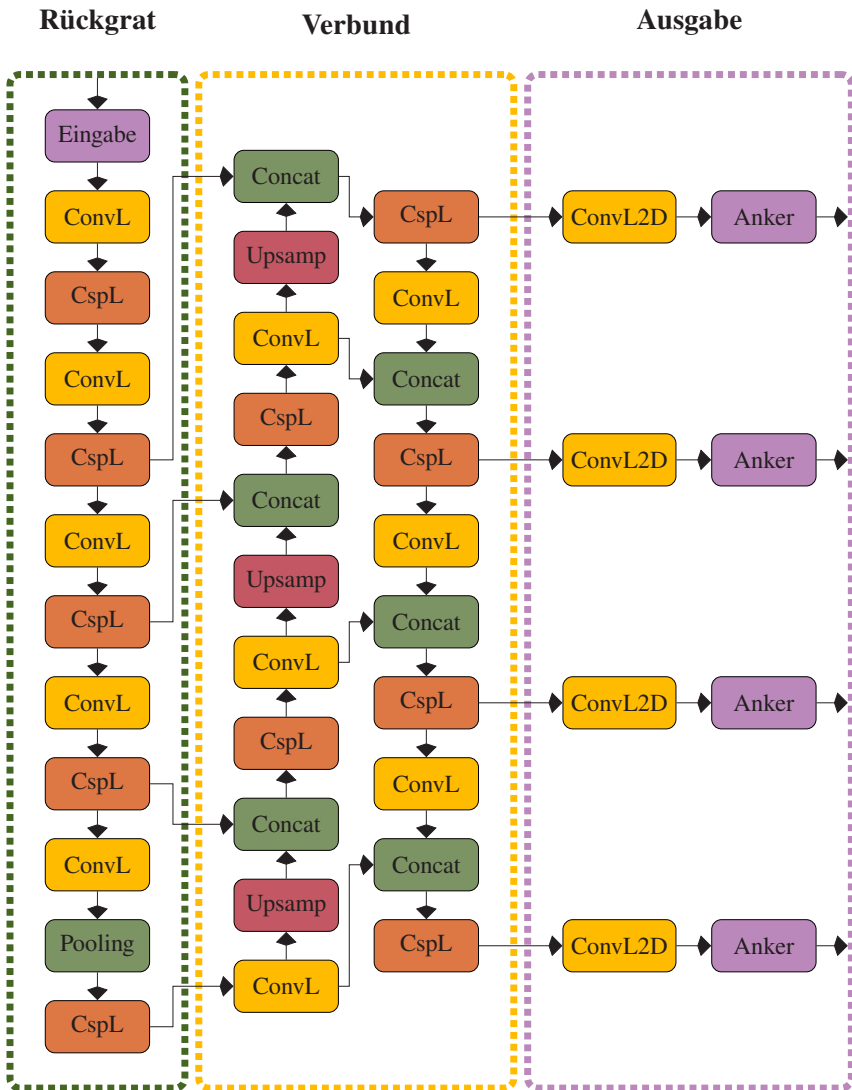


Abbildung 2.11 Detailansicht der Struktur von Rückgrat, Verbund und Ausgabe von YOLOv5

Eine weitere bedeutende Verbesserung ist die Modell-Skalierung. YOLOv5 bietet vorkonfigurierte Modelle in verschiedenen Größen (YOLOv5s, YOLOv5m, YOLOv5l und YOLOv5x), die eine einfache Anpassung an spezifische Anforderungen hinsichtlich Geschwindigkeit und Genauigkeit ermöglichen. Diese Modelle sind auf unterschiedliche Ressourcenanforderungen und Anwendungsfälle abgestimmt und bieten eine flexible Lösung für eine breite Palette von Einsatzszenarien.

### 2.2.3.3 Alternativen

Die meisten Alternativen zu YOLO basieren auf einem getrennten Ansatz, bei dem die Lokalisierung und die Klassifikation von zwei getrennten Strukturen durchgeführt werden. Dabei wird zur Klassifikation in den meisten Fällen VGG-16 (siehe Kapitel 2.2.2) verwendet.

**LRF (Local Receptive Fields)** VGG-16 verbessert die Fähigkeit der Objekterkennung durch das Hinzufügen lokaler Rezeptivfelder, die helfen, kleinere und detailliertere Objekte zu identifizieren, was gegenüber YOLO eine detailliertere Analyse auf Kosten der Geschwindigkeit bietet. RFBNet (Receptive Field Block Net) VGG-16 optimiert dies weiter durch eine spezielle Struktur, die die Rezeptivfelder variiert, um eine breitere Palette von Objektgrößen zu erkennen, was es vielseitiger als YOLO macht, aber auch komplexer in der Implementierung.

**SSD (Single Shot MultiBox Detector)** VGG-16 ist bekannt für seine Balance zwischen Geschwindigkeit und Genauigkeit [16]. Es kann Objekte in verschiedenen Größen durch die Verwendung multipler VGG-16-basierter Detektionsköpfe erkennen, die auf verschiedenen Auflösungen der Eingabebilder operieren. Dies macht es schneller als YOLO bei vergleichbarer Genauigkeit, vor allem in Umgebungen mit vielen kleinen Objekten. RefineDet VGG-16 baut auf SSD auf, indem es eine zweistufige Detektion einsetzt, die zuerst eine grobe und dann eine feinere Erkennung durchführt. Dies verbessert die Genauigkeit deutlich, was es besonders nützlich für Anwendungen macht, wo Präzision kritisch ist, allerdings auf Kosten der Latenz.

**EfficientDet-D0**, obwohl nicht auf VGG-16 basierend, ist eine EfficientNet-Architektur, die durch Skalierbarkeit der Netzwerktiefe, -breite und der Auflösung besticht [11]. Im Vergleich zu VGG-16-basierten Modellen ist EfficientDet in der Regel effizienter und skaliert besser auf niedrigeren Hardware-Ressourcen, bietet jedoch nicht die gleiche Genauigkeit auf komplexen Bildinhalten wie die spezifischeren VGG-16-Anpassungen.

## 2.3 Psychologische Grundlagen

Psychologische Grundlagen sind wichtig im Zusammenhang mit komplexen Aufgaben und zeigen, wie das menschliche Gehirn mit Belastung umgeht. Sie bilden die Grundlage für das Design effektiver Mensch-Maschinen-Interaktionen.

### 2.3.1 Kognitive Belastungstheorie

Die Kognitive Belastungstheorie (KBT) wurde im Bereich der kognitiven Psychologie entwickelt und bietet einen fundierten Rahmen zur Analyse der kognitiven Belastung während Lern- und Verarbeitungsprozessen [17]. Die Theorie geht von der Prämisse aus, dass die begrenzte Kapazität des Arbeitsgedächtnisses entscheidend für erfolgreiches Lernen und die Bewältigung von Aufgaben ist. Die KBT unterscheidet drei Hauptarten der kognitiven Belastung:

- **Intrinsische Belastung:** Bezieht sich auf die Komplexität der zu verarbeitenden Information selbst. Sie ist direkt mit der Aufgabe verbunden und hängt vom Schwierigkeitsgrad und dem Vorwissen der Person ab. Im Kontext des Medikamentenmanagements bezieht sich die intrinsische Belastung auf die grundlegende Herausforderung, Medikamente korrekt zu identifizieren und zuzuordnen, sowie auf die Bewältigung externer Störfaktoren.

- **Extrinsische Belastung:** Entsteht durch die Art der Informations- oder Aufgabenpräsentation. Übermäßig komplexe oder schlecht strukturierte Informationen können die extrinsische Belastung erhöhen. Ein schlecht gestalteter Medikamentenplan, der die Aufmerksamkeit auf nebensächliche Elemente lenkt, könnte im Bereich des Medikamentenmanagements zu einer gesteigerten extrinsischen Belastung führen.
- **Relevante Belastung:** Bezieht sich auf die kognitive Kapazität, die für ein tiefgehendes Verständnis und die Verarbeitung von Informationen eingesetzt wird. Sie ist die effektive Belastung, die zur Bildung neuer oder zur Anpassung bestehender Schemata im Langzeitgedächtnis beiträgt. Im Bereich des Medikamentenmanagements umfasst dies die kognitive Anstrengung, die in das Erlernen und Anwenden effizienter Strategien für die Sortierung von Medikamenten investiert wird.

Im Folgenden wird die KBT herangezogen, um zu verstehen, wie Systeme die kognitive Belastung beeinflussen könnten. Aus der KBT lässt sich ableiten, dass für eine effiziente Informationsverarbeitung und Aufgabenbewältigung die Gesamtbelastung des Arbeitsgedächtnisses optimiert werden sollte. Die intrinsische Belastung einer Aufgabe kann dabei nicht beeinflusst werden. Allerdings kann die extrinsische Belastung reduziert werden, um ausreichend Kapazität für die intrinsische und die relevante Belastung zu schaffen. Dies kann zum Beispiel durch eine klare und intuitive Nutzeroberfläche erreicht werden, was effektive Lern- und Arbeitsstrategien fördert, und dadurch die Gesamteffizienz und Genauigkeit beim Medikamentenmanagement verbessert.

### **2.3.2 Dual-Prozess-Theorie**

Die Dual-Prozess-Theorie (DPT), ein anerkanntes Konzept in der kognitiven Psychologie, geht von zwei fundamentalen Mechanismen der Informationsverarbeitung im menschlichen Denken aus, die oft als System 1 und System 2 klassifiziert werden [18]. Diese Systeme stehen für zwei unterschiedliche Modi der Verarbeitung, die sich in ihrer Arbeitsweise und ihrem Einsatz unterscheiden.

- System 1 ist für schnelle, automatische und intuitive Denkprozesse bekannt. Diese Abläufe finden oft unbewusst statt und benötigen nur geringe kognitive Anstrengungen. Es ist besonders effektiv in der Erkennung von Mustern und kann komplexe Situationen unverzüglich einschätzen, was häufig zu automatisierten und routinemäßigen Entscheidungen führt. Typische Beispiele für die Aktivierung von System 1 sind das Wiedererkennen von bekannten Gesichtern, das Verstehen einfacher Sprachstrukturen oder das Ausführen von Gewohnheitshandlungen, wie alltäglichen Aufgaben. Das Medikamentenmanagement ist für viele Pflegekräfte eine alltägliche Aufgabe, die vor allem durch System 1 abgearbeitet wird.
- System 2 hingegen ist verantwortlich für langsamere, bewusste und analytische Denkprozesse, die eine erhöhte kognitive Last mit sich bringen. Dieses System kommt bei komplexeren Aufgaben zum Einsatz, die gezielte Aufmerksamkeit und bewusste mentale Anstrengungen erfordern. Typische Einsatzgebiete für System 2 sind das Lösen von anspruchsvollen mathematischen Aufgaben, das Treffen durchdachter Entscheidungen oder das Erlernen neuer Fähigkeiten. Vor allem neue Arten von Medikamenten oder das Einlernen sind Beispiele, die im Medikamentenmanagement den Einsatz von System 2 erfordern.

Diese beiden Systeme interagieren im praktischen Einsatz oft miteinander: System 1 liefert schnelle, initiale Bewertungen und Urteile, die dann durch System 2 überprüft und bei Bedarf korrigiert werden. In alltäglichen Situationen neigen Menschen dazu, sich überwiegend auf System 1 zu verlassen, da es weniger kognitive Ressourcen beansprucht und schnelle Reaktionen ermöglicht.

Im Folgenden bieten die Erkenntnisse der DPT einen Erklärungsansatz für die Interaktion zwischen einem System und dem Nutzer. Nach der DPT kann die intuitive Informationsverarbeitung durch System 1 gefördert werden, indem klare visuelle Hinweise und eine leicht verständliche Benutzeroberfläche bereitgestellt werden. Damit kann der Bedarf an intensiveren analytischen Prozessen durch System 2 reduziert werden und somit die Effizienz und Genauigkeit gesteigert werden.

## 3 Stand der Technik

Der aktuelle Stand der Technik unterteilt sich in zwei Kategorien. Zunächst wird auf die Prozesse eingegangen, wie sie bereits in Pflegeheimen eingesetzt werden. Dazu gehören das klassische Medikamentenrichten, bei dem die einzelnen Medikamente für Patienten entsprechend ihrer Medikation manuell vorbereitet werden, aber auch maschinelle Verfahren, wie sie in großen Apotheken eingesetzt werden. Anschließend wird ein Überblick über den Stand der Wissenschaft gegeben. In der Forschung wird zunehmend der Einsatz moderner Technologien wie Künstliche Intelligenz, Augmented Reality und Automatisierung untersucht, aber auch konkrete Beispiele und Ideen, wie das Medikamentenmanagement verbessert werden kann.

### 3.1 Aktuelles Vorgehen beim Medikamentenmanagement

Nach §48 des Arzneimittelgesetzes dürfen Medikamente nur von Ärzten verschrieben werden [19]. Die genaue Vergabe wird dann in einem Medikationsplan festgehalten. Die Aufgabe von Pflegekräften ist es, die verordneten Medikamente dem Patienten zur richtigen Zeit zu verabreichen. Dazu werden diese in Tabletendosetten, wie in Abbildung 3.1 zu sehen, vorbereitet.



Abbildung 3.1 Tablettendosetten mit Fächern für Morgens, Mittags, Abends und Nachts

### 3.1.1 Händisches Einsortieren

Das händische Einsortieren ist die am weitesten verbreitete Methode des Medikamentenrichtens und wird in Lehrbüchern in einem standardisierten Prozess gelehrt [20]. Der Prozess beginnt mit der Vorbereitung der notwendigen Materialien durch das Pflegepersonal:

- **Medikamentenschrank:** Enthält alle verschriebenen Medikamente der Patienten.
- **Tablettendosetten:** Boxen mit Fächern für verschiedene Einnahmezeitpunkte (morgens, mittags, abends, nachts).
- **Medikationspläne:** Aktuelle ärztliche Verordnungen für jeden Patienten.
- **Hygienische Maßnahmen:** Das Personal wäscht sich die Hände und trägt eventuell Handschuhe, um eine Kontamination zu vermeiden.

Dabei muss darauf geachtet werden, dass die Medikationspläne aktuell sind und alle Änderungen berücksichtigt wurden. Anschließend werden die Medikamente entsprechend der ärztlichen Anordnung für jeden Bewohner einzeln aus den



Originalverpackungen entnommen und in die entsprechenden Fächer der Tabletendosetten einsortiert. Dieser Schritt erfolgt unter Beachtung der »5-R-Regel«:

- **Richtiger Patient:** Es wird sichergestellt, dass die Medikamente dem richtigen Bewohner zugeordnet werden.
- **Richtiges Medikament:** Es wird überprüft, dass das richtige Medikament ausgewählt wird.
- **Richtige Dosierung:** Die genaue Menge des Medikaments wird abgemessen.
- **Richtige Zeit:** Die Medikamente werden in das entsprechende Zeitfach gelegt.
- **Richtige Verabreichungsform:** Falls erforderlich, wird das Medikament zerkleinert oder auf eine andere Verabreichungsform vorbereitet.

Nach dem Richten der Medikamente erfolgt eine erneute Kontrolle, entweder durch dasselbe oder ein anderes Pflegepersonalmitglied. Diese doppelte Überprüfung ist ein wichtiger Schritt zur Vermeidung von Fehlern. Abschließend wird der Prozess dokumentiert. Dabei wird das Richten auf der Medikamentenkarte festgehalten und gegebenenfalls wird ein Kontrollblatt zur Verifizierung des Richtprozesses genutzt. Das Kontrollblatt kommt insbesondere bei Hochrisikomedikamenten zum Einsatz.

### 3.1.2 Maschinelles Verfahren

Beim maschinellen Verfahren werden in einer großen Apotheke die Medikamente vorverblistert. Dabei sortiert eine große Maschine die Medikamente bereits in die richtigen Segmente der Dosetten ein. Anschließend wird die Dosette versiegelt und an das Pflegeheim oder den Patienten ausgeliefert. Abbildung 3.2 zeigt eine Auswahl solcher Maschinen. Innerhalb der Maschine befinden sich Container, die die bestellten Medikamente enthalten. Mechanisch werden diese an die richtige Stelle in der Dosette geleitet.



**Abbildung 3.2** Auswahl an Verblisterungsautomaten. Von links nach rechts: Omnicell VBM200F [21], JVM Primeblister [22] und DB Rowa Smart [23]

Die angelieferten Dosetten müssen von den Pflegekräften vor der Verabreichung der Medikamente noch einmal überprüft werden. Im Falle einer Änderung der verabreichten Medikamente zwischen Bestellung und Verabreichung werden die Medikamente händisch angepasst. In diesem Fall wird der in Kapitel 3.1.1 beschriebene Prozess eingesetzt.

## 3.2 Stand der Wissenschaft

Der Stand der Wissenschaft lässt sich unterteilen in Betrachtungen von KI-Systemen zur Erkennung von Medikamenten und Ansätze, die das Medikamentenmanagement generell technisch unterstützen wollen. Diese enthalten sowohl AR-basierte Ansätze, als auch andere Konzepte. Relevante KI-Technologien wurde bereits in Kapitel 2.2 vorgestellt.

### 3.2.1 KI im Medikamentenmanagement

Tan et al. [24] versuchten bereits mit frühen Klassifikatoren Tabletten zu identifizieren. Dabei wurden konkret RetinaNet, Single Shot Multi-Box Detector (SSD) und YOLO v3 miteinander verglichen. Die Ergebnisse zeigten, dass SSD mit

einer  $mAP$  von 83% die höchste Genauigkeit aufwies, aber mit nur 30% der Geschwindigkeit verglichen mit YOLOv3 deutlich langsamer ist. YOLOv3 zeigte mit einer  $mAP$  von 80% jedoch immer noch eine sehr ähnliche Genauigkeit. RetinaNet platzierte sich sowohl von der Genauigkeit, als auch von der Geschwindigkeit, in der Mitte der beiden anderen Architekturen. Die Autoren empfahlen YOLOv3 für Echtzeitanwendungen, da die Genauigkeit sehr ähnlich ist, aber die Geschwindigkeit deutlich über den anderen Alternativen liegt.

Dhivya et al. [25] entwickelten einen Ansatz zur Identifikation von Tabletten auf Basis von Texterkennung mittels Support Vector Machine (SVM) und einer Fehlerkorrektur durch einen erweiterten N-Gramme-Algorithmus. Dieser Ansatz zielte darauf ab, die Genauigkeit der Erkennung von Texten auf Tablets zu verbessern, indem häufige Fehler in der Texterkennung durch den N-Gramme-Algorithmus korrigiert werden. Während das vorgeschlagene System ein vielversprechende Fehlerkorrekturtechnik vorstellt, müsste jeweils die Vorderseite der Tablette in hochauflösender Nahaufnahme in das System eingespeist werden, was einen Einsatz unter Realbedingungen schwierig macht.

Elyan et al. [26] gaben in ihrer Arbeit 2022 einen aktuellen Stand von verschiedenen KI-Algorithmen in der Medizintechnik und die damit verbundenen Herausforderungen. Eines der genannten Beispiele ist auch die kamerabasierte Erkennung von Medikamenten. So wurden unter anderem die dynamische Umgebung und die sich konstant weiterentwickelnde Medikamentenlandschaft, aber auch die generelle Schwierigkeit 3D-Objekte in 2D-Bildern abzubilden, als Herausforderungen genannt.

Nguyen et al. [27] schlugen ein System vor, das über Ähnlichkeitsmerkmale Bilder von Medikamenten mit denen in einer Datenbank paaren soll. Dabei wurden primär die Farbzusammensetzungen der Medikamente verwendet. Dieser Ansatz verwendet jedoch keine anderen Merkmale wie Schrift oder Gravuren auf den Medikamenten und macht so eine Unterscheidung zwischen Medikamenten gleicher Farbe und Form unmöglich.

### 3.2.2 Konzepte zum Medikamentenmanagement

Wie bereits in Kapitel 2.1 beschrieben, hat die Entwicklung der Augmented Reality Technologie in den letzten Jahren bedeutende Fortschritte gemacht und eröffnet somit vielfältige Anwendungsmöglichkeiten im Gesundheitswesen. Insbesondere im Bereich der Medikamentenabgabe werden AR-Brillen als innovative Werkzeuge zur Optimierung von Prozessen in Kliniken und Pflegeeinrichtungen gesehen.

Gansohr et al. [28] erarbeiteten ein Konzept, das jedoch noch nicht umgesetzt wurde. Bei diesem werden neue AR-Brillen in Zusammenarbeit mit Krankenhäusern und Pflegekräften entwickelt und die Aspekte bei der Medikamentenabgabe identifiziert, die das größte Optimierungspotenzial aufweisen. Das vorgeschlagene System soll unter anderem beim Richten von Medikamenten unterstützen. Dazu verwendet es Barcodes, um den Patienten zu identifizieren, für den das Medikament vorbereitet wird. In den Brillen werden zusätzliche Informationen angezeigt, wie die Namen der verschriebenen Medikamente, die Dosierung und der Standort des Medikaments innerhalb der Einrichtung. Der Weg zum Standort des Medikaments wird auf Anfrage wie bei einem Navigationsgerät angezeigt. Falls das Medikament nicht vorrätig ist, soll das System Alternativen vorschlagen. Während der Sortierung des Medikaments zeigt die AR-Brille zusätzliche Informationen ähnlich denen eines normalen Medikationsplans und ein Bild zum Vergleich mit dem physischen Gegenstück an. Ein System wie dieses würde einen hohen Grad an unterstützenden Technologien erfordern, sowie Veränderungen in der Krankenhausinfrastruktur, um realisiert zu werden. Viele der vorgestellten Ideen würden auch erhebliche Fortschritte in der Technologie selbst erfordern, bevor sie vollständig umsetzbar wären. Dennoch wird die Medikamentenvorbereitung als ein wichtiger Anwendungsfall identifiziert, bei dem technologische Unterstützung vorteilhaft wäre.

Aicher et al. [29] entwickelten ebenfalls eine Anwendung, die ein AR-System zum Unterstützen beim Medikamentenrichten verwendet. In diesem Vorschlag werden mehrere Medikamentendosetten in einer Vorrichtung platziert, um sicherzustellen, dass ihre Positionierung zueinander dem System bekannt ist. Dann wird ein Raster in Form der Medikamentendosetten über die physischen Objekte

gelegt. Um festzustellen, wie viel von einem Medikament hinzugefügt werden muss, scannt der Benutzer einen QR-Code auf dem Medikamentenbehälter. Die Benutzeroberfläche zeigt dann an, wie viel von dem Medikament in jedes Fach jeder Medikamentendosette gegeben werden muss. Allerdings wird die Benutzeroberfläche nicht aktualisiert, während der Benutzer die Medikamente verteilt, und potentielle Fehler des Benutzers werden nicht erkannt.

Chang et al. [30] konzentrierten sich in ihrer Arbeit auf die Entwicklung einer neuen Art von AR-Brille, die sehbehinderte Patienten bei der Verwaltung ihrer eigenen Medikation unterstützt. Das gesamte System besteht aus drei Teilen: der AR-Brille, die ständig getragen werden soll, einer cloudbasierten Managementplattform und einer Mobile-App. Nach dem Laden eines Medikationsplans über die App erinnert das AR-Gerät den Patienten zur richtigen Zeit über Sprachhinweise daran, seine Medikamente einzunehmen. Bei der Einnahme der Medikamente überprüft die AR-Brille dabei mithilfe eines KI-Bilderkennungssystems, ob mit den gerichteten Medikamenten alles stimmt. Der Medikamentenstatus des Patienten wird dann in das Managementsystem hochgeladen, wo zusätzliche Parteien wie die Familienangehörigen oder Ärzte des Patienten dies überprüfen und auf mögliche Unregelmäßigkeiten reagieren können. Ein Ansatz wie dieser würde wahrscheinlich gut für weitgehend unabhängige Patienten funktionieren und die Zeit verlängern, bevor sie unter ständige Aufsicht gestellt werden müssen, würde jedoch für die meisten Bewohner von Pflegeheimen nicht funktionieren.

### 3.3 Forschungsfrage

Die sichere und fehlerfreie Medikamentengabe ist ein zentraler Bestandteil der Patientenversorgung. Insbesondere im Pflegealltag, wo Pflegekräfte oft unter Zeitdruck arbeiten, kommt es bei der Vorbereitung und Verabreichung von Medikamenten immer wieder zu Fehlern, die ernsthafte gesundheitliche Folgen für die Patienten haben können. Die bisherigen Arbeiten zeigen, dass es noch keine geeignete Lösung für Pflegeheime kleiner und mittlerer Größe gibt und die Auswahl moderner Konzepte im Medikamentenmanagement sehr gering ist. Aus diesem Grund stellt sich die folgende Forschungsfrage:

**Forschungsfrage:** Wie kann beim Medikamentenrichten die Fehleranzahl reduziert werden?

**Hypothese:** Mit einem technischen System, das mithilfe von KI die Kameradaten eines AR-Systems auswertet und bei Fehlern sofort Rückmeldung gibt.

Die zunehmende Verfügbarkeit von Technologien wie KI und AR bietet vielversprechende Ansätze, um den Medikationsprozess zu verbessern. Durch den Einsatz von AR-Systemen können Pflegekräfte visuelle Unterstützung direkt im Arbeitsumfeld erhalten, während KI-Modelle eine schnelle und präzise Erkennung von Medikamenten und möglichen Fehlern ermöglichen.

Durch die Echtzeitanalyse der Kameradaten und die sofortige Rückmeldung bei Unregelmäßigkeiten soll die Genauigkeit des Medikationsprozesses gesteigert werden. In dieser Arbeit wird untersucht, inwiefern ein solches System entwickelt und implementiert werden kann und ob es dazu beiträgt, die Fehlerquote im Pflegealltag zu senken.

## 4 Konzept

Dieses Kapitel widmet sich zunächst der Erhebung der Anforderungen für ein Medikamentenmanagement-System (MMS) in Pflegeeinrichtungen. Anschließend wird ein Konzept für den Systementwurf vorgestellt, das den identifizierten Anforderungen gerecht wird. Dies umfasst die Betrachtung des Arbeitsablaufs, die Systemabgrenzung sowie die Interaktion mit den Nutzern. Besonderes Augenmerk liegt dabei auf dem Hardware-Aufbau, der effizienten Informationsvermittlung und der Auswahl geeigneter Medikamente.

### 4.1 Anforderungserhebung

Basierend auf den Bedürfnissen und Zielen verschiedener Interessengruppen werden Anwendungsfälle formuliert. Die Anwendungsfälle dienen dazu, die spezifischen Erwartungen und Bedürfnisse der einzelnen Gruppen zu erfassen und bilden die Grundlage für die Ableitung der funktionellen, qualitativen und sicherheitsrelevanten Anforderungen des MMS.

#### 4.1.1 Interessengruppen

Interessengruppen umfassen Einzelpersonen oder Organisationen, die durch das MMS beeinflusst werden oder Einfluss auf dessen Gestaltung nehmen können. Die Perspektiven und Bedürfnisse der im Folgenden beschriebenen Interessengruppen müssen bei der Konzeption des MMS berücksichtigt werden, um eine

erfolgreiche Integration in den Pflegealltag zu gewährleisten. Ihre unterschiedlichen Standpunkte spielen eine entscheidende Rolle für die Konzeption des MMS und tragen dazu bei, dessen Alltagstauglichkeit zu sichern.

- I.1 Pflegekräfte:** Sie sind die Hauptnutzer des MMS und interagieren direkt damit, um Medikamente zu verwalten und die Patienten zu versorgen.
- I.2 Pflegemanagement:** Diese Gruppe entscheidet über die Anschaffung und die Finanzierung neuer Systeme in Pflegeheimen und ist vornehmlich an der Wirtschaftlichkeit interessiert.
- I.3 Patienten:** Die Empfänger der Pflegedienstleistungen profitieren von der verbesserten Medikamentenverwaltung und geringeren Fehlerraten.
- I.4 Systemadministratoren:** Verantwortlich für die Integration und Wartung des MMS, einschließlich der Sicherstellung einer Koexistenz mit bereits bestehenden Systemen.
- I.5 Entwicklerteam:** Verantwortlich für die Entwicklung und kontinuierliche Verbesserung des MMS, einschließlich der Integration von KI-Modellen und Nutzerschnittstellen.
- I.6 Behörden und Regulierungsstellen:** Diese Gruppen sind für die Einhaltung gesetzlicher Vorschriften und Standards im Gesundheitswesen verantwortlich und überwachen die Implementierung neuer Technologien.

Weitere potenzielle Interessengruppen, wie Angehörige der Patienten, Datenschutzbeauftragte oder Investoren in Pflegeeinrichtungen, werden hier nicht aufgeführt, da ihre Interessen weitestgehend mit einem der oben genannten Interessengruppen übereinstimmen.

### **4.1.2 Anwendungsfälle**

Anwendungsfälle sind kurze, einfach verständliche Beschreibungen einer Funktionalität aus der Sicht einer Interessengruppe. Jeder Anwendungsfall folgt dem



Format: »Als [Rolle] möchte ich [Ziel/Wunsch], um [Nutzen].« Die folgenden Anwendungsfälle dienen dazu, die Anforderungen und Erwartungen der einzelnen Gruppen zu definieren. Die Inhalte der Anwendungsfälle sind aus Interviews mit verschiedenen Fachkräften abgeleitet.

**Als Pflegekraft möchte ich...**

- U.1** ... einfach und schnell auf den Medikationsplan eines Patienten zugreifen können, um Zeitverlust beim Richten der Medikamente zu vermeiden.
- U.2** ... zwischen den Medikationsplänen verschiedener Patienten wechseln können, um viele Patienten effizient nacheinander abarbeiten zu können.
- U.3** ... beim Richten der Medikamente eine Übersicht über die noch nicht gerichteten Tabletten haben, um sicherzustellen, dass keine Tablette vergessen wird.
- U.4** ... gewarnt werden, wenn ein Fehler beim Richten der Medikamente geschieht, um diesen verbessern zu können.
- U.5** ... eine klare, intuitive Darstellung der Informationen, um so wenig Zeit wie möglich zu verschwenden.
- U.6** ... mich auf das MMS verlassen können, um nicht gegen die angezeigten Fehler abzustumpfen.
- U.7** ... vor und nach dem Richten der Medikamente andere Aufgaben erledigen, ohne Zeit mit Auf- und Abbau von technischen Systemen zu verschwenden.

**Als Pflegeheimbetreiber möchte ich ...**

- U.8** ... die Kostenersparnis durch die Einführung des MMS nachvollziehen können, um die Wirtschaftlichkeit zu bewerten.
- U.9** ... die Integration des MMS in bestehende Arbeitsabläufe sicherstellen, um keine speziellen Arbeitsplätze einrichten zu müssen.

**U.10** ... sicherstellen, dass das MMS den gesetzlichen Anforderungen entspricht, um rechtliche Risiken zu minimieren.

**Als Patient möchte ich ...**

**U.11** ... sicher sein, dass ich meine Medikamente immer korrekt erhalte, um meine Gesundheit zu schützen.

**U.12** ... sicher sein, dass meine persönlichen Daten sicher und vertraulich behandelt werden, um meine Privatsphäre zu schützen.

**Als Systemadministrator möchte ich ...**

**U.13** ... das MMS einfach installieren und konfigurieren können, um einen reibungslosen Start zu gewährleisten.

**U.14** ... das MMS mit bestehenden IT-Systemen integrieren können, um Datenflüsse zu optimieren.

**U.15** ... Updates und Patches erhalten, um die Sicherheit und Funktionalität des MMS sicherzustellen.

**Als Entwickler möchte ich ...**

**U.16** ... regelmäßig Feedback von den Nutzern erhalten, um das MMS kontinuierlich verbessern zu können.

**U.17** ... sicherstellen, dass das MMS skalierbar ist, um zukünftige Erweiterungen, wie z.B. neue Medikamente, problemlos einbauen zu können.

**Als Behörde möchte ich ...**

**U.18** ... sicherstellen, dass das MMS alle gesetzlichen Anforderungen und Standards erfüllt, um die Patientensicherheit zu gewährleisten.

**U.19** ... sicherstellen, dass die Daten der Patienten sicher und vertraulich behandelt werden, um Datenschutzgesetze einzuhalten.

### 4.1.3 Formulierung der Anforderungen

Die Anforderungen basieren auf einer Analyse der Anwendungsfälle, die die spezifischen Bedürfnisse und Erwartungen der verschiedenen Interessengruppen repräsentieren. Bei Anwendungsfällen wird dabei untersucht, welche funktionellen, qualitativen und sicherheitsrelevanten Anforderungen daraus abgeleitet werden können. Eine nummerierte Liste aller Anforderungen folgt in Kapitel 4.1.4.

#### 4.1.3.1 Funktionelle Anforderungen

Die funktionellen Anforderungen werden direkt aus den primären Aufgaben abgeleitet, die das MMS erfüllen muss, um den Bedürfnissen der Nutzer gerecht zu werden. So führt der Anwendungsfall einer Pflegekraft, die einen schnellen Zugriff auf den Medikationsplan wünscht (U.1), direkt zur Anforderung, dass das MMS diesen Plan während des gesamten Richtprozesses anzeigen muss (A1.1). Die Tatsache, dass zwischen den Medikationsplänen verschiedener Patienten gewechselt werden muss (U.2), ist eine genau solche Anforderung (A1.2). Ebenso ergibt sich aus dem Wunsch, Fehler beim Richten zu vermeiden (U.4), die Notwendigkeit einer Validierungsfunktion für die Medikamentendosetten (A1.3). Die Anforderung, den aktuellen Stand des Richtprozesses visualisieren zu können (A1.4), ist ebenfalls eine direkte Folge der Anwendungsfälle, die eine klare Übersicht während der Arbeit verlangen (U.3).

Des Weiteren ist die Anforderung, dass das MMS vollständige und validierte Medikamentendosetten dokumentieren muss (A1.5), die sich aus dem Bedürfnis der Patienten nach Sicherheit (U.11) und dem Bedürfnis der Betreiber (U.10) und Behörden (U.18) nach Nachverfolgbarkeit ergeben. Die Fähigkeit des MMS, neue Medikamente schnell durch Updates zu erkennen (A1.6), stellt sicher, dass es immer auf dem neuesten Stand bleibt (U.15) und damit auch zukünftig alle Anforderungen erfüllen kann (U.17). Die Anforderung, dass das MMS auf verschiedenen Tischoberflächen funktionieren muss (A1.7), leitet sich aus der Notwendigkeit ab, das MMS in unterschiedlichen Pflegeumgebungen einsetzen

zu können (U.9)(U.13). Schließlich spiegeln die Anforderungen, dass der Arbeitsplatz auf ca.  $1\text{m}^2$ , der üblichen zur Verfügung stehenden Fläche, beschränkt sein soll (A1.8) und das MMS bei Nichtverwendung einen minimalen Platzbedarf aufweisen soll (A1.9), das Bedürfnis nach effizienter Nutzung von Arbeitsraum und Flexibilität wider (U.7).

### 4.1.3.2 Qualitative Anforderungen

Die qualitativen Anforderungen betreffen die Benutzerfreundlichkeit und Effizienz des MMS. Die Anwendungsfälle verlangen, dass die Benutzeroberfläche intuitiv erfassbar ist (A2.1), was eine benutzerfreundliche Gestaltung des MMS erfordert (U.5). Eine typische Pflegekraft muss mindestens genauso schnell mit dem MMS arbeiten können wie ohne (A2.2), und als Steigerung davon soll die Arbeit sogar um mindestens 20% schneller vonstattengehen (A2.3). Diese Anforderungen unterstützen, dass das MMS nicht nur funktional ist, sondern auch die Kosten senkt (U.8).

Die Fähigkeit des MMS, mindestens 100 Medikamente zu differenzieren (A2.4), und wünschenswerterweise sogar 10.000 Medikamente (A2.5), ist eine direkte Antwort auf die Vielfalt der Medikamente, die im Pflegealltag verwaltet werden müssen (U.11). So sind aktuell 50.786 verschreibungspflichtige Medikamente im Einsatz [31], mit 10.000 Medikamenten können jedoch über 99% der Patienten abgedeckt werden. Nur wenn die meisten Medikamente tatsächlich vom MMS erkannt werden können, kann es die Grundfunktion erfüllen (U.4). Die Reduktion von falsch-negativen Ausgaben auf unter 1% (A2.6) und die Minimierung von falsch-positiven Ausgaben (A2.7) stellen sicher, dass die Pflegekräfte nicht mit falschen Informationen überflutet werden (U.6). Die Anforderung, dass das MMS ohne Technikenkenntnisse in Betrieb genommen werden kann (A2.8), wird aus dem Bedürfnis nach einer schnellen und reibungslosen Implementierung abgeleitet (U.13). Schließlich spiegelt die Fähigkeit des MMS, Statistiken und Feedback der Nutzer an die Entwickler zu übertragen (A2.9), das Bedürfnis nach kontinuierlicher Verbesserung und Anpassung wider (U.16). Die Möglichkeit, das MMS mit minimalem Aufwand auf- und abbauen zu können (A2.10), ergänzt Anforderung

A1.9, indem sie die Mobilität und Anpassungsfähigkeit des MMS unterstützt (U.7).

#### **4.1.3.3 Sicherheitsanforderungen**

Die Sicherheitsanforderung, dass das MMS keine automatisierten Entscheidungen ohne Zustimmung des Nutzers treffen darf (A3.1), wird aus ethischen und rechtlichen Überlegungen abgeleitet, die in den Anwendungsfällen von Behörden und Regulierungsstellen implizit enthalten sind und eine Zertifizierungspflicht umgehen (U.10)(U.18). Die Verwendung zertifizierter Elektronik (A3.2) erfüllt hierbei das gleiche Ziel. Die Anforderungen, dass die Datenverarbeitung den rechtlichen Ansprüchen der Patienten, wie z.B. die DSGVO, gerecht werden muss (A3.3) und dass keine Daten auf fremden Servern gespeichert werden sollen (A3.4), gewährleisten den Schutz der Patientendaten und die Einhaltung von Datenschutzgesetzen (U.12)(U.19).

#### **4.1.4 Anforderungen**

Alle Anforderungen sind in der folgenden Liste zusammengefasst und mit einem eindeutigen Identifikator versehen.

##### **A1 Funktionelle Anforderungen**

- A1.1** Das MMS muss den Medikamentenplan während des ganzen Richtprozesses anzeigen.
- A1.2** Das MMS muss zwischen verschiedenen Patienten wechseln können.
- A1.3** Das MMS muss eine gerichtete Medikamentendosette anhand eines vorgegebenen Medikamentenplans validieren können.
- A1.4** Das MMS muss die aktuell fehlenden, falschen und korrekten Medikamente während des Richtprozesses visualisieren können.

**A1.5** Das MMS muss vollständige und validierte Medikamentendosetten dokumentieren.

**A1.6** Das MMS soll über Updates in kürzester Zeit neue Medikamente erkennen können.

**A1.7** Das MMS muss auf verschiedenen Tischoberflächen funktionieren.

**A1.8** Der Arbeitsplatz muss auf ca. 1m<sup>2</sup> beschränkt sein.

**A1.9** Das MMS soll bei Nichtverwendung einen minimalen Platzbedarf aufweisen.

## **A2 Qualitative Anforderungen**

**A2.1** Das Interface soll intuitiv erfassbar sein.

**A2.2** Eine typische Pflegekraft muss mit dem MMS mindestens genauso schnell arbeiten wie ohne.

**A2.3** Eine typische Pflegekraft soll mit dem MMS mindestens 20% schneller arbeiten als ohne.

**A2.4** Das MMS muss über 100 Medikamente differenzieren können.

**A2.5** Das MMS soll über 10.000 Medikamente differenzieren können.

**A2.6** Die Rate falsch-negativer Ausgaben muss unter 1% liegen

**A2.7** Die Rate falsch-positiven Ausgaben soll minimiert werden.

**A2.8** Das MMS soll ohne Technikenntnisse in Betrieb genommen werden können.

**A2.9** Das MMS soll Statistiken und Feedback der Nutzer zurück an die Entwickler übertragen können.

**A2.10** Das MMS soll mit minimalem Aufwand auf- und abgebaut werden können.

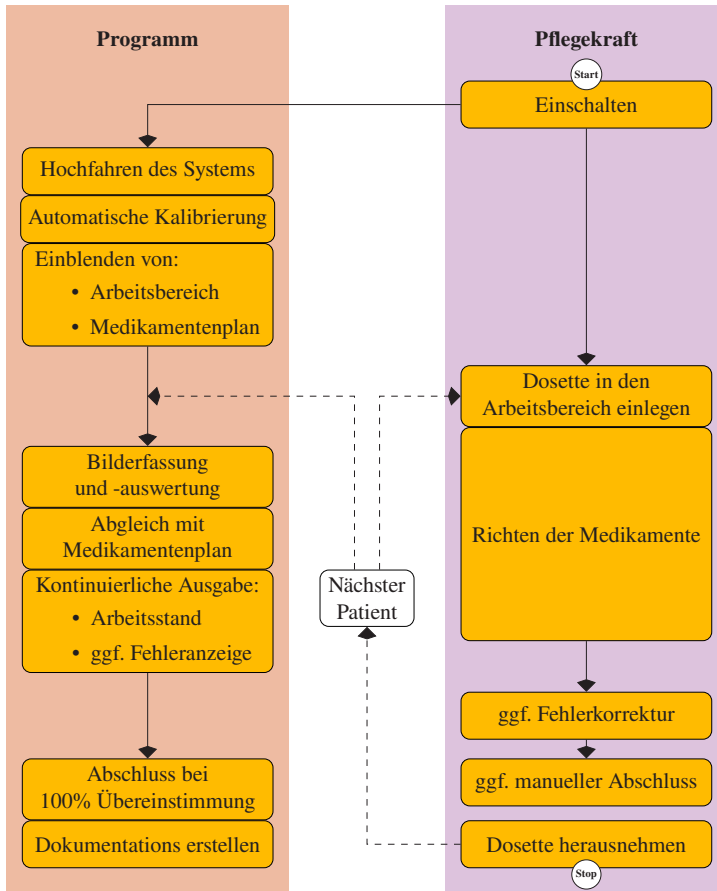
### **A3 Sicherheitsanforderungen**

- A3.1** Das MMS darf keine automatisierten Entscheidungen ohne Zustimmung des Nutzers treffen.
- A3.2** Es muss für Endnutzer zertifizierte Elektronik im MMS verwendet werden.
- A3.3** Die Datenverarbeitung muss den rechtlichen Ansprüchen der Patienten gerecht werden.
- A3.4** Es sollen keine Daten auf fremden Servern gespeichert werden.

## **4.2 Arbeitsablauf**

Der Arbeitsablauf mit dem MMS startet, wenn die Pflegekraft das System einschaltet (siehe Abbildung 4.1). Das Programm des MMS fährt daraufhin hoch und beginnt vollautomatisch mit der Kalibrierung. Sobald alle Komponenten einsatzbereit sind, wird der Medikamentenplan des ersten Patienten und der Arbeitsbereich angezeigt. Daraufhin kann die Pflegekraft die erste Medikamentendosette in den Arbeitsbereich einlegen und mit dem Richten der Medikamente beginnen. Der Arbeitsbereich ist dabei ein Teil des Tisches, auf dem die Pflegekraft arbeitet und der von der Kamera aufgenommen wird. Während des ganzen Richtprozesses nimmt das MMS kontinuierlich Bilder auf und erkennt die in der Dosette platzierten Medikamente. Diese werden mit dem Medikamentenplan abgeglichen und auf Fehler überprüft. Über den angezeigten Medikamentenplan wird der Pflegekraft sofort Rückmeldung gegeben, ob gegebenenfalls ein Fehler passiert ist. Sobald das MMS eine 100% Übereinstimmung zwischen Medikamentenplan und erkannten Medikamenten feststellt, wird der Prozess abgeschlossen und ein Bild wird zusammen mit den erkannten Medikamenten zur Dokumentation abgespeichert. Die Pflegekraft hat dabei die Möglichkeit, das System zu überstimmen (z.B. bei Fehlerkennungen oder nicht im Pflegemanagement-System hinterlegten Änderungen im Medikamentenplan) und den Prozess so manuell abzuschließen. Nachdem die

Dosette zur Übergabe an den Patienten aus dem Arbeitsbereich herausgenommen wird, kann die Pflegekraft entweder das Programm beenden oder mit dem nächsten Patienten fortfahren, indem eine neue Dosette in den Arbeitsbereich eingelegt wird. Das MMS lädt dazu automatisch den nächsten Patienten und zeigt dessen Medikamentenplan an.



**Abbildung 4.1** Zeitlicher Ablauf beim Medikamentenrichten mit dem Medikamentenmanagement-System



## 4.3 Systementwurf

Das Datenflussdiagramm in Abbildung 4.2 veranschaulicht die Verarbeitungsschritte zur Laufzeit. Die Richtstation hat einen Einplatinenrechner als Recheneinheit und ist direkt am Arbeitsplatz der Pflegekraft im Einsatz. Im Gegensatz dazu kann der Server an einem anderen Ort stehen und Anfragen mehrerer Richtstationen bearbeiten. An der Richtstation sendet die Kamera eine konstante Zeitreihe von Bildern an den Einplatinenrechner. Auf diesen Bildern ist der aktuelle Stand des Einsortierprozesses zu sehen. Der Einplatinenrechner wählt die Zeitabschnitte, zu denen eine neue Anfrage an den Server gestellt werden soll. Dabei hängt die Frequenz der Anfragen von der Geschwindigkeit der Vorverarbeitung auf dem Einplatinenrechner, und von der Verarbeitungszeit zwischen dem Stellen der Anfrage und dem Erhalten der Antwort ab. Die nächste Anfrage wird erst gestellt, wenn sowohl deren Vorverarbeitung abgeschlossen ist, als auch die Antwort des Servers auf die letzte Anfrage eingetroffen ist. Diese Anfrage enthält das aktuelle Bild, sowie potentiell eine Abfrage des nächsten Medikamentenplans. Der Medikamentenplan wird abgefragt, falls der aktuelle Einsortierprozess als abgeschlossen erkannt wird oder der Nutzer dies durch eine explizite Eingabe anfordert. Der Server verwaltet die Schnittstelle zum Pflegemanagement-System, über welche die Patientendaten abgefragt werden. Bevor es im KI-Netz evaluiert werden kann, wird das Bild auf das richtige Format umgewandelt und eine Normalisierung der Helligkeit wird vorgenommen. Nachdem das KI-Netz die Ergebnisse in Form von Erkennungswahrscheinlichkeiten mit korrespondierenden Dosetten-Koordinaten liefert, berechnet der Server daraus die Zuordnung der einzelnen Medikamente zu den Zeitabschnitten »Morgens«, »Mittags«, »Abends« und »Nachts«. Dafür wird eine euklidische Distanz zwischen dem Medikament und der Beschriftung auf der Medikamentendosette bestimmt und die Zuordnung zum Zeitabschnitt mit der geringsten Distanz vorgenommen. Diese Daten werden als Ist-Wert zurück an den Einplatinenrechner gesendet und falls dieser neue Patientendaten angefragt hat, werden zusätzlich die neuen Soll-Werte übermittelt. Der Einplatinenrechner füllt die Nutzerschnittstelle mit Informationen und zeigt diese dem Nutzer.

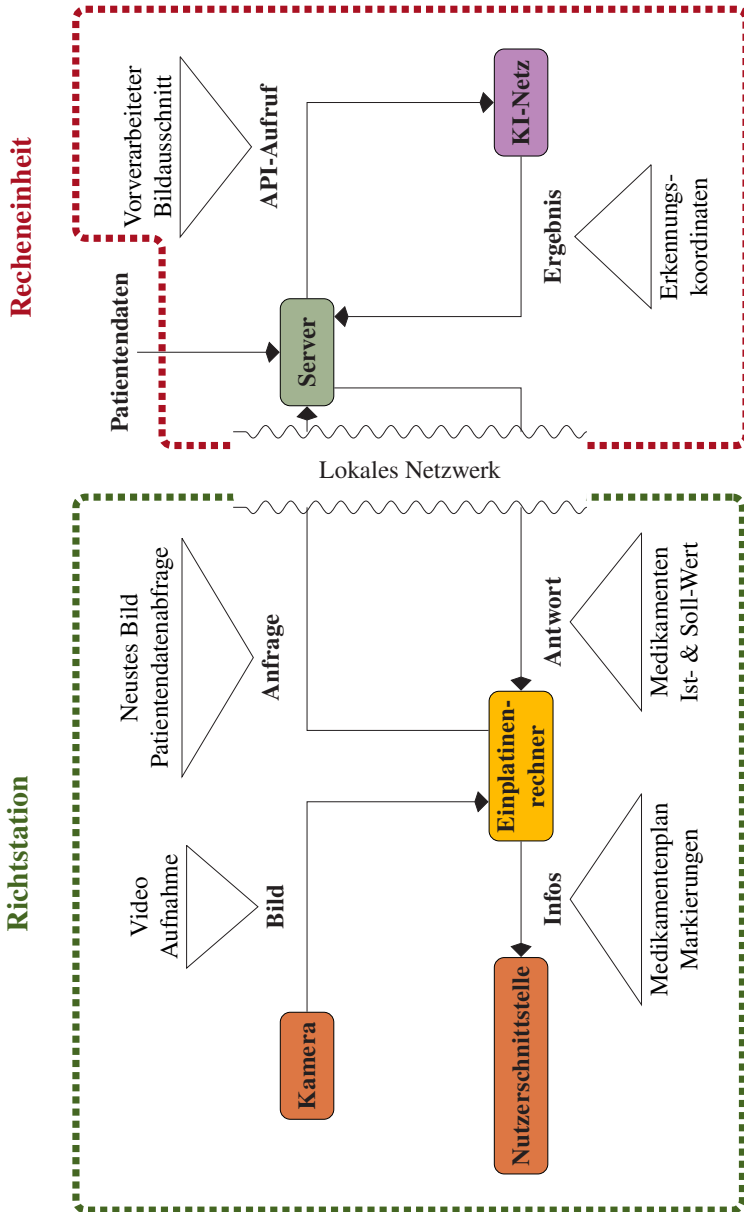
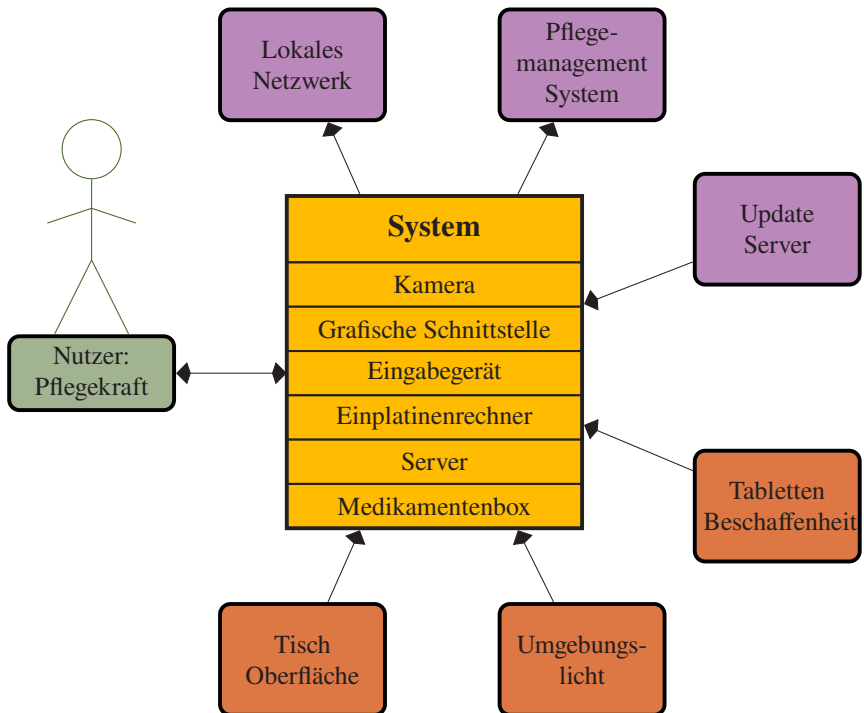


Abbildung 4.2 Datenflussdiagramm des Medikamentenmanagement-Systems

### 4.3.1 Bestandteile und Einflussgrößen

In diesem Abschnitt soll die Abgrenzung zwischen dem MMS und seiner Umgebung genauer betrachtet werden. Hierbei liegt nicht nur die Interaktion mit der Umgebung als solche im Fokus, sondern auch die genaue Systemgrenze und welche Komponenten als Teil des MMS aufgefasst werden. Wichtige Bestandteile



**Abbildung 4.3** Abgrenzung zwischen Medikantenmanagement-System und Einflussgrößen

des MMS sind dabei die Hardwarekomponenten, die für die Funktion des MMS notwendig sind. Dazu gehören zum einen die Recheneinheiten wie der Einplattinenrechner und der Server, die in ständiger Kommunikation zueinander stehen.

Andererseits sind die Schnittstellengeräte wie Kamera, Beamer und Eingabegerät ebenfalls von elementarer Wichtigkeit, um die volle Funktionalität des MMS zu gewährleisten. Des Weiteren soll hier noch die Entscheidung hervorgehoben werden, dass die Medikamentendosette als Teil des MMS betrachtet wird, sodass im Zuge dieser Arbeit die Form und Beschaffenheit frei gewählt werden kann. Abbildung 4.3 zeigt die Abgrenzung zwischen dem MMS und der Umgebung, mit der das MMS interagiert oder die das MMS beeinflusst. Die Interaktion, für die das MMS primär entwickelt wird, ist die Interaktion mit der Pflegekraft. Diese bedient das MMS mit dem Eingabegerät und indem die Medikamente in die Medikamentendosette einsortiert werden und erhält über den Beamer Informationen und Rückmeldung zu den Medikamenten. Das Pflegemanagement-System und das Lokale Netzwerk sind Infrastruktur, die bereits in den Pflegeheimen vorhanden ist und die vom MMS für den reibungsfreien Ablauf genutzt werden. Der Update Server stellt eine konzeptionelle Möglichkeit dar, neue KI-Netze und andere Softwareupdates einzuspeisen. Die Entwicklung einer solchen Funktion wird im Rahmen dieser Arbeit jedoch nicht erörtert und somit liegt diese außerhalb des beschriebenen MMS. Außer den gewünschten Interaktionen gibt es weitere Elemente, die einen vorwiegend negativen Einfluss auf das MMS nehmen können. Diese sind in Abbildung 4.3 in rot dargestellt. Vor allem die Medikamente sind hierbei ein erheblicher Einflussfaktor, der komplett außerhalb des kontrollierbaren Bereiches liegt. Aber sowohl die Lichtbedingungen, als auch Farbe und Reflexionsgrad der Oberfläche, auf der die Pflegekraft arbeitet, müssen berücksichtigt werden.

## 4.4 Repräsentation

Eine intuitive und klare Darstellung ist essenziell, um die Effizienz des MMS zu gewährleisten. Dabei spielen sowohl die Art und Weise der Informationsdarstellung als auch die Auswahl der Medikamente eine zentrale Rolle.

### 4.4.1 Informationsvermittlung

Seit dem 01.10.2016 haben Patienten nach § 31a im Sozialgesetzbuch Fünf [32] einen Anspruch auf die Ausstellung eines bundeseinheitlichen Medikationsplans. Der Aufbau wird, laut Gesetz, unter anderem von der Kassenärztlichen Bundesvereinigung festgelegt [33]. Abbildung 4.4 zeigt einen Musteraufbau, in dem die für das korrekte Richten der Medikamente essentiellen Informationen durch eine rote Markierung hervorgehoben werden. Diese zentralen Informationen umfassen hauptsächlich zwei Elemente: Zum einen den Namen des Medikaments (Punkt A in Abbildung 4.4). Dieser kann entweder der Handelsname oder der generische Name des Medikaments sein.

Medikationsplan		für: Dr. Michaela Freifrau von Musterhausen		geb. am: 13.12.1936			
Seite 1 von 1		ausgedruckt von: Dr. Manfred Überall Hauptstraße 55, 01234 Am Ort Tel: 04562-12345 E-Mail: m.ue@prokura-ueberall.de		Gew.: 85 kg Geschl.: w Allerg./Unv.: Penicillin ausgedruckt: 01.04.2023 12:00			
Wirkstoff	Handelsname	Stärke	Form	Einheit	Hinweise	Grund	
Ramipril		5 mg	Tabl	1 0 0 0	Stück	Während der Mahlzeit	Bluthochdruck
Hydrochlorothiazid	HCT 1A PHARMA 25 MG TABL	25 mg	Tabl	1 0 0 0	Stück	während der Mahlzeit	Bluthochdruck
Clopidogrel	PLAVIX	75 mg	Tabl	0 0 1 0	Stück	während der Mahlzeit	art. Verschluss
Simvastatin	SIMVA BASICS 20 mg	20 mg	Tabl	0 0 1 0	Stück	nach der Mahlzeit	erhöhte Blutfette
Insulin-Isophan	PROTAPHANE PENFILL ZAM	300 IE/3 ml	Ampulle	20 0 10 0	IE	subkutan	Diabetes
<b>Bedarfsmedikation</b>							
Glycerotriminrat (Nitroglycerin)	NITROLINGUAL Spray	0,4 mg	Spray	max. 3	Hub	akut	Herzschmerzen
Diphenhydramin	VIVINOX SLEEP SCHLAFTAB ST	50 mg	Tabl	0 0 0 1	Stück	bei Bedarf	Schlaflosigkeit
Kombipräparat	SINUPRET EXTRACT		Tabl	1 1 1 0	Stück		Erkältung/ Nasennebenhöhlen
<b>zeitlich befristet anzuwendende Medikamente</b>							
Cefaclor	CEFACTOR STADA 500 mg	500 mg	Kapsel	1 1 1 0	Stück	alle 8 Stunden einnehmen	Nebenhöhlenentzündung
Sofort mit Einnahme beginnen, für 10 Tage (bis 10.03.2023)							

**Wichtige Angaben**  
Bitte messen Sie Ihren Blutdruck täglich!

Für Vollständigkeit und Aktualität des Medikationsplans wird keine Gewähr übernommen.  
de-DE-Version 2.7

"Medikationsplan-Factory" ihr EDV-Partner

**Abbildung 4.4** Der bundeseinheitlichen Medikationsplan mit den wichtigsten Informationen in Rot Herausgegeben von der Kassenärztlichen Bundesvereinigung[33].

Zum anderen ist die Dosierung des Medikaments (Punkt B in Abbildung 4.4) von entscheidender Bedeutung. Diese gibt an, wie viel des Medikaments zu welchen Tageszeiten verabreicht werden soll. Dabei sind sowohl die Menge als auch die Einheit und der Formfaktor des Medikaments – ob Tablette, Kapsel oder eine andere Form – wichtig zu beachten.

Ein weiteres wichtiges Element ist das Hinweissfeld, das die Möglichkeit bietet, zusätzliche Informationen in Form von freiem Text einzufügen. Dies kann beispielsweise nützlich sein, um anzugeben, ob ein Medikament nur optional einzunehmen ist oder unter bestimmten Bedingungen weggelassen werden kann.

Für die eindeutige Identifizierung des Patienten, der die Medikamente erhalten soll, sind neben dem Namen des Patienten (Punkt D in Abbildung 4.4) auch die Zimmernummer oder die Patientennummer wichtig, um Verwechslungen zu vermeiden.

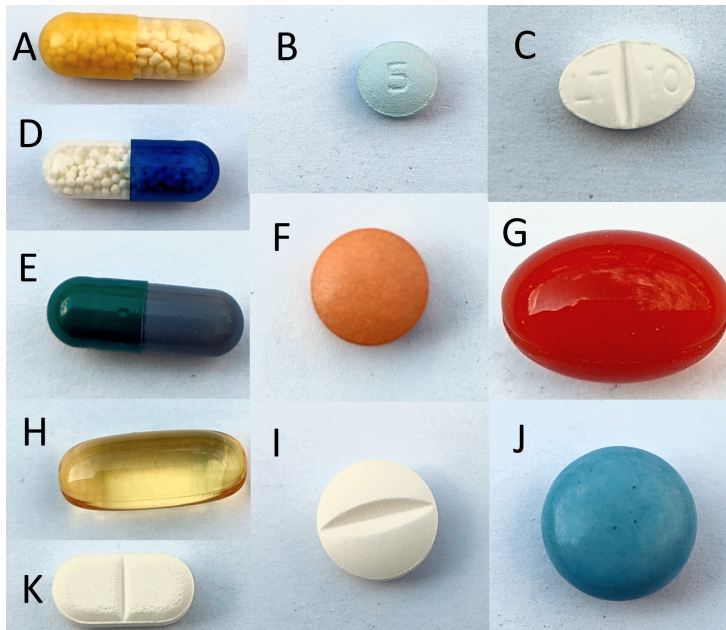
Beim Design der Nutzerschnittstelle für den Medikationsplan ist es zudem essenziell, die in Kapitel 2.3 erwähnten psychologischen Prinzipien zu berücksichtigen. Eine gut gestaltete Schnittstelle unterstützt das medizinische Fachpersonal dabei, zu erfassen, welche Medikamente in welcher Menge noch zu richten sind und bei welchen die Vorbereitung bereits abgeschlossen ist. Eine Möglichkeit, dies zu erreichen, ist die Verwendung von Farbcodes. So könnte beispielsweise eine grüne Markierung anzeigen, dass die Medikamentenvorbereitung für ein bestimmtes Medikament abgeschlossen ist, während eine gelbe oder rote Markierung darauf hinweist, dass hier noch Handlungsbedarf besteht. Abschließend ist es wichtig, die erforderliche Handlung aus der Markierung ableiten zu können. So können gleichzeitig Fehler minimiert und die Effizienz gesteigert werden.

### 4.4.2 Medikamentenauswahl

Bei der repräsentativen Auswahl der Medikamente sind zwei wesentliche Gesichtspunkte zu berücksichtigen. Zum einen ist es wichtig, die Diversität der in der Realität vorkommenden Medikamente zu repräsentieren. Dies beinhaltet eine

Bandbreite an Formfaktoren, sowie deren unterschiedliche Farbvariationen, um ein realistisches Spektrum abzudecken. Darüber hinaus ist es notwendig, potenzielle Verwechslungen und Ähnlichkeiten zwischen Medikamenten zu berücksichtigen, da diese in der Praxis eine signifikante Fehlerquelle darstellen können.

Zum anderen muss der Entwicklungs- und Trainingsaufwand in einem definierten Rahmen gehalten werden. Das bedeutet, dass die Anzahl der unterschiedlichen Medikamente während der Entwicklung bewusst begrenzt wird, um den Aufwand gering zu halten. Dies ermöglicht es, mehr Iterationen des Entwicklungsprozesses



**Abbildung 4.5** Übersicht der verwendeten Medikamente

A - apoday Vitamin C (Vitamin C); B - Deslora - 1A (Desloratadin); C - Lorano akut (Loratadine); D - Ambroxol 75 retard Heumann (Ambroxol); E - Imodium akut (Loperamid); F - Angocin (Kräuterextrakt); G - Femibion 3 (Vitamine); H - Doppel Herz Omega-3 Konzentrat (Omega-3 Fettsäuren); I - Paracetamol 500 - 1A (Paracetamol); J - Baldrian - Ratiopharm (Baldrian Wurzelextrakt); K - Loperamid akut aristo (Loperamid)

in kürzerer Zeit durchzuführen, ohne dabei die Qualität oder die Relevanz der Daten zu kompromittieren.

Abbildung 4.5 veranschaulicht die ausgewählten Medikamente, die die folgenden Formfaktoren aufweisen:

- **Tabletten:** Tabletten sind eine der am weitesten verbreiteten Formen von Medikamenten und bieten den Vorteil einer genauen Dosierung, langer Haltbarkeit und einfachen Lagerung. Sie kommen in verschiedenen Größen, Formen und Farben vor, was eine individuelle Anpassung an therapeutische Bedürfnisse ermöglicht. Bei der Auswahl sind Tabletten unter den Punkten B, C, I und K vertreten.
- **Kapseln:** Kapseln erlauben die Verpackung von Medikamenten in einer gelartigen Hülle, was sie besonders für schlecht schmeckende oder übelkeitsauslösende Wirkstoffe geeignet macht. Sie können sowohl feste als auch flüssige Inhalte aufnehmen und sind in einer Vielzahl von Größen und Farbkombinationen erhältlich. Kapseln sind unter den Punkten A, D, E, G und H vertreten.
- **Dragees:** Dragees bieten durch ihre glatte, oft zuckerhaltige Ummantelung eine angenehme Einnahme. Sie sind besonders bei Medikamenten beliebt, die im Mund nicht sofort zerfallen sollen, oder um einen bitteren Geschmack zu maskieren. Die Auswahl enthält Dragees unter den Punkten F und J.



## 5 Prototyp

In diesem Kapitel wird die Entwicklung des ersten Prototyps auf der Microsoft HoloLens 2 vorgestellt. Dies umfasst die Auswahl des passenden KI-Netzwerks und die Programmierung der Nutzerschnittstelle und des Servers.

### 5.1 Serverhardware

Für die Implementierung des MMS und das Training der neuronalen Netze wird folgende Serverhardware eingesetzt:

- **Prozessor:** Intel Core i7-14700k mit 20 Kernen.
- **Grafikkarte:** Nvidia RTX A6000 mit 10742 Cuda Kernen und 48 GB Videospeicher.
- **Speicher:** 32 GB Arbeitsspeicher und 1 TB SSD-Festplatte.

### 5.2 KI-Modellauswahl

Zu Beginn muss bei der Umsetzung des Konzepts die Frage beantwortet werden, ob mithilfe von künstlicher Intelligenz Medikamente automatisiert erkannt werden können. Um das richtige KI-Modell auszuwählen, sind zwei Schritte notwendig. Im ersten Schritt wird eine breite Auswahl an verschiedenen vortrainierten Netzen zusammengestellt. Diese werden anhand eines einheitlichen frei verfügbaren

Datensatzes evaluiert. Aus den besten werden dann Kandidaten für die Medikamentenerkennung ausgewählt. Im zweiten Schritt werden diese Kandidaten dann einem Test unterzogen. Dabei wird ein neuer Datensatz mit Medikamenten erstellt und zur Evaluation verwendet.

### 5.2.1 Vergleich verschiedener Architekturen

Für einen ersten Vergleich verschiedener Modelle wird eine Auswahl an Architekturen herangezogen und mit einem standardisierten Datensatz geprüft. Diese Auswahl besteht aus den in Kapitel 2.2.3 vorgestellten Alternativen zu YOLO, sowie aus den vergangenen Generationen von YOLO. Als Datensatz wird der MS-COCO (test-dev 2017) [34] verwendet, um die jeweiligen Architekturen zu testen. Als Testparameter wird die mittlere Genauigkeit bei einem Mindestwert der Schnittmenge über der Vereinigungsmenge von 50% gewählt. Außerdem wird die Geschwindigkeit in fps (Frames pro Sekunde) gemessen. Tabelle 5.1 zeigt, dass YOLOv5 die mit Abstand beste Kombination aus mittlerer Genauigkeit und

Modell	Bildgröße	$\varnothing PR_{50}$	fps
YOLOv3 Darknet-53	416 * 416	55,3 %	35
YOLOv4 CSPDarknet-53	512 * 512	64,9 %	31
YOLOv5s	640 * 640	68,9 %	140
LRF VGG-16	300 * 300	51,5 %	76,9
RFBNet VGG-16	300 * 300	49,3 %	66,7
SSD VGG-16	300 * 300	43,1 %	43
RefineDet VGG-16	320 * 320	49,2 %	38,7
EfficientDet-D0	512 * 512	52,2 %	62,5

**Tabelle 5.1** Vergleich verschiedener Lokalisierungsnetzwerke auf dem MS-COCO-Datensatz (test-dev 2017) [34].

Bildgröße: Dimensionen des Eingangsbildes in Pixel.

$\varnothing PR_{50}$ : mittlere Genauigkeit bei einem Mindestwert der Schnittmenge über der Vereinigungsmenge von 50%.

fps: Bearbeitungsgeschwindigkeit in Frames pro Sekunde.

Geschwindigkeit aufweist. Andere Ansätze können lediglich mit älteren YOLO Architekturen konkurrieren, aber haben auch hier lediglich einen Vorteil in der Geschwindigkeit.

## 5.2.2 Test im Medikamentenmanagement

Da die YOLOv5-Architektur im ersten Test einen so klaren Vorsprung vor anderen Architekturen aufweist, werden im zweiten Schritt lediglich die verschiedenen YOLOv5-Unterversionen evaluiert. Dafür werden zunächst spezifische Datensätze

Datensatz	Bilderanzahl	Objektanzahl	Augmentierung
Klein	400	1.931	<ul style="list-style-type: none"> <li>keine weitere Augmentierung</li> </ul>
Groß	1.200	5.820	<ul style="list-style-type: none"> <li>50% Chance horizontal zu spiegeln</li> <li>50% Chance vertikal zu spiegeln</li> <li>Verdrehung um bis zu <math>\pm 15^\circ</math></li> <li>Helligkeitsanpassung um bis zu <math>\pm 25\%</math></li> <li>Belichtungsanpassung um bis zu <math>\pm 25\%</math></li> </ul>

**Tabelle 5.2** Beschreibung von zwei Testdatensätzen zur Evaluation mit Medikamentenbildern

erstellt, die verschiedene Arten von Medikamenten in unterschiedlichen Umgebungen abbilden. Damit sollen die Herausforderungen der realen Anwendungsszenarien simuliert werden. In dieser ersten Datenerhebung wird noch nicht auf eine vollständige Repräsentation aller möglichen Formen und Größen geachtet.

Tabelle 5.2 zeigt die Zusammenstellung der beiden Datensätze. Diese beinhalten jeweils Aufnahmen von 5 verschiedenen Arten von Tabletten und der Medikamentendosette. Dabei enthält bereits der kleine Datensatz mehr als 300 Beispiele von jeder der 6 Objektklassen. Im großen Datensatz kommt jede Objektklasse mehr als 1000-mal vor. Der große Datensatz wird zusätzlich noch mit verschiedenen Augmentierungsmethoden synthetisch erweitert, was die effektive Anzahl an Bildern und Objekten verdreifacht.

Modell	Datensatz	Präzision	Trefferquote	$\emptyset PR_{50}$
YOLOv5s	Klein	0.9653	0.9688	0.9701
YOLOv5m	Klein	0.9960	0.9974	0.9956
YOLOv5l	Klein	0.9214	0.9180	0.9464
YOLOv5s	Groß	0.9903	0.9876	0.9929
YOLOv5m	Groß	0.9945	0.9888	0.9941
YOLOv5l	Groß	0.9953	0.9906	0.9948

**Tabelle 5.3** Ergebnisse des Vergleichs von drei verschiedenen YOLOv5 Architekturen auf zwei Datensätzen (siehe Tabelle 5.2)

Um die einzelnen Modelle zu testen, werden die drei vielversprechendsten Untertypen von YOLOv5 verwendet: YOLOv5s, YOLOv5m und YOLOv5l. Alle drei werden mit dem MS COCO (test-dev 2017) Datensatz vortrainiert [34]. Anschließend wird jedes Modell mit jedem Datensatz in der Feinabstimmung nachtrainiert. Feinabstimmung ermöglicht einem System, Wissen und Fähigkeiten, die in vorherigen Bereichen/Aufgaben erlernt wurden, zu erkennen und auf neue Bereiche/Aufgaben anzuwenden. In diesem Fall ermöglicht es dem Modell, den Aufbau und die Zusammensetzung von verschiedenen Objekten bereits aus dem MS COCO (test-dev 2017) Datensatz zu erlernen. Anschließend muss das

Modell lediglich den spezifischen Aufbau und das konkrete Aussehen der relevanten Medikamente und der Medikamentendosette aus den neuen Datensätzen erlernen.

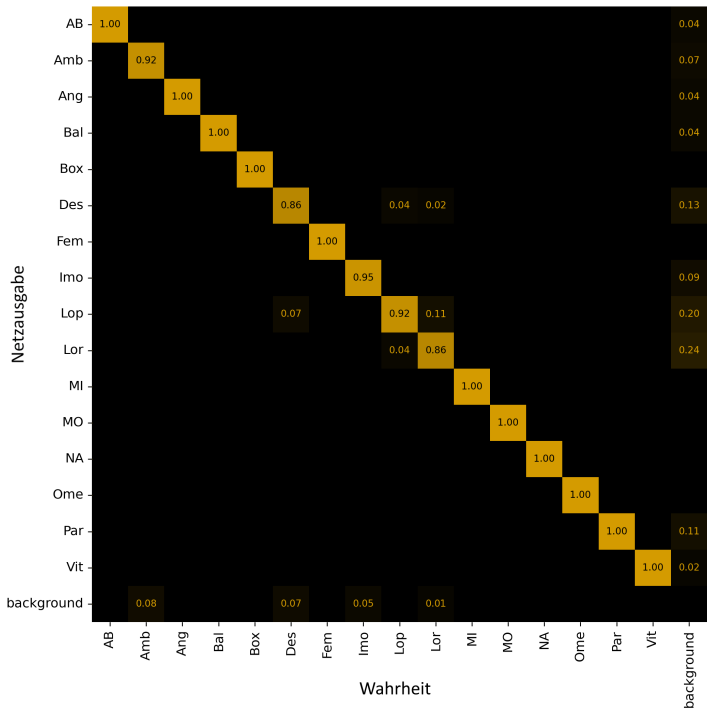
Die Ergebnisse, die in der Tabelle 5.3 dargestellt sind, zeigen deutlich die Überlegenheit der YOLOv5m und YOLOv5l Modelle im großen Datensatz, was auf die umfassenden Augmentierungen und die höhere Anzahl von Trainingsdaten zurückzuführen ist. Besonders bemerkenswert ist die Steigerung der Präzision und Trefferquote im Vergleich zu den kleineren Datensätzen, was die Bedeutung der Datenvielfalt für das Training von Deep-Learning-Modellen unterstreicht. Dies zeigt sich auch in der Tatsache, dass das YOLOv5l Modell im kleinen Datensatz eine deutliche Schwäche verglichen mit YOLOv5s und YOLOv5m aufweist. Da der Unterschied zwischen YOLOv5m und YOLOv5l im großen Datensatz nicht signifikant ist und im Medikamentenmanagement nicht immer auf große Datensätze zurückgegriffen werden kann, fällt die Wahl deswegen auf YOLOv5m für alle weiteren Entwicklungsschritte.

### 5.2.3 Training von YOLOv5

Um das Modell für den Betrieb zu verwenden, wird ein Datensatz aus den in Kapitel 4.4.2 vorgestellten Medikamenten zusammengestellt. In diesem Datensatz befinden sich 1614 Bilder mit 11331 Objekten. Das häufigste Objekt ist dabei die Medikamentendosette mit 1536 Vorkommen, das seltenste Objekt ist das Medikament Vitamin C mit 309 Vorkommen.

Zum Training wird der Datensatz unterteilt in einen Trainingsanteil, einen Validierungsanteil und einen Testanteil. Der Trainingsanteil wird verwendet, um die Gewichte des Modells zu ermitteln, und beträgt 75% des Datensatzes. Der Validierungsanteil wird verwendet, um den optimalen Zeitpunkt zu bestimmen, an dem das Training abgeschlossen ist. Nach jeder Epoche wird mit den Validierungsdaten die Genauigkeit des Modells gemessen. Der Anteil der Validierungsdaten beträgt 15%. Die Testdaten werden während des Trainingsprozesses überhaupt nicht verwendet, sondern werden nach Abschluss des Trainings benutzt, um die

Qualität des Modells zu bestimmen. Nachdem die Testdaten eingesetzt werden, darf das Modell nicht mehr angepasst werden, da sie sonst ihre Aussagekraft verlieren. Alle hier präsentierten Ergebnisse sind mit den Testdaten gemessen. Der Anteil der Testdaten liegt bei den übrigen 10%.



**Abbildung 5.1** Konfusionsmatrix der Testdaten von YOLOv5 mit 11 Medikamenten (siehe Kapitel 4.4.2) und der Dosette (Box) mit 4 Fächern (MO, MI, AB, NA)

Abbildung 5.1 zeigt die Konfusionsmatrix nach dem Training mit diesem Datensatz. Die Konfusionsmatrix stellt in jeder Spalte dar, wie häufig die in der Spalte angegebene Klasse als eine bestimmte Klasse (Zeilen) erkannt wird. Die Hauptdiagonale repräsentiert dabei die jeweils korrekten Zuordnungen. Die Klasse »Background« ist keine eigene Kategorie, sondern eine Fehlklassifizierung

eines irrelevanten Details im Bild (z.B. eine Reflexion). Es ist zu erkennen, dass sowohl die Medikamentendosette (BOX) als auch die einzelnen Fächer der Dosette in allen Fällen korrekt erkannt werden. Auch bei den Medikamenten ist die Erkennungsrate sehr hoch, wobei gelegentlich eine Verwechslung der weißen Tabletten »Deslora« (Des), »Lorano akut« (Lor) und »Loperamid akut« (Lop) untereinander auftritt.

## **5.3 AR-Brille**

Für die Implementierung der Augmented Reality (AR)-Funktionalität in diesem Projekt fiel die Wahl auf die Microsoft HoloLens 2. Dieses AR-Gerät wurde ausgewählt, weil es eine der fortschrittlichsten verfügbaren Technologien auf dem Markt darstellt.

### **5.3.1 Technische Vorteile der HoloLens**

Die Microsoft HoloLens 2 ist ein eigenständiges Mixed Reality-Headset, das ohne zusätzliche externe Hardware funktioniert. Sie verfügt über ein visuelles Displaysystem mit holografischen Linsen, die eine 2K-Auflösung pro Auge bieten, sowie eine Reihe von Lautsprechern, die Geräuschquellen virtuell überall im Raum positionieren können.

Die HoloLens 2 nutzt einen Flugzeit-Tiefensensor (englisch Time-of-Flight oder ToF), Infrarotkameras für die Augenpositionsbestimmung und eine Vielzahl von Mikrofonen, um die Umgebung, die Lage und die Interaktionen des Nutzers aufzuzeichnen. Der eingebaute Qualcomm Snapdragon 850 Prozessor und die holografische Recheneinheit (Holographic Processing Unit, HPU) bieten die nötige Rechenleistung für anspruchsvolle AR-Anwendungen. Zusätzlich ist die HoloLens 2 mit einem Akkupack ausgestattet, der je nach Auslastung 2-6 Stunden Betrieb ermöglicht.

### 5.3.2 Lokale Verarbeitung der KI-Modelle

Vor der Entwicklung einer Architektur, die die Datenverarbeitung zwischen Server und Nutzer-Anwendung aufteilt, wird zunächst überprüft, ob eine lokale Auswertung auf der HoloLens 2 ohne Server möglich ist. Um ein neuronales Netz auf der HoloLens 2 zum Laufen zu bringen, muss es zunächst auf die spezifische Hardwarearchitektur optimiert werden. Dies erfolgt durch die Verwendung der Bibliothek ONNX (Open Neural Network Exchange), die es ermöglicht, ein vortrainiertes Modell in ein kompatibles Format zu konvertieren. Anschließend kann das Modell mithilfe eines eigens maßgeschneiderten Plugins auf der HoloLens 2 zur Vorhersage genutzt werden.

Zur Einordnung wird der Test zunächst mit einem vortrainierten VGG16-Modell (vergleiche Kapitel 2.2.2.1) durchgeführt. Dieses Modell wird verwendet, da eine Konvertierung von YOLOv5 mithilfe von ONNX eine aufwändige Anpassung des Modells notwendig macht. Gleichzeitig kann die Verwendung von VGG16 eine Abschätzung über den unteren Rand der erwarteten Geschwindigkeit bei der Ausführung von YOLOv5 geben, da es ein kleineres und weniger komplexes Modell ist.

Für diesen Test werden außerdem alle anderen Prozesse außer der Vorhersage nicht betrachtet. Das zu inferierende Bild liegt bereits auf dem lokalen Speicher und muss nicht mehr über die Kamera aufgezeichnet werden. Außerdem findet keine Vor- oder Nachverarbeitung der Daten statt. Leider ergab dieser Test, dass die Vorhersage des Modells im Schnitt 42,4 Sekunden dauert. Eine lokale Verarbeitung ist also nicht möglich.

### 5.3.3 Entwicklungsprozess der Nutzer-Anwendung

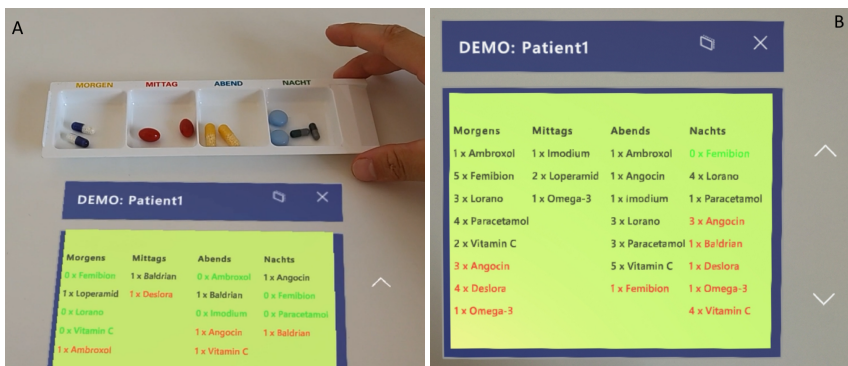
Für die Entwicklung einer Anwendung auf der HoloLens 2 stehen grundsätzlich drei Frameworks zur Auswahl: Unity, Unreal Engine und das Mixed Reality Toolkit (MRTK).



Das MRTK, ein Framework von Microsoft, stellt grundlegende Funktionen für die Entwicklung von Anwendungen auf Mixed Reality-Plattformen bereit. Technisch gesehen nutzen sowohl Unity als auch Unreal Engine das MRTK als Basis, bieten jedoch weitaus mehr Funktionen, ohne dabei die direkte Nutzung von MRTK-Funktionen einzuschränken. Daher besteht kein Vorteil darin, ausschließlich mit dem MRTK zu arbeiten.

Der funktionale Unterschied zwischen Unity und Unreal Engine ist gering, da für die geplante Anwendung keine komplexen Funktionen erforderlich sind. Für die Umsetzung wird die Unreal Engine gewählt, da Unity eine kostenpflichtige Lizenz für Forschungszwecke erfordert und somit hohe Kosten verursachen würde. Die Unreal Engine hingegen bietet eine kostenfreie Nutzung für Entwicklungs- und Forschungszwecke und ist daher die wirtschaftlichere Wahl.

Die Anwendung besteht aus einer grafischen Benutzeroberfläche und einer Anbindung an den Server. Die Anwendung überträgt regelmäßig die Bilder der Frontalkamera über ein lokales Netzwerk an den Server. Die Serverantworten werden anschließend an die Benutzeroberfläche weitergeleitet.



**Abbildung 5.2** Prototyp der Nutzerschnittstelle in der AR-Brille. In der Ansicht des Anwenders (A) und alleinstehend (B).

- fixiert das Fenster an der aktuellen Position
- schließt das Fenster und das Programm
- und wechseln zwischen den Patienten

Ein Prototyp der bildlichen Darstellung der Nutzerschnittstelle ist in Abbildung 5.2 abgebildet. Dabei ist das Fenster mit den Anzeigen zu Beginn frei schwebend und folgt dem Blick des Nutzers im unteren Bereich des Blickfeldes. Die zwei Schaltflächen im oberen Bereich des Bildes erlauben es dem Nutzer, dieses Fenster zu schließen oder es zu fixieren. Einmal fixiert, bleibt das Fenster an einem Ort im realen Raum. Eine Fixierung direkt unter dem Arbeitsbereich wird empfohlen, da so die Frontalkamera die Medikamente gut aufnehmen kann, während der Nutzer die Tabelle und das Feedback der KI überprüft. Mit den pfeilförmigen Schaltflächen am rechten Rand des Fensters können die Patientendaten ausgewählt werden.

## 5.4 Server

Da eine lokale Verarbeitung direkt auf der AR-Brille nicht möglich ist, wird eine Architektur aufgebaut, die die Verarbeitung zwischen der Nutzer-Anwendung auf der AR-Brille und Auswertung auf einem Server aufteilt.

### 5.4.1 Aufbau und Funktion des Servers

Für die Implementierung des Servers wird Python als Programmiersprache gewählt, eine Entscheidung, die sich durch die Spracheigenschaften und die umfangreiche Verfügbarkeit von Bibliotheken für maschinelles Lernen, Datenverarbeitung und Webentwicklung anbietet. Da auch YOLOv5 in Python implementiert ist, erleichtert diese Wahl zusätzlich die Programmierung der Schnittstelle zwischen Kommunikationsframework und KI-Modell.

Der Server nutzt die Flask-Bibliothek, ein Mikro-Netz-Framework für Python, das sich durch seine Einfachheit und Flexibilität auszeichnet. Sie ist ideal für Projekte, die eine maßgeschneiderte Serverlogik erfordern, wie bei der Integration von maschinellem Lernen in Echtzeit.

Die Integration des YOLOv5-Modells in den Server erfolgt durch eine angepasste Schnittstelle, die speziell für die Erkennung und Analyse der Bilder von Medikamenten entwickelt wird. Diese Schnittstelle empfängt Bilder, die von der Nutzer-Anwendung hochgeladen werden, verarbeitet sie und gibt die Erkennungsergebnisse zurück. Die Datenkommunikation zwischen der Nutzer-Anwendung und dem Server wird über das HTTP-Protokoll abgewickelt, wobei der Austausch von Informationen über POST-Anfragen erfolgt.

### 5.4.2 Datenfluss und -verarbeitung

Innerhalb des Flask-Servers werden die eingehenden Bilder in einem vordefinierten Verzeichnis gespeichert. Nach der Speicherung werden sie durch das YOLOv5-Modell analysiert. Die Ergebnisse dieser Analyse werden dann formatiert und nach einer Nachbearbeitung über die gleiche HTTP-Schnittstelle an die AR-Brille zurückgesendet, wodurch ein geschlossener Kreislauf für die Echtzeitdatenverarbeitung geschaffen wird.

Für die Nachbearbeitung wird das Ausgabeformat der YOLOv5-Schnittstelle zur Erkennung wie folgt angepasst: Nach der initialen Objekterkennung und -klassifizierung werden die Ergebnisse in zwei Kategorien eingeteilt: Medikamente und Medikamentendosette. Zunächst wird das Bild auf das Vorhandensein einer Medikamentendosette untersucht. Falls mehrere Medikamentendosetten erkannt werden, wählt das MMS diejenige aus, die am nächsten zum Zentrum des Bildes liegt. Anschließend wird diese Medikamentendosette geometrisch in die vier Zellen unterteilt: Morgens, Mittags, Abends und Nachts.

Die erkannten Medikamente werden dann anhand ihrer Position im Bild den entsprechenden Zellen zugeordnet. Abschließend zählt das MMS die Anzahl der Medikamente jeder Klasse in jeder Zelle. Diese segmentierten und gezählten Ergebnisse werden dann zur weiteren Verwendung an die Nutzer-Anwendung zurückgesendet. Der gesamte Prozess von der Bilderkennung bis zur spezifischen Zuweisung und Zählung der Medikamente wird durch eine einzige Schnittstelle

automatisiert und bei jedem Aufruf von Seiten der Nutzer-Anwendung durchgeführt.

## 5.5 Technische Auswertung

Für die technische Auswertung werden zwei verschiedene Messwerte erhoben. Zum einen wird die Genauigkeit und die Präzision der KI erhoben, um die Fehleranfälligkeit des MMS zu bewerten. Als zweites wird die Verarbeitungszeit betrachtet, um die Echtzeitfähigkeit des MMS zu bewerten.

Die Präzision liegt im vollständig trainierten Netzwerk bei durchschnittlich 0.9834 und die Trefferquote liegt bei durchschnittlich 0.9898. Der Wert für die Präzision-Trefferquote bei 50% Schnittmenge über Vereinigungsmenge liegt bei 0.993. Alle Werte liegen damit nahe an den in Kapitel 5.2.2 bestimmten Werten mit dem Datensatz von nur 5 Medikamenten. Dies zeigt eine gute Skalierbarkeit des Ansatzes.

Tabelle 5.4 zeigt die Zeitmessungen für jeden einzelnen Verarbeitungsschritt, die zur Gesamtlaufzeit des MMS beitragen.

Verarbeitungsschritt	Hardware	Zeit
Bild aufnehmen	HoloLens	1.400,0 ms
Vorverarbeitung	Server	0,1 ms
Datenübertragung	Netzwerk	31,0 ms
KI Vorhersage	Server	7,4 ms
Nachverarbeitung	Server	0,3 ms
Datenübertragung	Netzwerk	32,0 ms
Grafik rendern	HoloLens	35,0 ms

**Tabelle 5.4** Zeitdauer der einzelnen Verarbeitungsschritte, gerundet auf 2 signifikante Stellen

**Bild aufnehmen:** Der Prozess des Bilderfassens mit der HoloLens dauert 1,4 s. Dies umfasst das Aufrufen der asynchronen Schnittstelle, die von Microsoft dafür zur Verfügung gestellt wird und die Kameraaktivität.

**Datenübertragung:** Die Übertragung der Bilddaten vom Gerät zum Server und zurück erfolgt in insgesamt 63 ms. Dies kann zu einem entscheidenden Faktor werden, insbesondere bei der Nutzung eines Netzwerkes mit schlechter Konnektivität.

**KI Vorhersage:** Die Vorhersage des KI-Modells auf dem Server dauert 7,4 ms. Dies zeigt die Effizienz des YOLOv5-Modells in der Erkennung und Klassifizierung der Medikamente.

**Vorverarbeitung:** Die Vorverarbeitung auf dem Server benötigt lediglich 0,1 ms und ist somit für die Gesamtzeit vernachlässigbar.

**Nachverarbeitung:** Die Nachverarbeitung der Ergebnisse benötigt 0,3 ms und ist damit ebenfalls unerheblich für die Gesamtzeit.

**Grafik rendern:** Das Rendern der Ergebnisse auf der HoloLens dauert 35 ms, um die Informationen im Sichtfeld des Benutzers darzustellen.

Insbesondere der Schritt »Bild aufnehmen« benötigt sehr viel Zeit, was aufgrund der vorgegebenen Schnittstelle ohne direkten Kamerazugriff nicht umgangen werden kann. Die übrigen Verarbeitungsschritte dauern insgesamt 82 ms, was eine Bildwiederholungsrate von 12,2 Bildern pro Sekunde ermöglichen würde.

## 5.6 Zwischenfazit

Der Ansatz, die Microsoft HoloLens 2 zur Implementierung einer Augmented-Reality-Funktionalität zu nutzen, ist grundsätzlich vielversprechend, da die HoloLens 2 eine solide Grundlage für anspruchsvolle AR-Anwendungen bietet. Insbesondere die lokale Verarbeitung von KI-Modellen mithilfe von Windows ML

und ONNX zeigt theoretisch großes Potenzial, erweist sich jedoch in der Praxis als herausfordernd.

Die in Kapitel 7.2 beschriebene Nutzerstudie verdeutlicht dennoch eine Verbesserung gegenüber herkömmlichen Ansätzen. Die Anwender profitieren von den AR-Informationen und der intuitiven Benutzeroberfläche, die die Interaktion mit den Medikamenten vereinfacht und visuell unterstützt. Zudem zeigt die Studie, dass die Integration der HoloLens 2 eine weniger fehleranfällige Bearbeitung im Vergleich zu traditionellen Methoden ermöglicht.

Allerdings wurde im Feedback deutlich, dass das Tragen der Brille als unangenehm empfunden wird, insbesondere, da sich das Gerät bei längerer Nutzung stark erhitzt. Darüber hinaus wurde die langsame Verarbeitung als wesentlicher Nachteil benannt. Diese beiden Aspekte schmälern die insgesamt positive Nutzererfahrung und unterstreichen die Notwendigkeit, die lokale Rechenleistung zu verbessern oder alternative Lösungsansätze in Betracht zu ziehen.

Zusammenfassend lässt sich sagen, dass der Einsatz von AR-Technologie vielversprechende Ergebnisse liefert, jedoch durch die begrenzte Rechenleistung des Geräts eingeschränkt wird. Ansätze, die die Vorteile lokaler und serverbasierter Verarbeitung kombinieren, könnten dazu beitragen, die Effizienz und Benutzerfreundlichkeit weiter zu steigern.

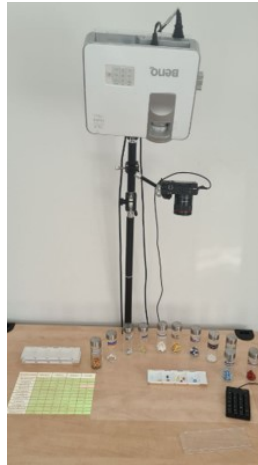
## 6 Technische Entwicklung

In diesem Kapitel wird die Weiterentwicklung des Prototyps zu einem vollständigen System beschrieben. Dabei werden vor allem die Rückmeldungen aus der Zwischenevaluation genutzt, um das Konzept anzupassen. Außerdem wird die Skalierbarkeit des MMS gesteigert, sodass mehr verschiedene Arten von Medikamenten erkannt werden können.

### 6.1 Hardwareumbau

Die Entwicklung des endgültigen Hardwareaufbaus erfolgt nach den in Kapitel 4.1 gezeigten Anforderungen. Insbesondere die Anforderungen A1.2 »Der Arbeitsplatz muss auf ca.  $1\text{m}^2$  beschränkt sein«, A1.3 »Das MMS soll bei Nichtverwendung die alltägliche Arbeit nicht einschränken« und A3.2 »Es muss für Endnutzer zertifizierte Elektronik verwendet werden« sind für diesen Aufbau relevant. Um Anforderungen A1.2 und A1.3 zu erfüllen, wird ein Metallgestänge gewählt, um die benötigten elektrischen Geräte zu fixieren. Konkret wird das TARION Overhead Tischstativ verwendet, das an der Rückseite des Arbeitsbereichs am Tisch befestigt werden kann. Das Stativ hat eine Gesamtlänge von 104 cm mit zwei Gelenken.

An dem Tischstativ werden eine Kamera, ein Beamer und ein Einplatinenrechner befestigt. Abbildung 6.1 zeigt diesen Aufbau inklusive der genauen Positionierung der Komponenten. Der Beamer projiziert die Benutzeroberfläche auf den Arbeitsplatz, während die Kamera Bilder des Arbeitsbereichs aufnimmt. Diese Bilder werden an den Einplatinenrechner gesendet, der als zentrale Recheneinheit



**Abbildung 6.1** Hardwareaufbau mit Tischstativ, Kamera, Beamer und Einplatinenrechner

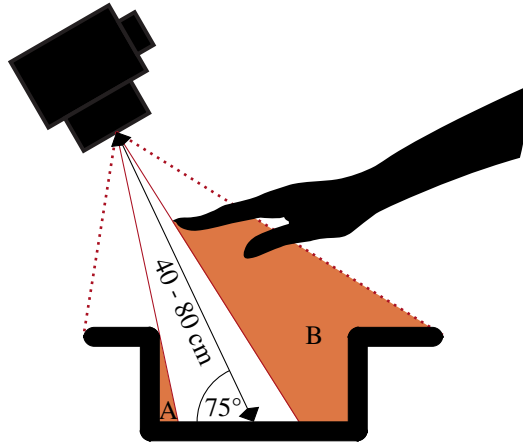
zur Ansteuerung von Beamer und Kamera fungiert und die Nutzerschnittstelle zusammenstellt. Nach der Verarbeitung der Daten kommuniziert der Einplatinenrechner mit dem Server, damit dieser die Erkennung und Klassifizierung der Medikamente durchführt.

### 6.1.1 Kameraplatzierung

Beim Aufbau der Hardware ist vor allem der Winkel der Kamera besonders relevant. Da das MMS während der aktiven Arbeit des Nutzers eine vollständige Aufnahme der Medikamentendosette benötigt, verdecken die Hände des Nutzers oft den Arbeitsbereich. Abbildung 6.2 zeigt diese Abdeckung in Bereich B, sowie die Abdeckung durch die Seitenwände der Medikamentendosette bei einem flachen Kamerawinkel in Bereich A. Dies würde für eine Kameraposition direkt über dem Arbeitsbereich sprechen. Allerdings hat dies den gravierenden Nachteil, dass die Tiefeninformation der Tabletten verloren geht. Das Bild hat in diesem Fall keine Informationen mehr über die Dimensionen oder die Form der übrigen Seiten. Aus diesem Grund ist die Anordnung in einem Winkel von ungefähr  $75^\circ$ ,



aber nicht direkt über der Medikamentendosette, zu bevorzugen. Die Höhe ist ein Designparameter und sollte so gewählt werden, dass der gesamte Arbeitsbereich immer aufgenommen werden kann.



**Abbildung 6.2** Kameraaufbau über der Medikamentendosette und potentielle Abschattungen  
 A Abschattung durch die Wände der Medikamentendosette  
 B Abschattung durch die Hände des Nutzers

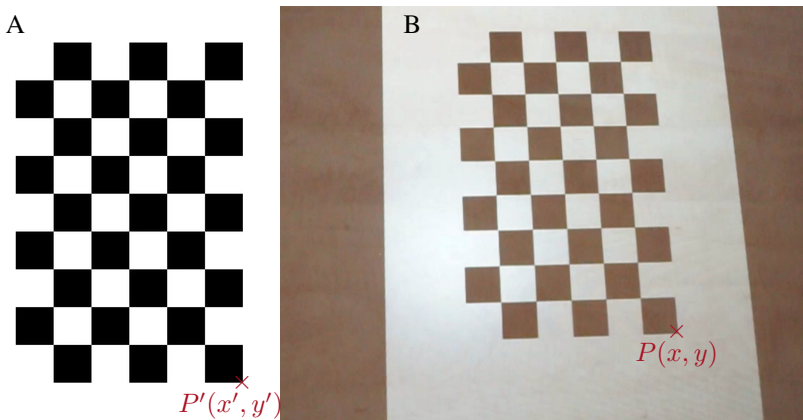
## 6.2 Koordinatentransformation

Eine Koordinatentransformation ist ein Verfahren, bei dem Koordinaten eines Punktes in einem Koordinatensystem in ein anderes Koordinatensystem umgewandelt werden. Diese Transformation ist notwendig, da der Winkel zwischen Kamera, Tisch und Beamer an jedem Einsatzort verschieden ist.

Außerdem kann auch die Größe des Tisches nicht als bekannt angenommen werden. Aus diesem Grund wird in einer zweiten Kalibrierungsphase die Größe des Tisches mithilfe von projizierten QR-Codes bestimmt. Zur Umsetzung werden Funktionen aus der OpenCV Bibliothek [35] verwendet.

## 6.2.1 Homographie Berechnen

In diesem Schritt wird ein Schachbrettmuster auf die Arbeitsfläche projiziert, das von der Kamera anschließend erfasst wird. Dieses Schachbrettmuster besteht aus regelmäßigen, kontrastreichen Kanten und Ecken, die leicht zu erkennen und zu verfolgen sind. Anhand der erfassten Bilder des Schachbrettmusters kann das MMS die Lage und Orientierung der Arbeitsfläche relativ zur Kamera bestimmen. Die Berechnung der Homographie läuft in 4 Schritten:



**Abbildung 6.3** Original (A) und Projektion (B) des Schachbrettmusters der Homographieberechnung.  $P(x, y)$  ist ein Korrespondenzpunkt zu  $P(x', y')$

1. Zunächst werden die Korrespondenzpunkte bestimmt. Dabei handelt es sich um bekannte Punkte, die für das Muster und die aufgezeichnete Projektion übereinstimmen. Hierfür werden die Ecken zwischen den einzelnen Feldern des Schachbrettmusters, das in Abbildung 6.3 gezeigt wird, verwendet. Die Punkte werden dafür über eine Canny-Kantenerkennung detektiert. Obwohl ein Schachbrett dieser Größe 70 Eckpunkte hat, werden nur 68 Punkte verwendet, da die weißen Ecken oben und unten links nicht verwendet werden können. Die große Anzahl der Punkte ermöglicht es, nicht-lineare Unebenheiten auf der Tischoberfläche auszugleichen.

2. Für das Homographie-Gleichungssystem werden die Koordinaten der Punkte  $P'_i(x', y')$  im Muster den korrespondierenden Punkten  $P_i(x, y)$  auf der Aufnahme gegenübergestellt und mit der Homographiematrix  $H$  (Dimension  $3 \times 3$ ) verknüpft. Dabei wird ein Skalierungsfaktor  $s$  hinzugenommen, um eine Normierung der Matrix zu ermöglichen.

$$s \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = H \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (6.1)$$

3. Eine eindeutige Lösung des Homographie-Gleichungssystems ist nicht möglich, da es sich um ein überbestimmtes Gleichungssystem handelt. Um die best mögliche Lösung zu finden, wird die Methode des minimalen quadratischen Fehlers eingesetzt. Dazu wird Gleichung 6.1 ohne Skalierungsfaktor in Matrixform gebracht:

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x'_1 x_1 & -x'_1 y_1 & -x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y'_1 x_1 & -y'_1 y_1 & -y'_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x'_2 x_2 & -x'_2 y_2 & -x'_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -y'_2 x_2 & -y'_2 y_2 & -y'_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & 1 & 0 & 0 & 0 & -x'_n x_n & -x'_n y_n & -x'_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -y'_n x_n & -y'_n y_n & -y'_n \end{pmatrix} \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad (6.2)$$

Die Matrix in Gleichung 6.2 wird im Weiteren als  $A$  bezeichnet und der Vektor als  $h$ . Um die Lösung zu finden wird eine Singulärwertzerlegung durchgeführt:

$$A = UDV^T \quad (6.3)$$

In Gleichung 6.3 ist  $U$  eine orthogonale  $2n \times 2n$ -Matrix,  $D$  eine Diagonalmatrix der Dimension  $2n \times 9$  mit den Singulärwerten auf der Diagonalen und  $V$  eine orthogonale  $9 \times 9$ -Matrix.

Der Lösungsvektor  $\mathbf{h}$  entspricht dem Spaltenvektor von  $A$ , der zum kleinsten Singulärwert von  $D$  gehört.

4. Der Vektor  $\mathbf{h}$  wird zum Schluss wieder in eine  $3 \times 3$ -Matrix umgeformt, um die Homographiematrix  $H$  zu erhalten:

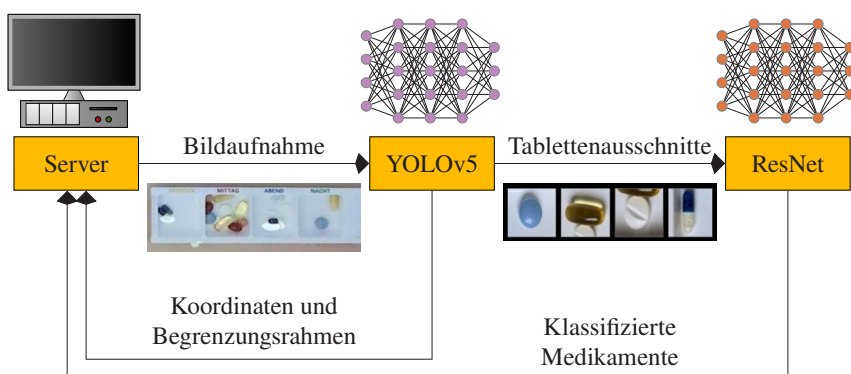
$$H = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \quad (6.4)$$

## 6.2.2 Tischgrößenbestimmung

Nach der Homographieberechnung folgt die Bestimmung des Arbeitsbereichs. Der Arbeitsbereich liegt direkt neben der Nutzeroberfläche und enthält die Medikamentendosette während des Richtprozesses. Zur Bestimmung wird ein Gitter aus QR-Codes auf den Tisch projiziert. In diesem ist jeweils die Koordinate codiert. Anhand der erkannten QR-Codes auf dem aufgenommenen Kamerabild, kann dann bestimmt werden, welcher Bereich des Kamerabildes für die Medikamentenerkennung genutzt werden kann. Dies ist möglich, da die QR Codes nur auf freier, ebener Tischfläche ohne Hindernisse erkannt werden können. Der Arbeitsbereich wird in zwei Koordinatensystemen gespeichert: dem Nutzeroberflächen-Koordinatensystem (NOK) und dem Kamera-Koordinatensystem (KAK). Das NOK enthält die Streckung und Verdrehung von Tisch zu Nutzeroberfläche und wird für die Berechnung der Begrenzungsrahmen auf den Medikamenten benötigt. Das KAK beschreibt das Verhältnis von Tisch zu Kamerabild. Dieses wird verwendet, um das Kamerabild auf den Arbeitsbereich mit der Medikamentendosette zuzuschneiden, bevor es an den Server und die KI weitergeleitet wird. Dieses Koordinatensystem wird auch in der weiteren Verarbeitung von der KI verwendet.

## 6.3 KI Weiterentwicklung

Um die Skalierbarkeit bei der Erkennung einer großen Anzahl von Medikamenten zu verbessern, wird eine Kaskade aus zwei verschiedenen KI-Modellen eingesetzt (siehe Abbildung 6.4). Die Kombination von YOLOv5 und anschließend ResNet erhöht die Effizienz des MMS, da beide Modelle spezifische Stärken besitzen, die in Kombination eine effektivere Verarbeitung ermöglichen.



**Abbildung 6.4** KI-Kaskade von YOLOv5 und ResNet zur Effizienzerhöhung bei großer Artenzahl an Medikamenten

Um als Echtzeit-Objektdetektor die erforderliche Verarbeitungsgeschwindigkeit zu erreichen, verarbeitet YOLOv5 das Bild in einer reduzierten Auflösung von  $640 \times 640$  Pixeln. Diese Auflösung ist ausreichend, um sowohl die Medikamentendosette als auch die einzelnen Medikamente zu lokalisieren.

Nachdem YOLOv5 die Positionen der Medikamente im Bild identifiziert hat, werden die entsprechenden Bildausschnitte aus dem Originalbild extrahiert. Da diese Ausschnitte deutlich kleiner als das Gesamtbild sind und nur die relevanten Bereiche enthalten, ist es möglich, sie in höherer Auflösung zu verarbeiten, ohne die Echtzeitfähigkeit des Gesamtsystems zu beeinträchtigen. ResNet kann

dadurch detailliertere Merkmale analysieren und eine präzisere Klassifikation der Medikamente durchführen.

Die Verwendung der Kaskade bietet auch bei der Skalierung auf eine große Anzahl von Medikamentenklassen erhebliche Vorteile. Da YOLOv5 die allgemeine Objektdetektion übernimmt, muss es besonders bei sehr ähnlichen Medikamenten nicht für jede spezifische Klasse umfassend trainiert werden. Dies reduziert den Bedarf an Trainingsdaten für die Detektion. ResNet hingegen spezialisiert sich auf die Feinklassifikation der bereits detektierten Medikamente. Durch diese Aufgabenteilung kann das MMS effizient erweitert werden, indem neue Klassen hauptsächlich in ResNet integriert werden.

### 6.3.1 Training von ResNet

ResNet wird wie schon YOLOv5 mit einem großen, öffentlichen Datensatz vor-trainiert - in diesem Fall ImageNet 1k V2 [36] - und durch Transfer Learning

Medikament	Abkürzung	Datensatz 1	Datensatz 2	Datensatz 3	Insgesamt
Ambroxol	Amb	2	16	71	575
Angocin	Ang	5	23	77	603
Baldrian	Bal	8	24	88	682
Deslora	Des	3	22	77	599
Femibion	Fem	2	8	55	441
Imodium	Imo	4	16	65	513
Loperamid	Lop	5	18	69	535
Lorano	Lor	5	33	93	720
Omega 3	Ome	2	17	80	638
Paracetamol	Par	4	22	94	733
Vitamin C	Vit	4	18	98	769

**Tabelle 6.1** Zusammensetzung der Trainingsdatensätze zur Bestimmung der optimalen Datenmenge

auf einem eigenen Datensatz feinabgestimmt. Für diesen werden die Trainingsbilder für YOLOv5 in einzelne Ausschnitte mit jeweils einem Medikament pro Bild aufgeteilt. Im Gegensatz zu YOLOv5 werden für ResNet die Trainingsbilder in voller 1080p-Auflösung verarbeitet. Mit Transfer Learning können die bereits gelernten Merkmalsrepräsentationen von ResNet genutzt und dadurch der Trainingsaufwand erheblich reduziert werden. Nur der letzte Block des Modells muss für die spezifischen Medikamentenklassen feinabgestimmt werden.

Zum Testen der optimalen Trainingsdatenmenge wird das ResNet-Modell mit drei verschiedenen Trainingsdatensätzen der in Kapitel 4.4.2 vorgestellten Medikamente trainiert. Die einzelnen Bilder wurden dabei zufällig aus allen verfügbaren Trainingsdaten ausgewählt. Tabelle 6.1 zeigt den genauen Aufbau der verschiedenen Datensätze, dabei wurde die Anzahl der Bilder zwischen jedem Datensatz jeweils vervierfacht.

In Datensatz 1 ist eine generelle Tendenz zur korrekten Klassifikation der Medikamente zu erkennen (siehe Abbildung 6.5). Einzelne Medikamente wie Des, einer einfachen weißen Tablette, werden jedoch kaum richtig zugeordnet. Aber auch andere Medikamente im unteren Drittel der Konfusionsmatrix zeigen eine hohe Anzahl an Falschzuweisungen.

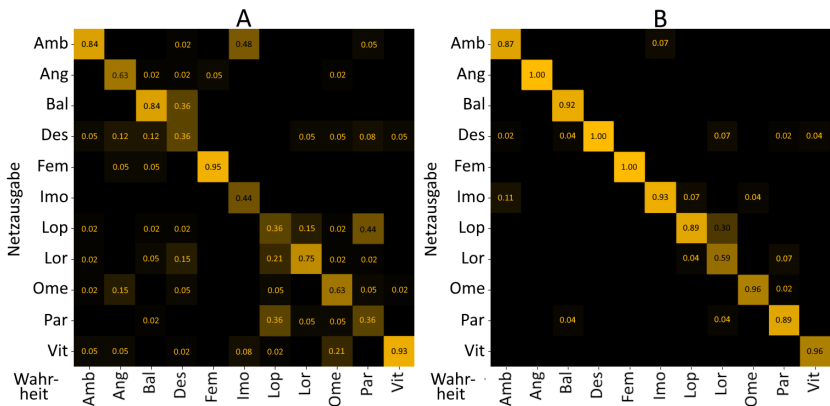


Abbildung 6.5 Konfusionsmatrizen: Datensatz 1 (A) und Datensatz 2 (B)

Im Gegensatz dazu sind in Datensatz 2 bereits fast alle Medikamente korrekt zugeordnet (siehe Abbildung 6.6). Lediglich die Unterscheidung zwischen Loperamid akut und Lorano akut, ebenfalls zwei einfache weiße Tabletten, ist unzureichend.

Erst in Datensatz 3 sind alle Medikamente bis auf einzelne Ausnahmen korrekt zugeordnet. Dies bedeutet, dass für das initiale Training jede Art von Medikament in einer Mindestanzahl von 70-100 Bildern vorhanden sein sollte, um eine korrekte Erkennung zu erreichen.

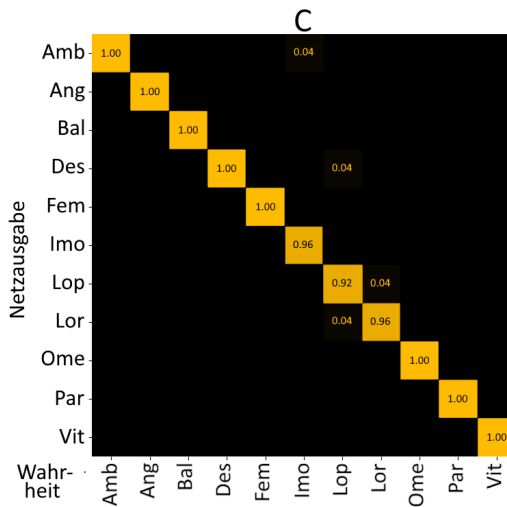


Abbildung 6.6 Konfusionsmatrizen von Datensatz 3 (C)

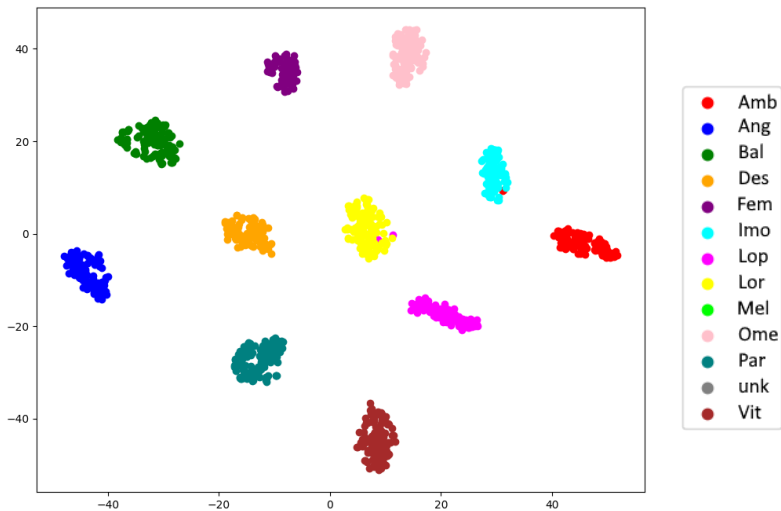
### 6.3.2 t-SNE Auswertung

Die t-distributed Stochastic Neighbor Embedding (t-SNE) ist eine Methode zur Dimensionsreduktion, die besonders für die Visualisierung hochdimensionaler Daten geeignet ist. Sie projiziert Datenpunkte aus einem hochdimensionalen Merkmalsraum in einen zweidimensionalen Raum. Dies geschieht, indem die Abstände im hochdimensionalen Raum in eine t-Verteilung umgerechnet wird.



Anschließend wird eine Lösung im zweidimensionalen Raum berechnet, bei der die Divergenz dieser Wahrscheinlichkeiten minimiert wird.

Im trainierten ResNet-Modell wird t-SNE verwendet, um die von den letzten Schichten des Netzwerks extrahierten Merkmalsrepräsentationen der Eingabedaten zu visualisieren. Indem die hochdimensionalen Merkmalsvektoren der Validierungs- oder Testdaten mittels t-SNE auf zwei Dimensionen reduziert werden, lässt sich erkennen, wie gut die verschiedenen Medikamentenklassen im Merkmalsraum voneinander getrennt sind (siehe Abbildung 6.7).

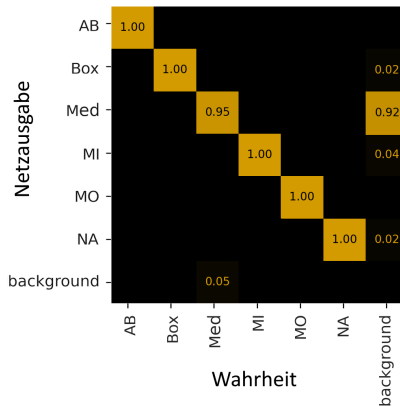


**Abbildung 6.7** t-SNE Visualisierung der Merkmalsrepräsentationen der Testdaten im trainierten ResNet zur Darstellung der Klassentrennung

Eine klare Trennung der Klassen deutet darauf hin, dass das Modell effektive Merkmale gelernt hat, um die Klassen zu unterscheiden, was auf eine hohe Trainingsqualität hinweist. Nur wenige Datenpunkte verschiedener Klassen überlappen, wie z.B. die Loperamid-Datenpunkte (Lop in pink) in der Lorano- (Lor in gelb) und der Imodium-Gruppierung (Imo in türkis). Dies zeigt, dass bis auf wenige Ausnahmen eine eindeutige Trennung der Klassen vorgenommen werden kann.

### 6.3.3 Training von YOLOv5

Für das Training von YOLOv5 werden die Gesamtbilder des vollen Trainingsdatensatzes verwendet. Dieser Datensatz besteht aus 1.614 Bildern. Insgesamt werden so 8.532 einzelne Objekte im Training des Modells verwendet. Im vorherigen Modell sind insbesondere Schwächen bei verschiedener Beleuchtung aufgefallen. So ist die korrekte Erkennung bei direkter Sonneneinstrahlung, oder einer anderen hellen Lichtquelle direkt über dem Arbeitsbereich, besonders schlecht und deswegen im neuen Datensatz verstärkt vorhanden. Die Aufteilung in 70% Trainingsdaten, 15% Validierungsdaten und 10% Testdaten bleibt unverändert.



**Abbildung 6.8** Konfusionsmatrix von YOLOv5 trainiert mit allen Medikamenten als eine Klasse.

AB - Abends; MI - Mittags; MO - Morgens; NA - Nachts; Box - Medikamentendosette; Med - Medikamente

Abbildung 6.8 zeigt die Konfusionsmatrix nach dem Training mit diesem neuen Datensatz. Da für YOLOv5 alle Medikamente zu einer Klasse zusammengefasst werden, sind diese in der Matrix unter »Med« zusammengefasst. An der Diagonale ist klar zu sehen, dass eine Verwechslung unter den vorhandenen Klassen gar nicht mehr stattfindet. Allerdings sind durch die erschwerten Belichtungsbedingungen einige Medikamente nicht erkannt worden. Im realen Einsatz kann dieses Problem leicht behoben werden, indem die entsprechenden Medikamente bewegt werden.

## 6.4 Handerkennung

Die Handerkennung wird eingesetzt, um die Interaktion des Benutzers mit dem MMS zu optimieren. Da die Hand des Benutzers einzelne oder mehrere Medikamente verdecken kann, was die genaue Erkennung erschwert, ist eine zuverlässige Handerkennung notwendig. Um dies zu verhindern, wird der Erkennungsprozess unterbrochen, wenn eine Hand erkannt wird, die mindestens 15% der Arbeitsfläche verdeckt. Ein solcher Schwellwert verhindert, dass einzelne Tabletten in Hautfarbe die Unterbrechung auslösen. Die Abdeckung wird über die Gesamtanzahl der Pixel berechnet, die innerhalb des Arbeitsbereichs liegen.

Für die Handerkennung bietet der HSV-Farbraum (Hue, Saturation, Value) eine geeignete Darstellung zur Charakterisierung von Farbvariationen im Bild. Im Vergleich zum RGB-Farbraum trennt der HSV-Farbraum Farbton, Sättigung und Helligkeit, was die Segmentierung von Hauttönen erleichtert. Bei der Handerkennung können die drei Hauptkomponenten des HSV-Farbraums wie folgt genutzt werden [37]:

**Hue** (Farbton): Der Farbton wird auf einer Skala von  $0^\circ$  bis  $360^\circ$  gemessen. Er ermöglicht die Identifizierung charakteristischer Hautfarben. Durch die Analyse des Farbtons können spezifische Farbmerkmale der menschlichen Haut extrahiert werden, was die Unterscheidung von anderen Objekten im Bild erleichtert. Der Wert muss für eine erkannte Hand im roten Bereich zwischen  $0^\circ$  und  $50^\circ$  liegen.

**Saturation** (Sättigung): Die Sättigung wird auf einer Skala von 0% bis 100% gemessen und gibt Aufschluss über die Echtheit der Hautfarbe. Die Filterung stark gesättigter Bereiche unterstützt die Reduzierung von Hintergrundstörungen, da menschliche Haut oft eine geringe Sättigung aufweist. Der Wert muss für eine erkannte Hand zwischen 20% und 60% liegen.

**Value** (Wert): Der Wert wird auf einer Skala von 0% bis 100% gemessen und repräsentiert die Helligkeit der Farbe. Haut kann von sehr dunklen bis zu sehr hellen Tönen variieren. Der Wert muss für eine erkannte Hand zwischen 5% und 90% liegen.

## 6.5 Nutzerschnittstelle

Die initiale Version der Nutzerschnittstelle in Kapitel 5.3.3 konzentrierte sich auf die Anzeige grundlegender Informationen wie den Medikamentennamen und die vorgesehenen Einnahmezeitpunkte. Diese reduzierte Darstellung reichte für erste Funktionstests und Konzeptnachweise aus, erfüllt jedoch nicht die umfassenden Anforderungen für eine sichere Medikamentenverwaltung. Der Medikationsplan nach dem E-Health-Gesetz aus Kapitel 4.4.1 bietet hingegen eine strukturierte Darstellung aller relevanten Informationen zu den verordneten Medikamenten eines Patienten. Die Integration dieses Plans in die Nutzerschnittstelle ermöglicht es, folgende Informationen übersichtlich darzustellen:

- **Medikamentenname:** Anzeige des Handelsnamens und des Wirkstoffs.
- **Dosierung und Einheit:** Genaue Angabe der einzunehmenden Menge und der entsprechenden Einheit (z.,B. mg, ml).
- **Darreichungsform:** Information über die Form des Medikaments (z.,B. Tablette, Kapsel, Dragee).
- **Einnahmezeitpunkte:** Klare Darstellung der vorgesehenen Einnahmezeitpunkte (Morgens, Mittags, Abends, Nachts).
- **Hinweise zur Einnahme:** Spezielle Anweisungen wie »vor dem Essen«.
- **Grund der Medikation:** Angabe des therapeutischen Ziels oder der Diagnose.
- **Patienteninformationen:** Name, Geburtsdatum und andere relevante Daten.

Die Darstellung dieser Informationen erfolgt in einer klar strukturierten Tabelle, die basierend auf den Grundlagen aus Kapitel 2.3 eine schnelle Erfassung und Verarbeitung ermöglicht. Dabei muss beachtet werden, dass die Integration zusätzlicher Daten die Komplexität der Anzeige erhöhen kann. Daher gilt es, eine Balance zwischen Informationsfülle und Übersichtlichkeit zu finden.

**Medikationsplan**

Name: Max Mustermann | Geburtsdatum: 10. Januar 1985 | Wochentag: Montag

Wirkstoff	Handelsname	Stärke	Form	Morgen	Mittag	Abend	Nacht	Einheit	Hinweise
Ambroxol	Ambroxol	Verschiedene	Tablette	1 / 1	0 / 0	0 / 0	0 / 0		
Desloratadin	Deslora	Verschiedene	Tablette	0 / 0	0 / 0	1 / 1	0 / 0	mg	Schleimbildung, Hustenlinderung
Omega-3 Fettsäuren	Omega-3	Verschiedene	Tablette	1 / 1	0 / 0	0 / 0	0 / 0	mg	Arztstärkung, Allergien
Ascorbinsäure	Vitamin C	Verschiedene	Tablette	0 / 0	0 / 0	1 / 1	0 / 0	mg, g	Herzgesundheit, Gehirnfunktion

Geschlecht	männlich
Allergien	Keine bekannt
Vorerkrankungen	Ashma
Hinweise zum Patienten	Alle Medikamente nach den Mahlzeiten einnehmen.
Anmerkungen	

Zu entfernende Tabletten

Femiblon (Nacht)

Abbildung 6.9 Nutzerschnittstelle mit allen integrierten Informationen

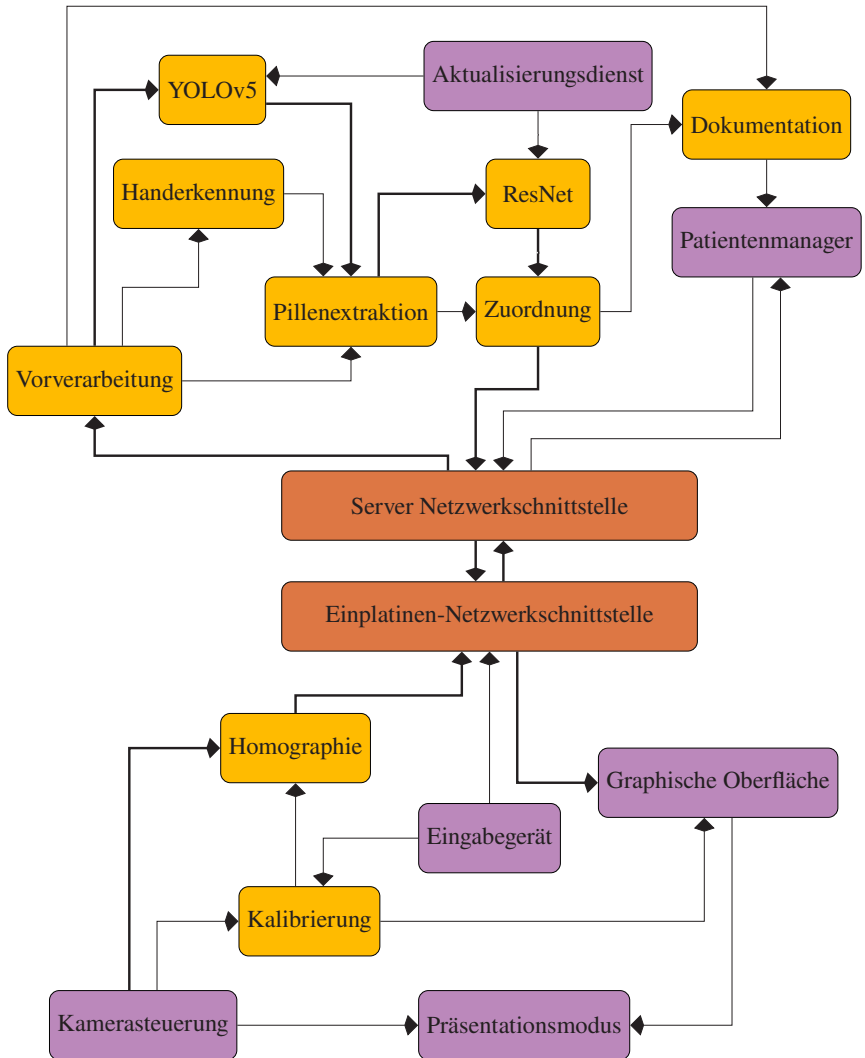
Um diese Anforderungen umzusetzen, wird die Nutzerschnittstelle wie folgt umgesetzt (siehe Abbildung 6.9):

- Unter Verwendung von gut lesbaren Tabellen, wird eine schnelle Erfassung der Informationen ermöglicht. Diese Strukturierung erleichtert die Orientierung.
- Die Tabelle wird farblich kodiert, Grün für korrekt einsortierte Medikamente und Rot für fehlende oder fehlerhaft einsortierte Medikamente. Dies ermöglicht eine sofortige Erkennung des Medikationsstatus.
- Eine separate Tabelle für fehleinsortierte Medikamente wird eingeführt, die es dem Pflegepersonal ermöglicht, nicht benötigte Medikamente schnell zu identifizieren und zu korrigieren ohne die Hauptanzeige mit Informationen zu überlasten.

## 6.6 Gesamtübersicht

Das in Abbildung 6.10 dargestellte Diagramm illustriert die verschiedenen Komponenten des vollständigen MMS mit dessen Interaktionen. Die Komponenten sind farblich kodiert: lila für Schnittstellen nach Außen, gelb für Verarbeitungseinheiten und orange für Netzwerkschnittstellen.

1. Der Datenfluss beginnt mit dem Eingabegerät. Dieses sendet die Daten sowohl an die Kalibrierung, um beim Start des MMS unterbrechen zu können (z.B. für Fehlerbehebung), als auch an die Netzwerkschnittstelle, um im Server das Laden eines neuen Patienten starten zu können.
2. Die Kamerasteuerung liefert aktuelle Bilder beim Start des MMS an die Kalibrierung, und im laufenden Betrieb an die Homographie und den Präsentationsmodus.
3. Der Präsentationsmodus ermöglicht es, die aktuelle Arbeit über die Kameraaufnahme und den aktuellen Stand des Richtprozesses über die Graphische Oberfläche an einen externen Bildschirm zu spiegeln. Dies ermöglicht ein Zuschauen bei der Arbeit.
4. Die Kalibrierung sorgt für die in Kapitel 6.2 gezeigte automatische Erkennung des Arbeitsbereichs und für die Berechnung der Homographiematrix, die dann an die Homographie weitergeleitet wird.
5. Die Homographie beginnt nach der Kalibrierung und berechnet die Koordinatentransformation kontinuierlich aus den Bildern der Kamerasteuerung. Diese werden anschließend über die Netzwerkschnittstelle an den Server weitergeleitet.
6. Die Graphische Oberfläche empfängt beim Systemstart den Arbeitsbereich und die Homographiematrix von der Kalibrierung, um im laufenden Betrieb die Nutzerschnittstelle und die farbigen Begrenzungsrahmen unverzerrt anzuzeigen.



**Abbildung 6.10** Komponenten-Diagramm mit farblicher Kodierung: lila für Schnittstellen, gelb für Verarbeitungseinheiten und orange für Netzwerkkommunikation. Die dicken Pfeile geben den normalen Prozessablauf an.

7. Die Einplatinen-Netzwerkschnittstelle kommuniziert bidirektional über das lokale Netzwerk mit der Server-Netzwerkschnittstelle.
8. Die Vorverarbeitung bereitet die eingehenden Bilder für die weitere Analyse vor und sendet sie an YOLOv5, die Handerkennung, die Pillenextraktion und bei Bedarf an die Dokumentation.
9. YOLOv5 führt die Objekterkennung durch und leitet die erkannten Begrenzungsrahmen an die Pillenextraktion.
10. Die Handerkennung gibt im Fall einer Überschattung des Arbeitsbereichs ein Stoppsignal an die Pillenextraktion.
11. Die Pillenextraktion sendet sowohl die mit den erkannten Begrenzungsrahmen ausgeschnittenen Bilder in Originalauflösung an ResNet, als auch die Bildkoordinaten aller erkannten Objekte an die Zuordnung.
12. Bei ResNet handelt es sich um das in Kapitel 2.2.2 vorgestellte Modell, welches die extrahierten Pillen klassifiziert und die Informationen an die Zuordnung übergibt.
13. Die Zuordnung ordnet die klassifizierten Daten den entsprechenden Einnahmezeitpunkten zu und kommuniziert die Ergebnisse an die Dokumentation und über die Netzwerkschnittstellen an die grafische Oberfläche.
14. Die Dokumentation speichert ein Bild für die Dokumentation im Patientenmanager ab, wenn alle Pillen korrekt einsortiert sind.
15. Der Patientenmanager interagiert mit dem Pflegemanagementsystem, um die Dokumentation abzulegen und um bei Bedarf neue Medikationspläne abzufragen.
16. Der Aktualisierungsdienst hält die Modellgewichte von YOLOv5 und ResNet auf dem neuesten Stand, damit die Modelle stets mit den aktuellsten Daten arbeiten.



## 7 Nutzerevaluation

Im Laufe der Arbeit wurden zwei Evaluationen mit potenziellen Nutzern durchgeführt. Die erste Zwischenevaluation wurde mit der AR-Brille durchgeführt. Bei der Endevaluation, nach Abschluss der Entwicklung an dem AR-System mit Kamera und Beamer, ist die Probandenzahl deutlich höher als bei der Zwischenevaluation. In beiden Fällen wurden sowohl professionelle Pflegekräfte als auch Laien eingeladen.

### 7.1 Methodik der Evaluation

In diesem Kapitel wird zunächst der Ablauf der Evaluation erklärt, inklusive der erhobenen Messwerte. Anschließend wird auf die Fragebögen eingegangen, welche von den Probanden ausgefüllt werden sollen.

#### 7.1.1 Versuchsablauf

Die Evaluation ist für jeden Probanden in zwei Phasen unterteilt. Jeder Proband erhält Papier-Medikamentenpläne (PMP), einige Pillendosen und einen Vorrat an Tabletten. Die Medikamentenpläne enthalten zufällig zusammengestellte Tabletten für imaginäre Patienten. Dabei hat jeder Patient für jeden Wochentag einen komplett anderen Medikamentenplan. Der Tablettenvorrat enthält mehr Tabletten als für das Richten der einzelnen Medikamentenpläne benötigt wird. Um jedoch über die volle Evaluationszeit stets genug Vorrat zu haben, werden die fertig

angerichteten Pillendosen nach der Dokumentation sofort zurück in den Vorrat sortiert.

Bei der Durchführung haben die Probanden zunächst ca. 5 Minuten Zeit, um sich in die Aufgabe des Richtens gemäß des gegebenen Medikamentenplans einzuarbeiten. Anschließend beginnt die erste Evaluationsphase. Die Probanden werden gebeten, so viele Tabletten wie möglich innerhalb von 15 Minuten zu richten, während gleichzeitig eine geistige Ablenkung simuliert wird. Dazu wird eine verbale Version des OSPAN-Tests durchgeführt. Der Operation Span Test (OSPAN) fordert den ständigen Einsatz von Gedächtnis und mathematischem Denken [38]. Er wird verwendet, um eine kognitive Überlastung bei den Teilnehmern zu erzeugen, indem die Probanden zwischen dem Lösen von Gleichungen und Gedächtnisübungen abwechseln. Während jeder Sequenz erhalten die Probanden alternierend drei Buchstaben und zwei Gleichungen. Die drei Zeichen sollen sie sich merken und nach dem Ende der Sequenz wiederholen. Die beiden mathematischen Gleichungen sollen innerhalb von 3 Sekunden als »richtig« oder »falsch« beurteilt werden. Einige Beispielsequenzen sind in Abbildung 7.1 zu sehen. Jede Sequenz wird mit einem Takt von 5 Sekunden je Spiegelstrich vorgelesen. Nach einer vorgelesenen Gleichung soll der Proband innerhalb dieser Zeit sofort antworten. Bei der ersten Sequenz sind die richtigen Antworten »falsch« gefolgt von »richtig« für die Gleichungen und anschließend »C« gefolgt von »X« und »A« um die gemerkten Buchstaben korrekt wiederzugeben.

C - $3 + 3 = 5$ - X - $7 - 2 = 5$ - A	H - $1 + 2 = 3$ - H - $6 - 2 = 3$ - D
B - $4 * 2 = 9$ - Q - $5 + 3 = 7$ - Z	U - $8 / 2 = 6$ - K - $4 - 1 = 3$ - J
G - $9 / 3 = 3$ - R - $4 + 2 = 6$ - P	W - $2 * 3 = 6$ - V - $4 - 3 = 2$ - A

**Abbildung 7.1** 6 Beispiele für Sequenzen aus einem OSPAN-Test

Bei einem Fehler wurde die ganze Sequenz als falsch markiert, zwischen nur einem oder mehreren Fehlern in einer Sequenz wird nicht unterschieden. In jedem

15-minütigen Durchlauf werden insgesamt 30 solcher Sequenzen vorgelesen, also eine Sequenz pro 30 Sekunden. Parallel dazu sind die Tabletten anzurichten.

Für die Evaluation sind immer zwei Personen zum Durchführen zusätzlich zum Probanden erforderlich. Eine Person liest die OSPAN-Tests vor und protokolliert die dabei auftretenden Fehler. Die andere Person dokumentiert die fertig angeordneten Pillendosen und sortiert anschließend die Medikamente zurück in den Vorrat.

Zur Berechnung der Fehlermetrik werden in der Evaluation die Gesamtanzahl an einsortierten Medikamenten, die Anzahl der Sortierfehler und die Anzahl der Abweichungen im OSPAN-Test aufgezeichnet. Jede fehlende, überschüssige oder verwechselte Pille wird als ein Fehler gezählt. Ebenso werden übersprungene oder wiederholte Medikamentenpläne als Fehler gezählt und für eine separate Auswertung extra markiert (siehe Kapitel 8.3.1). Vor der nächsten Aufgabe wurde den Teilnehmern eine Pause von 10 Minuten gewährt. Die zweite Phase der Evaluation begann mit einer ausführlichen Erklärung des AR-Systems und dessen Anwendung. Anschließend wurde das gleiche Schema wie in Phase eins durchgeführt, wobei das MMS anstelle von PMP verwendet wurde. Die Reihenfolge der Phasen wird für verschiedene Probanden variiert.

## **7.1.2 Fragebögen Auswahl**

Nach Versuchsende wurden die Probanden gebeten, folgende drei Fragebögen auszufüllen:

- System-Usability-Scale-Fragebogen
- NASA-Task-Load-Index-Fragen
- Freitext für offenes Feedback.

### **7.1.2.1 System Usability Scale**

Der angepasste System-Usability-Scale-Fragebogen (SUS) beinhaltet die in Tabelle 7.1 dargestellten Fragen und Optionen [39].

Zur Auswertung werden die einzelnen Antworten in Punkte konvertiert, wobei negative Fragen invertiert werden:

- »Stimme absolut nicht zu«      0 Punkte
- »Stimme nicht zu«                1 Punkt
- »Stimme zu«                        3 Punkte
- »Stimme voll und ganz zu«      4 Punkte

Fragen	Stimme absolut nicht zu	Stimme nicht zu	Stimme zu	Stimme voll und ganz zu
Ich kann mir gut vorstellen, das System regelmäßig zu nutzen.	(0)	(1)	(3)	(4)
Ich empfinde das System als unnötig komplex.	(4)	(3)	(1)	(0)
Ich habe mich bei der Nutzung des Systems sehr sicher gefühlt.	(0)	(1)	(3)	(4)
Ich denke, dass ich technischen Support brauchen würde, um das System zu nutzen.	(4)	(3)	(1)	(0)
Ich finde, dass die verschiedenen Funktionen des Systems gut integriert sind.	(0)	(1)	(3)	(4)
Die Verwendung des Systems erleichtert mir die Erfüllung der Aufgabe.	(0)	(1)	(3)	(4)
Ich glaube, dass die meisten Leute das System schnell zu beherrschen lernen.	(0)	(1)	(3)	(4)
Ich empfinde die Bedienung als sehr umständlich.	(4)	(3)	(1)	(0)
Ich empfinde das System als einfach zu nutzen.	(0)	(1)	(3)	(4)
Durch das System fühle ich mich bei der Ausführung der Aufgabe behindert.	(4)	(3)	(1)	(0)

**Tabelle 7.1** System-Usability-Scale-Fragebogen mit den Punkten zu jeder Antwortmöglichkeit

Alle Punkte werden zusammengezählt und mit 2,5 multipliziert, um eine Skala von 0 bis 100 zu erhalten. Allgemein gilt, dass Ergebnisse von 68 den Durchschnitt bilden und Ergebnisse von 73 eine gute Nutzererfahrung widerspiegeln [40].

### 7.1.2.2 NASA Task Load Index

Der NASA Task Load Index (NASA-TLX) ist ein Fragebogen zur Bewertung der Arbeitsbelastung, der die in Tabelle 7.2 dargestellten Faktoren und Bewertungsoptionen umfasst [41]. Teilnehmer bewerten jeden Faktor auf einer Skala von 0 bis 10.

Faktor	0	1	2	3	4	5	6	7	8	9	10
Geistige Anforderungen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Körperliche Anforderungen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zeitliche Anforderungen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Leistungsdruck	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Anstrengung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Frustration	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Tabelle 7.2** NASA-Task-Load-Index-Fragebogen

Die Bewertungen der einzelnen Faktoren werden anschließend mit 10 multipliziert, gewichtet und gemittelt, um eine Gesamtnote für die Arbeitsbelastung zu erhalten. Für die Gewichtung werden die Probanden paarweise zu jeder Kombination von Faktoren nach dem belastenderen der beiden gefragt. Das Ergebnis wird dem belastenderen Faktor zugeordnet und aufsummiert. Zu jedem Faktor entsteht so ein Gewicht, je einer mit einem Gewicht von 0, 1, 2, 3, 4 und 5. Das Ergebnis ist

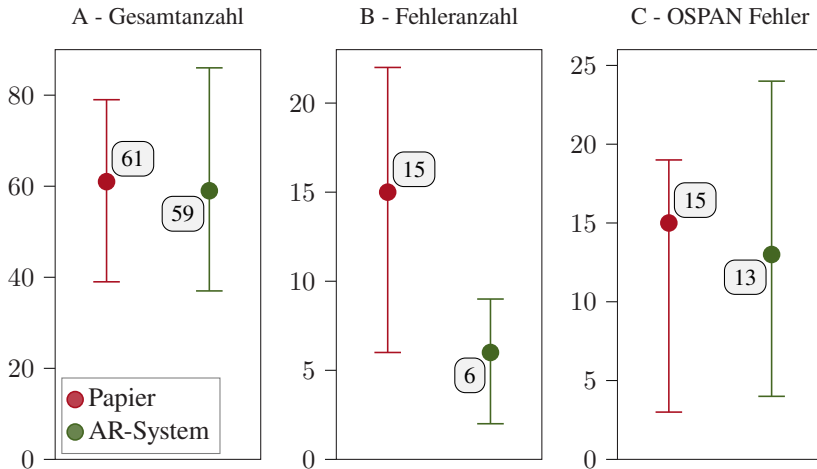
Wert	Bewertung
0 - 20	Sehr Geringe Belastung
20 - 40	Geringe Belastung
40 - 60	Moderate Belastung
60 - 80	Starke Belastung
80 - 100	Extreme Belastung

**Tabelle 7.3** Bewertungstabelle für den NASA-TLX [42]

dann eine Skala von 0 bis 100, wobei höhere Werte eine höhere wahrgenommene Arbeitsbelastung anzeigen. Tabelle 7.3 zeigt, wie das Ergebnis interpretiert wird [42].

## 7.2 Zwischenevaluation

Die Zwischenevaluation wurde nach der Fertigstellung des Prototyps in Kapitel 5 durchgeführt. Im Durchschnitt wurden mit Unterstützung der AR-Brille (61 Pillen) und dem PMP (59 Pillen) ähnlich viele Medikamente einsortiert (siehe Abbildung 7.2 A), was auf eine vergleichbare Leistung hindeutet. Die Varianz bei der Verwendung der AR-Brille war höher, was durch die individuelle Vertrautheit mit der Technologie erklärt werden kann.



**Abbildung 7.2** Evaluationsergebnisse der Zwischenevaluation mit der Gesamtanzahl der einsortierten Medikamente (A), der Anzahl der Sortier- und Verwechslungsfehler (B) und den Fehlern im OSPAN-Test (C) von maximal 30

Ein klarer Unterschied ist jedoch bei der Anzahl der Sortierfehler zu sehen (siehe Abbildung 7.2 B). Das Sortieren auf Basis von PMP führte im Median zu 15 Fehlern, während die Unterstützung durch die AR-Brille zu 6 Fehlern führte.

Bei der Lösung der OSPAN-Aufgabe (siehe Abbildung 7.2 C) waren die durchschnittlichen Fehlermengen in beiden Evaluationsaufgaben nahezu identisch (15 gegenüber 13). Obwohl einige Probanden eine höhere Fähigkeit zeigten, OSPAN und Sortieren parallel auszuführen als andere, konnte eine klare Korrelation zwischen einer niedrigeren OSPAN-Fehlerrate und einer höheren Rate von Pillensortierfehlern festgestellt werden. Allerdings erzielte kein Teilnehmer eine geringere Anzahl von Sortierfehlern mit dem PMP als mit der AR-Brille.

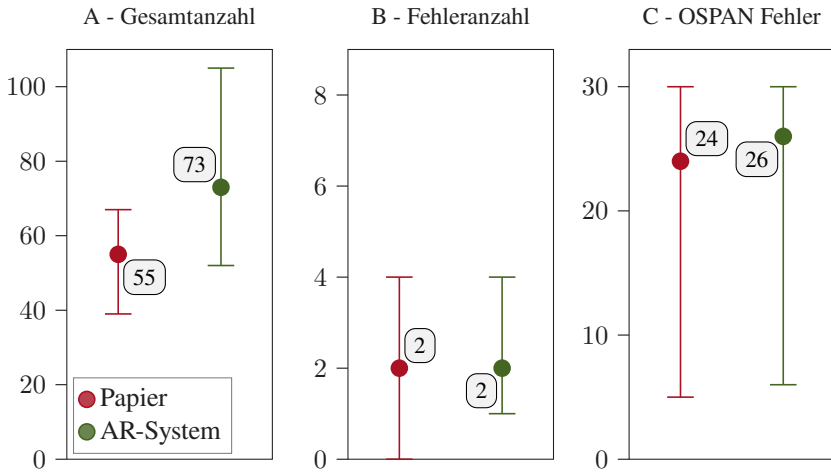
## **7.3 Endevaluation**

Die Endevaluation wurde mit 5 professionellen Pflegekräften und 19 Laien durchgeführt. Das Alter der Probanden deckt eine Spanne von Anfang 20 bis Ende 50 ab, wobei die Verteilung eine Tendenz zu jüngeren Probanden zeigt. Von den Probanden sind 11 männlich und 13 weiblich, von den professionellen Pflegekräften ist 1 männlich und 4 sind weiblich. Genau die Hälfte der Probanden gibt an, vor der Studie bereits mehr als einmal Tabletten gerichtet zu haben, dazu gehören alle 5 professionellen Pflegekräfte. Außerdem geben 15 Probanden an, neuer Technik gegenüber aufgeschlossen zu sein, davon allerdings nur 1 professionelle Pflegekraft.

### **7.3.1 Professionelle Pflegekräfte**

Die Ergebnisse der professionellen Pflegekräfte zeigen in Abbildung 7.3 einen klaren Anstieg der Gesamtanzahl an einsortierten Medikamenten von 55 Tabletten im Median mit dem PMP auf 73 Tabletten im Median bei Unterstützung durch das MMS. Auch in der Mindestanzahl, also der langsamsten Pflegekraft, zeigt sich der gleiche Trend in einer Steigerung von 39 Tabletten auf 51 Tabletten. Beim

Maximum, also der schnellsten Pflegekraft, ist der Trend sogar noch stärker mit einer Steigerung von 67 Tabletten auf 105 Tabletten.



**Abbildung 7.3** Evaluationsergebnisse der professionellen Pflegekräfte mit der Gesamtanzahl der einsortierten Medikamente (A), der Anzahl der Sortier- und Verwechslungsfehler (B) und den Fehlern im OSPAN-Test (C) von maximal 30

In der Fehleranzahl ist hingegen kein signifikanter Unterschied zu sehen. Sowohl mit dem PMP, als auch bei Unterstützung durch das MMS ist ein Median von 2 Fehlern zu sehen. Auch die maximale Fehleranzahl ist in beiden Fällen mit 4 Fehlern gleich. Lediglich in der kleinsten Fehleranzahl ist ein Unterschied zu sehen, hier steigt diese von 0 Fehlern auf 1 Fehler.

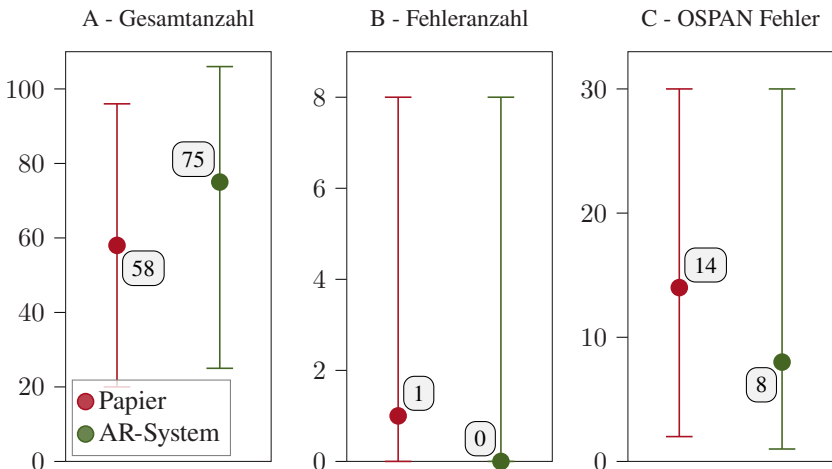
Das Maximum der OSPAN-Fehler liegt sowohl bei dem PMP als auch bei der Unterstützung durch das MMS auf dem Maximalwert von 30 falschen Antworten. Im Median liegen die Abweichungen bei 24 falschen Antworten mit dem PMP und bei 26 falschen Antworten mit der MMS Unterstützung. Allerdings gibt es starke Ausreißer nach unten. Der Minimalwert, also die Pflegekraft mit den geringsten Abweichungen, liegt beim PMP bei 5 falschen Antworten und bei der MMS Unterstützung bei 6 falschen Antworten.



### 7.3.2 Laien in der Pflege

Auch bei den Laien ist in Abbildung 7.4 der selbe Trend wie bei den professionellen Pflegekräften in der Gesamtanzahl der sortierten Medikamente zu beobachten. Hier steigt der Median von 58 Tabletten mit dem PMP auf 75 Tabletten bei Unterstützung durch das MMS. Das Minimum liegt bei den Laien sehr niedrig und die Steigerung ist weniger stark ausgeprägt. Mit dem PMP liegt das Minimum bei 19 Tabletten und bei Unterstützung durch das MMS bei 25. Der gleiche Trend ist auch beim Maximalwert zu beobachten, wo das Maximum mit dem PMP bei 96 Tabletten liegt und bei Unterstützung durch das MMS bei 106 Tabletten.

Bei der Fehlerrate ist genau wie bei den professionellen Pflegekräften kein großer Unterschied zwischen dem PMP und Unterstützung durch das MMS festzustellen. Aber, dem Trend der Gesamtanzahl an sortierten Medikamenten folgend, ist der Unterschied zwischen dem besten und dem schlechtesten erreichten Ergebnis deutlich größer als bei den professionellen Pflegekräften. Sowohl mit dem PMP als



**Abbildung 7.4** Evaluationsergebnisse der Laien mit der Gesamtanzahl der einsortierten Medikamente (A), der Anzahl der Sortier- und Verwechslungsfehler (B) und den Fehlern im OSPAN-Test (C) von maximal 30

auch bei der Unterstützung durch das MMS liegt die beste erreichte Fehleranzahl bei 0 Fehlern und die schlechteste bei 8 Fehlern. Lediglich im Median ist eine Verbesserung von 1 Fehler mit dem PMP auf 0 Fehler bei Unterstützung durch das MMS festzustellen.

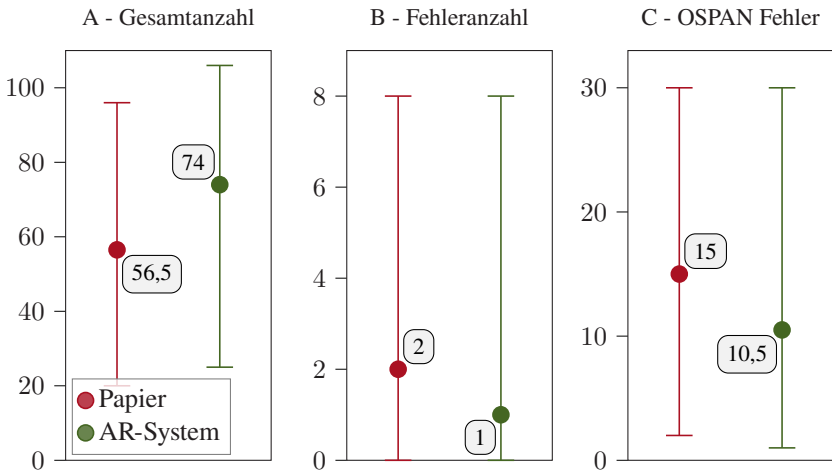
Die Anzahl an OSPAN-Fehlern ist bei den Laien, im Vergleich zu den professionellen Pflegekräften, deutlicher ausgeprägt. Der Medianwert sinkt von 14 falschen Antworten mit dem PMP auf 8 falsche Antworten bei Unterstützung durch das MMS. Allerdings liegt in beiden Fällen der schlechteste Wert auf der maximal erreichbaren Anzahl von 30 falschen Antworten. Dies ist das gleiche Ergebnis, das auch die professionellen Pflegekräfte erreicht haben. Die besten Ergebnisse liegen jedoch im Vergleich zu den professionellen Pflegekräften deutlich tiefer. Mit dem PMP lag die Abweichung bei 2 falschen Antworten, wohingegen bei Unterstützung durch das MMS die Abweichung bei 1 falschen Antwort lag.

### 7.3.3 Gesamtergebnisse

Zusammen genommen ist in Abbildung 7.5 bei den Laien und den professionellen Pflegekräften in allen drei Kategorien eine klare Verbesserung zu sehen. Die Gesamtanzahl der sortierten Medikamente ist relativ gesehen um 31% gestiegen. Allerdings zeigt sich die hohe Varianz der Laien in allen drei Kategorien auch in den Gesamtergebnissen. Klar zu sehen ist, wie die Anzahl in den Sortier- und Verwechslungsfehlern relativ gesehen um 50% und in den OSPAN-Fehlern um 30% gesunken ist, was die Effizienz in diesem Szenario eindeutig beweist.

## 7.4 2-Stunden-Evaluation

Zusätzlich zur Evaluation unter Stressbedingungen wird in einer weiteren Evaluation untersucht, ob die Ergebnisse unter störungsfreien Bedingungen ebenfalls eine Verbesserung zeigen. Außerdem soll festgestellt werden, ob das System auch

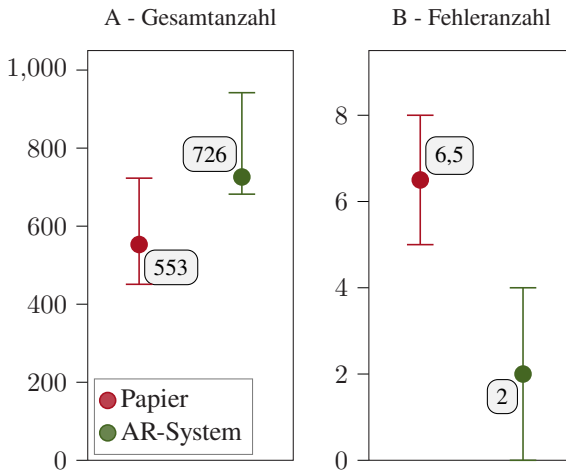


**Abbildung 7.5** Evaluationsergebnisse der Laien und professionellen Pflegekräfte kombiniert mit der Gesamtanzahl der einsortierten Medikamente (A), der Anzahl der Sortier- und Verwechslungsfehler (B) und den Fehlern im OSPAN-Test (C) von maximal 30

über einen längeren Zeitraum echte Unterstützung leistet, um Anfangsschwierigkeiten und Lerneffekte ausschließen zu können. Dazu wurden insgesamt vier Probanden gebeten jeweils zwei Stunden mit dem PMP und mit der Unterstützung durch das MMS Tabletten zu richten.

In Abbildung 7.6 ist zu sehen, dass die Gesamtanzahl der sortierten Medikamente vorhersehbarerweise deutlich höher ist als bei der 15-minütigen Evaluation. Trotzdem ist derselbe Trend zu beobachten und der Median steigt von 553 Tabletten mit dem PMP auf 726 Tabletten bei Unterstützung durch das MMS. Die Mindestanzahl steigt von 451 Tabletten mit dem PMP auf 682 Tabletten bei Unterstützung durch das MMS und der Maximalwert von 723 Tabletten mit dem PMP auf 942 Tabletten bei Unterstützung durch das MMS.

Die Fehleranzahlen liegen trotz längerer Bearbeitungszeit in einem ähnlichen Rahmen wie bei der 15-minütigen Evaluation (siehe Abbildung 7.6 B). Dabei zeigen sowohl der Median als auch die jeweiligen Maximal- und Minimalwerte klare Verbesserungen durch den Einsatz des MMS.



**Abbildung 7.6** 2-Stunden-Evaluationsergebnisse der Laien und professionellen Pflegekräfte kombiniert mit der Gesamtanzahl der einsortierten Medikamente (A) und der Anzahl der Sortier- und Verwechslungsfehler (B)

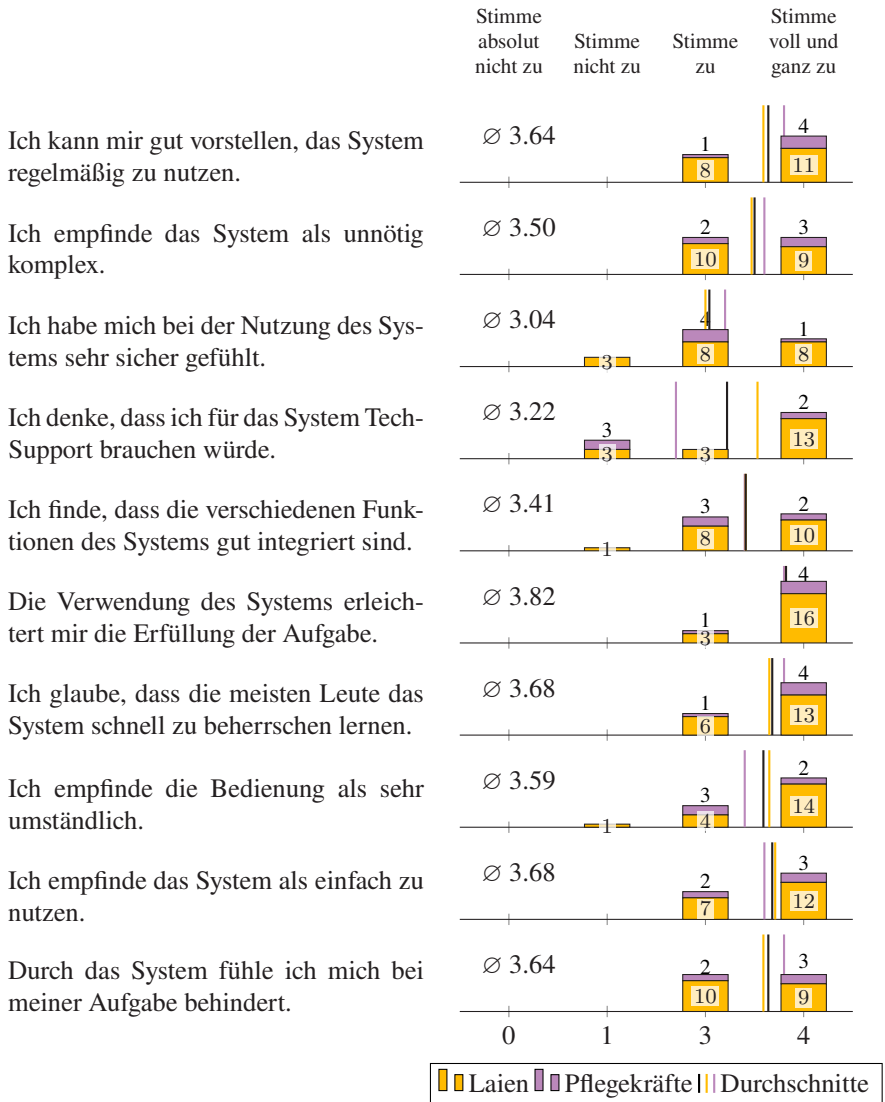
## 7.5 Fragebögen Antworten

Jeder Proband wurde gebeten einen SUS-Fragebogen, einen NASA-TLX Fragebogen auszufüllen. Da die Freitextfelder jedoch optional waren, gibt es hierbei nicht von jedem Proband Rückmeldung.

### 7.5.1 System Usability Scale

Tabelle 7.4 zeigt die Ergebnisse der System Usability Scale (SUS) Befragung für die zwei verschiedenen Benutzergruppen: Laien und professionelle Pflegekräfte. Hierbei sind negative Aussagen bereits invertiert.

Für die Laien lagen die höchsten Bewertungen in den Aussagen »Erleichtert die Aufgabe« (3,84), »Kann schnell beherrscht werden« (3,68) und »Bedienung ist umständlich« sowie »Einfache Nutzung« (jeweils 3,63). Dies deutet darauf



**Tabelle 7.4** Ergebnisse des SUS-Fragebogens als Histogramme mit Mittelwerten. Negative Fragen sind bereits invertiert. Ø und die schwarzen Markierungen geben die Durchschnittswerte aller Befragten an. Das Endergebnis ist ein SUS-Wert von 88,02.

hin, dass die Laien die Benutzerfreundlichkeit des Systems insgesamt als positiv einschätzen.

Die höchste Zustimmung bei den Pflegekräften erhielten die Aussagen »Würde ich regelmäßig nutzen« , »Erleichtert die Aufgabe« , »Kann schnell beherrscht werden« und »Aufgabe wird behindert« (jeweils 3,8). Die Bewertung für »Bräuchte Technischen Support« war mit 2,2 am niedrigsten, was darauf hindeutet, dass die Pflegekräfte hier den größten Verbesserungsbedarf sehen.

Beim Vergleich der beiden Gruppen fällt auf, dass Pflegekräfte tendenziell höhere Bewertungen für die Sinnhaftigkeit des Systems abgegeben haben. Insbesondere bei den Aussagen »Würde ich regelmäßig nutzen« und »Aufgabe wird behindert« erzielten Pflegekräfte höhere Durchschnittswerte als Laien. Im Gegensatz dazu ist die Benutzerfreundlichkeit bei den Pflegekräften jedoch niedriger bewertet. Insbesondere die Aussagen »Bedienung ist umständlich« und »Benötige Tech-Support« fallen hier auf und zeigen, dass die Nutzererfahrung noch Verbesserungspotenzial aufweist.

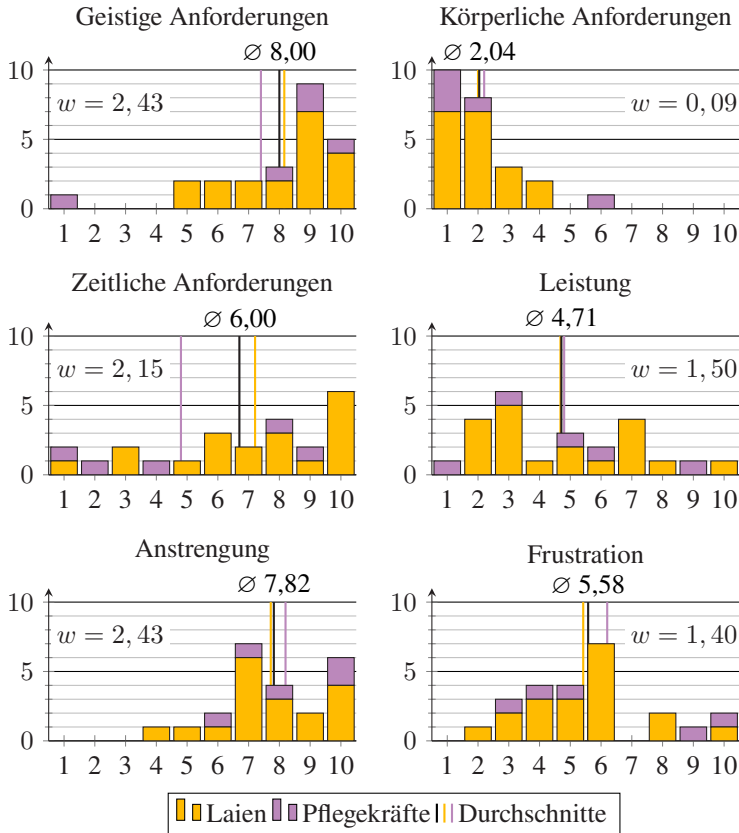
Insgesamt deuten die Ergebnisse darauf hin, dass beide Gruppen das System als insgesamt benutzerfreundlich empfinden, wobei Pflegekräfte insgesamt etwas zufriedener mit dem System sind als die Laien.

### 7.5.2 NASA Task Load Index

Die Ergebnisse in Abbildung 7.7 zeigen, dass sowohl Laien als auch Pflegekräfte hohe geistige Anforderungen und eine hohe Anstrengung mit  $w = 2,43$  wahrnehmen. Die sehr niedrige Gewichtung von  $w = 0,09$  für körperliche Anforderungen verdeutlicht, dass die körperliche Belastung als gering empfunden wird.

Im Vergleich zwischen Laien und Pflegekräften ist zu beobachten, dass die geistigen Anforderungen - mit  $\varnothing = 7,4$  für Pflegekräfte und  $\varnothing = 8,2$  für Laien - und die zeitlichen Anforderungen - mit  $\varnothing = 4,8$  für Pflegekräfte und  $\varnothing = 7,2$  für Laien - von den Pflegekräften deutlich weniger intensiv wahrgenommen wurden.

Im Gegensatz dazu zeigen die Kategorien Anstrengung - mit  $\bar{x} = 8,2$  für Pflegekräfte und  $\bar{x} = 7,7$  für Laien - und Frustration - mit  $\bar{x} = 6,2$  für Pflegekräfte und  $\bar{x} = 5,4$  für Laien - einen Anstieg in der Wahrnehmung der Pflegekräfte im Vergleich zu den Laien.



**Abbildung 7.7** Ergebnisse der Anforderungs- und Belastungsanalyse als Histogramme.  $\bar{x}$  und die schwarzen Markierungen geben die Mittelwerte aller Befragten auf der x-Achse an, die gelbe Markierung die Mittelwerte für Laien und lila für die Pflegekräfte.  $w$  gibt die berechneten Gewichtungsfaktoren an. Das Endergebnis ist ein Gesamtwert von 67,93.

Das Endergebnis von 67,93 liegt nach Tabelle 7.3 in Kapitel 7.1 eindeutig in der Kategorie »Starke Belastung«. Dies zeigt, dass vor allem die Konzentration bei der Erfüllung der Aufgabe enorm beansprucht wird und nicht über längere Zeiträume aufrechterhalten werden kann.

### 7.5.3 Freitextrückmeldungen

Eine vollständige Übersicht der Freitextrückmeldungen ist im Anhang A.1 angegeben. Die Rückmeldungen lassen sich in drei Kategorien einordnen (das Evaluationsvorgehen, die Erkennungsleistung des MMS und die Nutzerschnittstelle).

Die Probanden empfanden bei der **Evaluation** ein hohes Maß an Stress. Jedoch meldete eine Pflegekraft zurück, dass es sich bei der Intensität des OSPAN-Tests um »realistische Ablenkungen« handle.

Die Rückmeldung zur **Erkennungsleistung** war geprägt von Meldungen über langsame Verarbeitungszeiten und Fehlerkennungen. Obwohl das System insgesamt ein Gefühl der Sicherheit vermittelt, können einzelne Fehlerkennungen dennoch verunsichern.

Zur **Nutzerschnittstelle** wurden die meisten Rückmeldungen gegeben. Viele Vorschläge wie »Fotos der Medikamente im Interface neben den Namen würden helfen« und »Anstatt von Nullen sollten im Interface leere Felder angezeigt werden.« wurden abgegeben. Die häufigste Anmerkung war, dass die Nutzerschnittstelle sehr langsam aktualisiert wird, wenn neue Daten geladen werden.



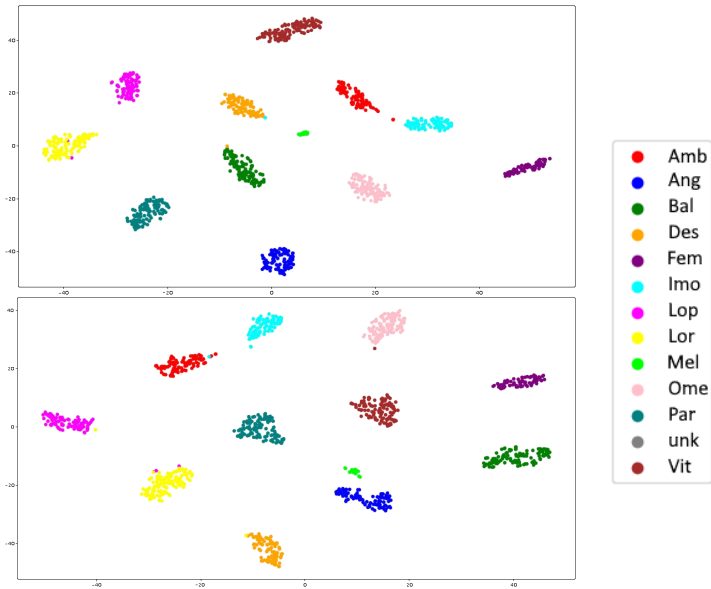
## 8 Ergebnisse

In diesem Kapitel wird zunächst die Skalierbarkeit des MMS präsentiert und die Anforderungserfüllung überprüft, gefolgt von einer Diskussion der Ergebnisse der Nutzerumfrage. Abschließend werden die möglichen nächsten Schritte und eine mögliche wirtschaftliche Verwertung aufgezeigt.

### 8.1 Skalierbarkeit

Um die notwendigen Medikamente für einen Realeinsatz des MMS zu erkennen, müssen neue Datenklassen mit in das Modell eintrainiert werden. Zum Hinzufügen einer neuen Datenklasse wird jedoch nicht die gleiche Menge an Trainingsdaten benötigt, wie es beim allerersten Training der Fall ist. Um zu testen, wie viele Bilder genau notwendig sind, wird das Netz mit unterschiedlichen Datenmengen trainiert. Anschließend wird die in Kapitel 6.3.2 gezeigte t-SNE-Visualisierung verwendet, um festzustellen, wie gut die Differenzierung bei unterschiedlichen Datenmengen gelingt.

Abbildung 8.1 zeigt die t-SNE Visualisierung der Versuche mit 0 Bildern der neuen Klasse und mit 5 Bildern der neuen Klasse in den Trainingsdaten. Zur Visualisierung werden Exemplare der neuen Daten in den Testdaten verwendet. Dabei ist zu sehen, dass bereits ohne ein Bild in den Trainingsdaten, das Modell die neue Medikamentenklasse größtenteils korrekt gruppiert, aber einzelne Datenpunkte zusammen mit anderen Klassen gruppiert werden. Erst bei fünf Exemplaren der neuen Klasse in den Trainingsdaten werden alle Vorkommen korrekt klassifiziert und gruppiert.



**Abbildung 8.1** Die t-SNE Visualisierung der Merkmalsrepräsentationen des ResNet mit 0 Exemplaren (oben) und 5 Exemplaren (unten) der neuen Klasse (hellgrün).

## 8.2 Überprüfen der Anforderungen

Um zu bewerten, ob alle Ziele erreicht worden sind, werden die in Kapitel 4.1 erarbeiteten Anforderungen noch einmal überprüft.

So konnten unter den funktionellen Anforderungen fast alle erfolgreich umgesetzt werden. Insbesondere die Forderung nach einer platzsparenden Umsetzung (A1.8 und A1.9) konnte den geforderten Arbeitsbereich von  $1\text{m}^2$  weit unterschritten werden. Lediglich die Updateinfrastruktur für das System (A1.6) wurde im Rahmen dieser Arbeit nicht umgesetzt. Im Bereich der qualitativen Anforderungen wurden ebenfalls die meisten Ziele erreicht. Die Benutzeroberfläche erwies sich als intuitiv bedienbar (A2.1), und Pflegekräfte konnten ihre Arbeitsgeschwindigkeit durch den Einsatz des MMS signifikant steigern (A2.2 und A2.3), wobei die angestrebte

Zeitersparnis von 20% sogar übertroffen wurde, während die Fehlerrate unter den geforderten Grenzwerten blieb (A2.6 und A2.7). Die Skalierung des MMS (A2.4 und A2.5) wurde in dieser Arbeit theoretisch gezeigt, jedoch ohne die geforderte Anzahl von 100 Medikamenten zu erreichen. Die Sicherheitsanforderungen hingegen wurden vollständig erfüllt, insbesondere hinsichtlich der Datenverarbeitung auf lokalen Servern (A3.3 und A3.4).

## 8.3 Interpretation der Nutzerumfrage

Die Analyse der Gesamtergebnisse zeigt deutlich, dass das MMS in allen gemessenen Kategorien die Leistung sowohl bei Laien als auch bei professionellen Pflegekräften verbessert. Besonders hervorzuheben ist die signifikante Steigerung der sortierten Medikamentenanzahl, was auf eine effizientere Nutzung der verfügbaren Zeit hinweist. Diese Steigerung verdeutlicht, dass das MMS den Sortierprozess beschleunigt und intuitiver gestaltet, was seine wesentliche Fähigkeit zur Unterstützung unterstreicht.

Die hohe Varianz, insbesondere bei den Laien, deutet auf unterschiedliche Fähigkeiten und Vertrautheit mit dem Prozess hin, was auf eine individuell stark variierende Lernkurve schließen lässt. Dennoch ist positiv anzumerken, dass die minimalen und maximalen Werte unter AR-Unterstützung durchweg höher ausfallen als ohne, was die Flexibilität und Anpassungsfähigkeit des MMS an verschiedene Nutzergruppen zeigt. Die Reduzierung der Fehleranzahl, besonders im Median, bestätigt den praktischen Nutzen des MMS. Diese Verbesserung bei beiden Nutzergruppen unterstreicht die allgemeine Effektivität der AR-Unterstützung.

Bemerkenswert ist, dass die OSPAN-Fehler zwischen den Zeitabschnitten weniger stark schwanken. Dies könnte darauf hindeuten, dass die kognitive Belastung beim Medikamentenrichten auch mit AR-Unterstützung weiterhin hoch bleibt. Eine Weiterentwicklung des MMS könnte darauf abzielen, diese mentale Belastung weiter zu reduzieren. Dennoch zeigt sich bereits eine messbare Entlastung im Vergleich zum PMP.

Die 2-Stunden-Evaluation verdeutlicht, dass das MMS unter störungsfreien Bedingungen effizienter und konsistenter arbeitet, indem es sowohl die Medikamentensortierung als auch die Fehleranzahl verbessert. Dabei bestätigt die 2-Stunden-Evaluation die positiven Trends aus den stressbasierten Kurzzeittests, wodurch deren Ergebnisse validiert werden. Während die 15-Minuten-Tests Schwächen bei der Tablettenerkennung und der Interface-Stabilität aufdeckten, arbeiten die Probanden in weniger hektischen Umgebungen zuverlässiger mit dem MMS.

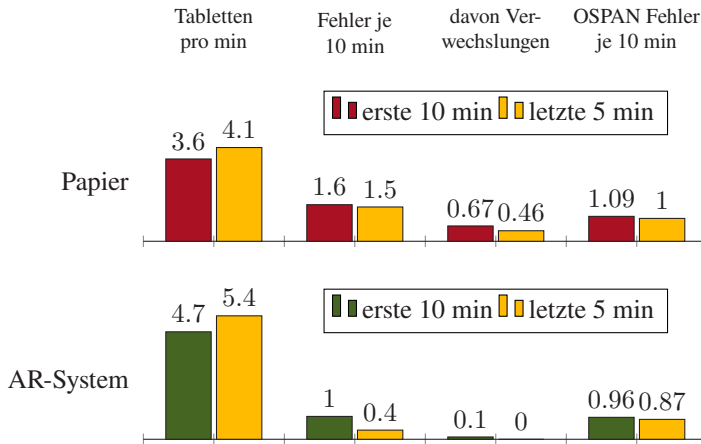
Insgesamt bieten die Ergebnisse klare Hinweise darauf, dass das MMS einen erheblichen Mehrwert im Medikationsmanagement darstellt. Die gesteigerte Effizienz und Genauigkeit sprechen für eine Implementierung in der Praxis. Weitere Optimierungen, insbesondere zur Reduktion der kognitiven Belastung, könnten die Effektivität des MMS noch weiter steigern.

### 8.3.1 Zeitabhängigkeit der Probandenleistung

In der folgenden Analyse der Medikamentensortierung wird die Konzentration der Teilnehmer durch einen Vergleich der ersten 10 Minuten mit den letzten 5 Minuten jedes Durchgangs bewertet. In Abbildung 8.2 sind die jeweiligen Raten von Tabletten und Fehlern des gemeinsamen Datensatzes aus Kapitel 7.3.3 zu sehen.

Die Auswertung zeigt, dass bereits zu Beginn der Tabletten-Sortierung das MMS mit durchschnittlich 4,65 einsortierten Tabletten pro Minute gegenüber 3,58 Tabletten pro Minute im papierbasierten System überlegen war. Auch in den letzten 5 Minuten blieb das MMS effizienter, mit einer durchschnittlichen Anzahl von 5,4 einsortierten Tabletten pro Minute im Vergleich zu 4,1 beim PMP. Dies zeigt, dass es sowohl beim Durchgang mit dem PMP als auch beim Durchgang mit dem MMS gleichermaßen einen Lerneffekt innerhalb der Durchgänge gibt. Der prozentuale Anteil liegt jedoch mit 14% in beiden Fällen gleichauf.

Die ersten 10 Minuten kennzeichneten sich auch durch eine geringere Fehlerzahl beim MMS, mit einem Mittelwert von 1,0 Fehler je 10 Minuten, verglichen mit 1,6



**Abbildung 8.2** Vergleich der Ergebnisse der ersten 10 Minuten mit den letzten 5 Minuten, wobei die Skala zum Vergleich normiert ist.

beim PMP. Die Fehleranzahl sank in der Endphase beim MMS weiter auf einen Durchschnitt von 0,4 Fehler je 10 Minuten, während der PMP noch immer einen Wert von 1,5 Fehler je 10 Minuten aufwies. Hier zeigt sich, dass der Lernerffekt mit dem MMS sogar noch stärker ist, als mit dem PMP. Insbesondere fällt auf, dass die Verwechslungsfehler mit dem MMS quasi komplett eliminiert werden konnten. Bei diesen richtet ein Proband aus Versehen den gleichen Patienten zweimal hintereinander oder lässt einen Patienten ganz aus.

Die Anzahl der Fehler im OSPAN-Test blieb über beide Methoden hinweg gleich und zeigt auch kaum Unterschiede zwischen den ersten 10 oder den letzten 5 Minuten. Da der OSPAN-Test aber nur zur Ablenkung beim Richten eingesetzt wurde, ohne dass die genaue Natur der Ablenkung eine größere Rolle spielt, werden diese Werte nicht weiter interpretiert.

### 8.3.2 Interpretation der Rückmeldungen

Die Ergebnisse des SUS- und des NASA-TLX-Fragebogens in Kapitel 7.1.2 bieten wertvolle Einblicke in die Benutzerfreundlichkeit und Arbeitsbelastung des MMS. In den Freitextrückmeldungen hatten die Probanden die Gelegenheit, gezielt auf für sie wichtige Aspekte einzugehen.

Im Rahmen der SUS-Befragung bewerteten sowohl Laien als auch professionelle Pflegekräfte das MMS überwiegend positiv. Besonders die Erleichterung der Aufgabe und die schnelle Erlernbarkeit des MMS erhielten hohe Zustimmungswerte. Dies zeigt, dass die grundlegende Benutzerfreundlichkeit des MMS anerkannt wird, wobei insbesondere Pflegekräfte häufig angaben, das MMS regelmäßig nutzen zu wollen. Trotz der positiven Rückmeldungen wurden Unterschiede deutlich: Pflegekräfte kritisierten die Bedienung als teils umständlich und sahen Bedarf an technischem Support, was auf spezifische Anforderungen dieser Nutzergruppe hinweist.

Die NASA-TLX-Ergebnisse verdeutlichen die wahrgenommene Arbeitsbelastung während der Nutzung des MMS. Sowohl Laien als auch Pflegekräfte empfanden die geistigen Anforderungen und den Aufwand als hoch, wobei Pflegekräfte die zeitlichen Anforderungen als weniger belastend wahrnahmen – möglicherweise aufgrund ihrer größeren Erfahrung. Gleichzeitig berichteten sie jedoch von höherer Anstrengung und Frustration, insbesondere mit flickernden (Fehl-)Klassifikationen und langsamer Benutzeroberfläche. Dies weist darauf hin, dass das MMS trotz seiner Effizienzsteigerungen zusätzliche kognitive Belastungen verursacht.

Die Freitextrückmeldungen bestätigen, dass das MMS als hilfreich und sicher wahrgenommen wird, jedoch Schwächen bei der Benutzeroberfläche und der Tablettenerkennung aufweist. Kritisiert wurden insbesondere die Größe und Langsamkeit des Interfaces sowie Schwierigkeiten bei der Erkennung weißer oder horizontal liegender Tabletten. Dennoch lobten viele Nutzer die Übersichtlichkeit

und die Unterstützung bei hektischen Aufgaben. Vorschläge zur Verbesserung umfassten eine flexiblere Skalierung der Benutzeroberfläche und eine Optimierung der Erkennung.

Zusammengefasst zeigen alle drei Evaluationsinstrumente, dass das MMS insgesamt positiv bewertet wird, aber noch Verbesserungspotenzial besitzt. Besonders die Benutzerfreundlichkeit und die Reduktion der subjektiven Frustration sollten weiter vorangetrieben werden, um das MMS langfristig für alle Anwendergruppen attraktiv und effizient zu gestalten.

## 8.4 Nächste Schritte

Bei einem abgeschlossenen System, wie es hier vorgestellt wird, ist ein offensichtliches Ziel die Umsetzung in der Praxis. Für eine wirtschaftliche Verwertung ist jedoch eine Weiterentwicklung unabdingbar.

### 8.4.1 Mögliche Weiterentwicklungen

Die möglichen Weiterentwicklungen lassen sich unterteilen in notwendige Schritte zur wirtschaftlichen Verwertung und interessante, aber optionale Erweiterungen. Um das MMS erfolgreich in der Praxis einsetzen zu können, sollten zunächst die folgenden Weiterentwicklungen durchgeführt werden:

- **Skalierung:** Um das MMS für eine breite Palette von Medikamenten nutzbar zu machen, sollte eine umfassende Datenbank mit Bildern aller relevanten Tabletten erstellt werden. Besonders bei ähnlich aussehenden Medikamenten sollte dabei auf eine breite Datenbasis geachtet werden.
- **Hardwareaufbau:** Das MMS soll für den Einsatz in Pflegeeinrichtungen kosteneffizient sein. Daher sollte ein modularer und robuster Hardwareaufbau entwickelt werden, der einfach zu installieren ist und sich an die räumlichen Gegebenheiten der jeweiligen Einrichtung anpassen lässt.

- **Pflegemanagement-Systeme:** Zur Einbindung in das Pflegemanagement-System sollte eine Schnittstelle entwickelt werden, die den Medikationsplan automatisch mit dem MMS synchronisiert. Diese Schnittstelle sollte nicht nur die Abfrage der richtigen Medikation ermöglichen, sondern auch die Dokumentation ablegen. Eine solche Schnittstelle muss für jeden Pflegemanagementsoftwareanbieter einzeln entwickelt werden, da in diesem Bereich keine Standards existieren.
- **Vernetzung:** Um die Verteilung von neuen Modellen zu ermöglichen, sollte eine Infrastruktur zur Aktualisierung entwickelt werden. Diese kann nicht nur ein MMS im Betrieb zur Erkennung neuer Medikamente befähigen, sondern möglicherweise auch Feedback einsammeln.

Des Weiteren ist es möglich, den Einsatz von föderiertem Lernen (englisch. federated Learning) zu untersuchen, um fehlende Medikamente direkt in den Einrichtungen zum MMS hinzufügen zu können. Solche weitertrainierten Modelle können dann mit der Infrastruktur zur Aktualisierung zentral zusammengefasst werden und an alle Einrichtungen verteilt werden.

Eine andere Möglichkeit ist die Entwicklung oder Einbindung einer Datenbank für Medikamentenwechselwirkungen. Da viele Medikamente von unterschiedlichen Ärzten verschrieben werden, können erst die Pflegekräfte bei der Vergabe auf Wechselwirkungen eingehen, und eine solche Erweiterung kann sie dabei unterstützen.

## 8.4.2 Verwertungsplan

Um eine wirtschaftliche Verwertung zu erreichen, sollten zunächst die in Kapitel 8.4.1 genannten Punkte umgesetzt werden. Nach der Entwicklung kann die Verwertung nach folgendem Plan erfolgen:

Kurzfristig (1-2 Jahre) wird schrittweise durch (bezahlte) Pilotprojekte in ausgewählten Pflegeeinrichtungen die Technologie unter realen Bedingungen getestet



und optimiert. Schulungsprogramme für Pflegekräfte mit praktischen Anleitungen und technischem Support fördern die Akzeptanz. Kooperationen mit Gesundheitsinstitutionen, Pflegeverbänden und Technologieanbietern sind entscheidend für den breiten Markteintritt. Die Technologie wird flexibel und skalierbar entwickelt, um verschiedenen Benutzerbedürfnissen gerecht zu werden.

Mittelfristig (3-5 Jahre): Der Markteintritt ist erfolgt und die Technologie wird durch Geräteverkauf und Lizenzverträge in weiteren Pflegeeinrichtungen implementiert. Partnerschaften mit großen Pflegeeinrichtungen und weiteren Multiplikatoren beschleunigen dabei die Verbreitung.

Langfristig (5+ Jahre): Langfristig wird die Technologie international vermarktet, zunächst in Europa. Die Anpassung an gesetzliche Anforderungen und Marktbefürfnisse erfolgt durch lokale Partnerschaften. Durch kontinuierliche Weiterentwicklung und Anpassung wird ein nachhaltiger wirtschaftlicher Erfolg sichergestellt.

### **8.4.3 Wirtschaftlichkeit**

Der Mindestlohn für eine ausgebildete Pflegefachkraft ist im Jahr 2024 auf 19,50 € pro Stunde gestiegen und soll 2025 weiter auf 20,50 € steigen [15]. Viele Pflegekräfte können, je nach individuellen Vereinbarungen, darüber hinausgehend bis zu 30 € pro Stunde verdienen. Interviews mit Pflegefachkräften haben ergeben, dass das Richten von Medikamenten jeden Tag zwischen 30 und 60 Minuten in Anspruch nimmt. Bei den in Kapitel 8.3 genannten Einsparungen von 30% der benötigten Zeit kann im Schnitt von einer Einsparung von mindestens 10 Minuten pro Tag ausgegangen werden. Pro Jahr kann ein Pflegeheimbetreiber so pro Pflegekraft etwa 1.200 € einsparen. Dabei werden die Vorteile der verbesserten Patientensicherheit nicht betrachtet.

Umsatz	Jahr 1	Jahr 2	Jahr 3	Jahr 4	Jahr 5
Systemverkäufe	10	70	1.40	2.20	3.20
Lizenzeinnahmen	10 tsd.€	80 tsd.€	220 tsd.€	440 tsd.€	760 tsd.€
Gesamtumsatz	110 tsd.€	780 tsd.€	1.620 tsd.€	2.640 tsd.€	3.960 tsd.€

**Tabelle 8.1** Vorhersage des Unternehmensumsatzes in den ersten Jahren nach einer Ausgründung.  
Preis pro System: 10.000N.

Damit können Lizenzkosten von 1.000€pro Jahr allein durch die Einsparungen finanziert werden. Dazu kommen initiale Anschaffungskosten von 10.000 €, die in ersten Interviews mit einzelnen Pflegeheimbetreibern bereits als realistisch eingeordnet wurden.

Auch wenn die initialen Anschaffungskosten sich rechnerisch nur sehr langsam amortisieren, kann der Mehrwert durch die gestiegene Sicherheit und konsistente Dokumentation hier als Argument ausreichen.

In Deutschland gab es 2021 16.115 Pflegeheime [43]. Langfristig wird ein Wachstum der Marktdurchdringung von 2% pro Jahr angestrebt. Mit diesen Annahmen kann der in Tabelle 8.1 gezeigte Umsatz erzielt werden. Damit können ab dem zweiten Jahr die Personalkosten für ein initiales Gründerteam von 3 Leuten getragen werden und ab dem dritten Jahr kann das Team gewinnbringend expandiert werden.

## 9 Zusammenfassung

In dieser Arbeit wurde ein System entwickelt, das die manuelle Medikamentensortierung durch den Einsatz von Augmented Reality und künstlicher Intelligenz unterstützt. Ziel war es, eine Lösung zu schaffen, mit der es möglich ist, Medikamente effizienter und mit geringerer Fehlerquote zu sortieren. Dazu nutzt das System visuelle Anweisungen, die direkt auf den Arbeitsbereich projiziert werden, während KI-Algorithmen die Erkennung und Klassifikation der Medikamente übernehmen.

Die Systemarchitektur besteht aus mehreren Komponenten: Eine Kamera erfasst den Arbeitsbereich, während ein Beamer Anweisungen direkt auf die Arbeitsfläche projiziert. Ein lokaler Server übernimmt die Bildverarbeitung und -analyse, verarbeitet die Daten und sendet sie an die Benutzeranwendung zurück.

Die Implementierung erfolgt in Python und nutzt das Mikro-Web-Framework Flask, um die Kommunikation zwischen Benutzeranwendung und Server zu gewährleisten. Auf dem Server wird in einer Kaskade aus Vorverarbeitungsschritten und mehreren KI-Modellen die Erkennung ausgeführt.

Die Kalibrierung des Systems, einschließlich der Berechnung der Homographiematrix zur Ausrichtung des Projektions-Interfaces, verläuft vollautomatisch. Die Erkennung der Arbeitsfläche und die daraus folgende Berechnung der Transformationsmatrix ermöglichen eine präzise Projektion und Interaktion.

Die Wahl der YOLOv5-Architektur erfolgte aufgrund ihrer hohen Geschwindigkeit und Genauigkeit. Das Modell erlaubt eine effiziente Objekterkennung, wodurch Echtzeitfähigkeit gewährleistet ist. Da YOLOv5 das Bild in zu geringer

Auflösung analysiert, werden relevante Bildausschnitte extrahiert und durch ResNet weiterverarbeitet. ResNet übernimmt die exakte Klassifikation der einzelnen Medikamente.



**Abbildung 9.1** Pflegekraft beim Richten von Medikamenten mit dem vorgestellten System. Verwendung mit Zustimmung der abgebildeten Person.

## 9.1 Ergebniszusammenfassung

Auf technischer Ebene erzielten die eingesetzten KI-Modelle vielversprechende Ergebnisse und die Kombination von YOLOv5 zur Detektion und ResNet zur Klassifikation der Medikamente erwies sich als effektiv. YOLOv5 konnte die Position der Medikamente im Bild zuverlässig identifizieren, während ResNet zur genauen Klassifikation der einzelnen Medikamente eingesetzt wurde. Dies führte zu einer hohen Erkennungsgenauigkeit bei minimalem Rechenaufwand.

Das System zeigte während der Testläufe eine stabile Leistung und konnte in unterschiedlichen Szenarien zuverlässig arbeiten. Es wurde festgestellt, dass das System

skalierbar ist und durch Erweiterungen der Trainingsdatensätze sowie die Integration zusätzlicher Datenklassen weiter optimiert werden kann. Eine vollständige Skalierung auf die benötigte Anzahl an Medikamenten für den praktischen Einsatz konnte im Rahmen dieser Arbeit jedoch noch nicht realisiert werden.

Um die praktische Wirksamkeit des Systems zu bewerten, wurden umfangreiche Nutzerevaluationen mit professionellen Pflegekräften und Laien durchgeführt. Die Teilnehmer sortierten Medikamente sowohl mit als auch ohne Unterstützung des AR-Systems. Bewertet wurde die Leistung anhand verschiedener Metriken, wie der Anzahl korrekt sortierter Medikamente und der Anzahl von Fehlern.

Die Ergebnisse zeigen, dass das AR-unterstützte System die Leistung der Nutzer in allen getesteten Kategorien signifikant verbessert hat:

- **Effizienz:** Die Anzahl korrekt sortierter Medikamente stieg signifikant. Bei professionellen Pflegekräften erhöhte sich die durchschnittliche Anzahl sortierter Medikamente von 55 auf 73 pro Testdurchgang.
- **Fehlerquote:** Die Fehleranzahl wurde drastisch reduziert, von durchschnittlich 2,4 auf 1,25 Fehler pro Testdurchgang. Insbesondere Laien machten mit Unterstützung des Systems deutlich weniger Fehler.
- **Benutzerfreundlichkeit:** Die Teilnehmer bewerteten das System als nützlich und unterstützend, identifizierten jedoch Verbesserungspotenzial bei der Benutzeroberfläche und der Systemstabilität.

Eine detaillierte Analyse zeigte außerdem, dass die Systemleistung auch über längere Zeiträume hinweg stabil blieb. Darüber hinaus konnte die kognitive Belastung der Nutzer reduziert werden, was sich in positiven subjektiven Bewertungen der Teilnehmer widerspiegelte.

## 9.2 Fazit

In dieser Arbeit wurde ein innovatives System zur Unterstützung der manuellen Medikamentensortierung vorgestellt, das Augmented Reality und künstliche Intelligenz kombiniert, um die Effizienz und Sicherheit des Prozesses zu steigern. Die zentrale Forschungsfrage, wie die Fehleranzahl beim Medikamentenrichten verringert werden kann, wurde erfolgreich beantwortet und die zentrale Hypothese bestätigt. Die Ergebnisse verdeutlichen, dass die Integration von visuellen AR-Anweisungen und modernen KI-Algorithmen, wie YOLOv5 zur Objekterkennung und ResNet zur exakten Klassifikation, zu einer signifikanten Leistungssteigerung führt.

Das System überzeugte in den Tests durch Stabilität, hohe Erkennungsgenauigkeit und Benutzerfreundlichkeit. Besonders hervorzuheben ist die deutliche Reduzierung der Fehlerquote sowie die Effizienzsteigerung bei professionellen Pflegekräften. Die durchgeführten Nutzerevaluationen bestätigten, dass das System nicht nur die korrekte Sortierung von Medikamenten verbessert, sondern auch die kognitive Belastung der Nutzer reduziert.

Trotz dieser Erfolge besteht weiterhin Potenzial für Verbesserungen. Die Skalierbarkeit des Systems für den Einsatz in der Praxis sollte weiter untersucht werden, und es gibt Möglichkeiten, die Benutzeroberfläche und die Systemstabilität weiter zu optimieren. Besonders wichtig ist die Erweiterung der Trainingsdatensätze und die Integration zusätzlicher Medikamentenklassen, um das System in der Praxis vollumfänglich einsetzbar zu machen.

Abschließend leistet diese Arbeit einen wertvollen Beitrag zur Entwicklung unterstützender Systeme im Gesundheitswesen und zeigt auf, wie durch den Einsatz moderner Technologien die Sicherheit und Effizienz im Medikamentenmanagement erheblich gesteigert werden können. Zukünftige Forschungen könnten auf den hier gewonnenen Erkenntnissen aufbauen, um das System weiter zu skalieren und auf einen breiten Einsatz im Pflegebereich vorzubereiten.

# A Anhang

## A.1 Freitext Rückmeldungen

Im folgenden sind alle Rückmeldungen auf den Fragebogen der Probanden (siehe Kapitel 7.5.3) aufgelistet. Dabei stehe die Pflegekräfte zu Beginn, gefolgt von den Laien.

Pflegekraft 1

- Das Interface ist zu groß

Pflegekraft 2

- Das System gibt Sicherheit

Pflegekraft 3

- Die Buchstabenmerkaufgabe ist ein realistischer Stresstest
- Das System hat Probleme die weißen Tabletten zu erkennen

Pflegekraft 4

- Sehr stressiger Test

Pflegekraft 5

- Imodium hat Loperamid als Wirkstoff, aber Loperamid ist auch ein separates Medikament

Probant 1

- Das Interface ist zu langsam
- Die farbliche Markierungen sollten unterschiedlich sein für zu viele oder zu wenig Medikamente

Probant 2

- Das System hat Schwierigkeiten bei der Unterscheidung von weißen Tabletten
- Es dauert zu lange beim Patientenwechsel, bis das Interface aktualisiert wird.

Probant 3

- 

Probant 4

- 

Probant 5

- Das Interface flackert.
- Eine Angabe über die Gesamtangabe pro Tageszeit wäre wünschenswert.
- Anstatt von Nullen sollten im Interface leere Felder angezeigt werden.

Probant 6

- Die Erkennung ist nicht stabil genug.
- Die Erkennung wechselt bei Unsicherheit sehr oft zwischen verschiedenen Angaben.



Probant 7

- Fotos der Medikamente im Interface neben den Namen würden helfen.
- Die Steuerung der Patientenwechsel muss verbessert werden.

Probant 8

- Bunte Pillen sind gut fürs System, aber wollen Menschen sie essen?
- Die Angaben auf dem Papier sind zu klein, aber das Interface des Systems ist zu groß
- Der Arbeitsbereich sollte näher sein
- Kann der Aufbau an jede Person anpassbar sein?

Probant 9

- Sehr übersichtlich
- Einfache Bestätigung

Probant 10

- Falsch negative Rückmeldungen bringen einen aus dem Konzept und die Fehlersuche irritiert.
- Bilder der Medikamente würden verwirren.

Probant 11

- Das System hilft, die parallelen Aufgaben besser zu handlen.
- Man muss weniger nachdenken.
- Man fühlt sich abgesichert.
- Farbliche Codierung hilft.

Probant 12

- 

Probant 13

- 

Probant 14

- System als hilfreich empfunden, insbesondere im Vergleich zu den Zetteln.

Probant 15

- Das System muss die Vorbereitung der ganzen Woche auf einmal unterstützen.

Probant 16

- Guter Eindruck
- Das Interface ist langsamer als der Wechsel des Papiers

Probant 17

- Das Interface ist langsam bei neuen Pillen.
- Horizontal liegende Pillen werden schlechter erkannt.

Probant 18

- Man gewöhnt sich an die Tabletten und erkennt sie dann.
- Die Matheaufgaben waren alle schnell und korrekt zu beantworten.
- Die Konzentration auf zwei Sachen ging gut mit dem System.
- positiver Gesamteindruck





Proband 19

- Übersichtlich und groß
- Keine Verwirrung
- Gute Überprüfung am Ende
- Die Erkennung könnte besser sein.



# Abbildungsverzeichnis

1.1	Die deutsche Bevölkerung in Tausend gruppiert in jeweils fünf Jahre pro Balken. A-D zeigt die Bevölkerungsverteilung in den Jahren 1957 (A), 1991 (B), 2021 (C) und die vorhergesagte Bevölkerungsverteilung im Jahr 2040 (D) [3]. . . . .	2
2.1	Übersicht über das Spektrum von Extended Reality . . . . .	7
2.2	Illustration eines künstlichen Neurons . . . . .	9
2.3	Die häufigsten Aktivierungsfunktionen $\sigma(a)$ . . . . .	11
2.4	Funktion einer Faltungsschicht: Jeder Filter wird schrittweise über den Eingang gezogen und erzeugt bei jedem Schritt einen Wert im Ausgang. . . . .	12
2.5	Aufbau der Querphasen-Teilschicht . . . . .	13
2.6	Die Schnittmenge über der Vereinigungsmenge von Modellausgabe (A) und Grundwahrheitswert (B) auf einem Trainingsbild. . . . .	16
2.7	VGG16-Architektur mit individuellen Schichten . . . . .	20
2.8	ResNet-Architektur mit vier Blöcken und deren Beipassverbindungen. . . . .	21
2.9	Dimensionales Clusters von Ankerrahmen mit gemeinsamen Zentrum. . . . .	23
2.10	YOLOv5 Aufbau . . . . .	25
2.11	Detailansicht der Struktur von Rückgrat, Verbund und Ausgabe von YOLOv5 . . . . .	26
3.1	Tablettendosetten mit Fächern für Morgens, Mittags, Abends und Nachts . . . . .	32
3.2	Auswahl an Verblistungsautomaten. Von links nach rechts: Omnicell VBM200F [21], JVM Primeblisters [22] und DB Rowa Smart [23] . . . . .	34

4.1	Zeitlicher Ablauf beim Medikamentenrichten mit dem Medikamentenmanagement-System . . . . .	48
4.2	Datenflussdiagramm des Medikamentenmanagement-Systems . . . . .	50
4.3	Abgrenzung zwischen Medikamentenmanagement-System und Einflussgrößen . . . . .	51
4.4	Der bundeseinheitlichen Medikationsplan mit den wichtigsten Informationen in Rot Herausgegeben von der Kassenärztlichen Bundesvereinigung[33]. . . . .	53
4.5	Übersicht der verwendeten Medikamente A - apoday Vitamin C (Vitamin C); B - Deslora - 1A (Desloratadin); C - Lorano akut (Loratadine); D - Ambroxol 75 retard Heumann (Ambroxol); E - Imodium akut (Loperamid); F - Angocin (Kräuterextrakt); G - Femibion 3 (Vitamine); H - Doppel Herz Omega-3 Konzentrat (Omega-3 Fettsäuren); I - Paracetamol 500 - 1A (Paracetamol); J - Baldrian - Ratiopharm (Baldrian Wurzelextrakt); K - Loperamid akut aristo (Loperamid) . . . . .	55
5.1	Konfusionsmatrix der Testdaten von YOLOv5 mit 11 Medikamenten (siehe Kapitel 4.4.2) und der Dosette (Box) mit 4 Fächern (MO, MI, AB, NA) . . . . .	62
5.2	Prototyp der Nutzerschnittstelle in der AR-Brille. In der Ansicht des Anwenders (A) und alleinstehend (B).  fixiert das Fenster an der aktuellen Position  schließt das Fenster und das Programm  und  wechseln zwischen den Patienten . . . . .	65
6.1	Hardwareaufbau mit Tischstativ, Kamera, Beamer und Einplatinenrechner . . . . .	72
6.2	Kameraaufbau über der Medikamentendosette und potentielle Abschattungen A Abschattung durch die Wände der Medikamentendosette B Abschattung durch die Hände des Nutzers . . . . .	73
6.3	Original (A) und Projektion (B) des Schachbrettmusters der Homographieberechnung. $P(x, y)$ ist ein Korrespondenzpunkt zu $P(x', y')$ . . . . .	74
6.4	KI-Kaskade von YOLOv5 und ResNet zur Effizienzerhöhung bei großer Artenzahl an Medikamenten . . . . .	77

6.5	Konfusionsmatrizen: Datensatz 1 (A) und Datensatz 2 (B) . . . . .	79
6.6	Konfusionsmatrizen von Datensatz 3 (C) . . . . .	80
6.7	t-SNE Visualisierung der Merkmalsrepräsentationen der Testdaten im trainierten ResNet zur Darstellung der Klassentrennung .	81
6.8	Konfusionsmatrix von YOLOv5 trainiert mit allen Medikamenten als eine Klasse. AB - Abends; MI - Mittags; MO - Morgens; NA - Nachts; Box - Medikamentendosette; Med - Medikamente . . . . .	82
6.9	Nutzerschnittstelle mit allen integrierten Informationen . . . . .	85
6.10	Komponenten-Diagramm mit farblicher Kodierung: lila für Schnittstellen, gelb für Verarbeitungseinheiten und orange für Netzwerkcommunication. Die dicken Pfeile geben den normalen Prozessablauf an. . . . .	87
7.1	6 Beispiele für Sequenzen aus einem OSPAN-Test . . . . .	90
7.2	Evaluationsergebnisse der Zwischenevaluation mit der Gesamtanzahl der einsortierten Medikamente (A), der Anzahl der Sortier- und Verwechslungsfehler (B) und den Fehlern im OSPA-Test (C) von maximal 30 . . . . .	94
7.3	Evaluationsergebnisse der professionellen Pflegekräfte mit der Gesamtanzahl der einsortierten Medikamente (A), der Anzahl der Sortier- und Verwechslungsfehler (B) und den Fehlern im OSPA-Test (C) von maximal 30 . . . . .	96
7.4	Evaluationsergebnisse der Laien mit der Gesamtanzahl der einsortierten Medikamente (A), der Anzahl der Sortier- und Verwechslungsfehler (B) und den Fehlern im OSPA-Test (C) von maximal 30 . . . . .	97
7.5	Evaluationsergebnisse der Laien und professionellen Pflegekräfte kombiniert mit der Gesamtanzahl der einsortierten Medikamente (A), der Anzahl der Sortier- und Verwechslungsfehler (B) und den Fehlern im OSPA-Test (C) von maximal 30 . . . . .	99
7.6	2-Stunden-Evaluationsergebnisse der Laien und professionellen Pflegekräfte kombiniert mit der Gesamtanzahl der einsortierten Medikamente (A) und der Anzahl der Sortier- und Verwechslungsfehler (B) . . . . .	100

7.7	Ergebnisse der Anforderungs- und Belastungsanalyse als Histogramme. $\emptyset$ und die schwarzen Markierungen geben die Mittelwerte aller Befragten auf der x-Achse an, die gelbe Markierung die Mittelwerte für Laien und lila für die Pflegekräfte. $w$ gibt die berechneten Gewichtungsfaktoren an. Das Endergebnis ist ein Gesamtwert von 67,93. . . . .	103
8.1	Die t-SNE Visualisierung der Merkmalsrepräsentationen des ResNet mit 0 Exemplaren (oben) und 5 Exemplaren (unten) der neuen Klasse (hellgrün). . . . .	106
8.2	Vergleich der Ergebnisse der ersten 10 Minuten mit den letzten 5 Minuten, wobei die Skala zum Vergleich normiert ist. . . . .	109
9.1	Pflegekraft beim Richten von Medikamenten mit dem vorgestellten System. Verwendung mit Zustimmung der abgebildeten Person. . . . .	116



# Tabellenverzeichnis

5.1	Vergleich verschiedener Lokalisierungsnetzwerke auf dem MS-COCO-Datensatz (test-dev 2017) [34]. Bildgröße: Dimensionen des Eingangsbildes in Pixel. $\varnothing PR_{50}$ : mittlere Genauigkeit bei einem Mindestwert der Schnittmenge über der Vereinigungsmenge von 50%. fps: Bearbeitungsgeschwindigkeit in Frames pro Sekunde. . . . .	58
5.2	Beschreibung von zwei Testdatensätzen zur Evaluation mit Medikamentenbildern . . . . .	59
5.3	Ergebnisse des Vergleichs von drei verschiedenen YOLOv5 Architekturen auf zwei Datensätzen (siehe Tabelle 5.2) . . . . .	60
5.4	Zeitdauer der einzelnen Verarbeitungsschritte, gerundet auf 2 signifikante Stellen . . . . .	68
6.1	Zusammensetzung der Trainingsdatensätze zur Bestimmung der optimalen Datenmenge . . . . .	78
7.1	System-Usability-Scale-Fragebogen mit den Punkten zu jeder Antwortmöglichkeit . . . . .	92
7.2	NASA-Task-Load-Index-Fragebogen . . . . .	93
7.3	Bewertungstabelle für den NASA-TLX [42] . . . . .	93
7.4	Ergebnisse des SUS-Fragebogens als Histogramme mit Mittelwerten. Negative Fragen sind bereits invertiert. $\varnothing$ und die schwarzen Markierungen geben die Durchschnittswerte aller Befragten an. Das Endergebnis ist ein SUS-Wert von 88,02. . . . .	101
8.1	Vorhersage des Unternehmensumsatzes in den ersten Jahren nach einer Ausgründung. Preis pro System: 10.000N. . . . .	114



# Literaturverzeichnis

- [1] EPRS, *Demographic Outlook for the European Union: PE 690.528*, Published: Veröffentlichung, 2021. Adresse: [https://www.europarl.europa.eu/thinktank/en/document/EPRS\\_STU\(2021\)690528](https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2021)690528).
- [2] H. G. Hockerts, *Sozialpolitische Entscheidungen im Nachkriegsdeutschland: Alliierte u. deutsche Sozialversicherungspolitik 1945 bis 1957* (Forschungen und Quellen zur Zeitgeschichte). Stuttgart: Klett-Cotta, 1980, Bd. 1, ISBN: 978-3-12-912910-4.
- [3] DESTATIS, *Bevölkerung in Deutschland*, Published: Website, 2023. Adresse: <https://service.destatis.de/bevoelkerungspyramide/index.html>.
- [4] H. Dormann, *Medikamenten-Nebenwirkungen: krank statt geheilt*, Plusminus, 8. Feb. 2017.
- [5] S. Erzkamp und O. Rose, „Medikamentenmanagement oder Medikationsmanagement? Was bringt die Neuverblisterung für die Arzneimitteltherapiesicherheit in Alten- und Pflegeheimen?“, *Deutsche Apotheker Zeitung*, 2016. Adresse: <https://www.deutsche-apotheker-zeitung.de/daz-az/2016/daz-17-2016/medikamentenmanagement-oder-medikationsmanagement> (besucht am 16.07.2024).
- [6] Y. Lecun, L. Bottou, Y. Bengio und P. Haffner, „Gradient-based learning applied to document recognition“, *Proceedings of the IEEE*, Jg. 86, Nr. 11, S. 2278–2324, 1998, ISSN: 00189219. DOI: 10.1109/5.726791.
- [7] A. Krizhevsky, I. Sutskever und G. E. Hinton, „ImageNet Classification with Deep Convolutional Neural Networks“, in *Advances in Neural Information Processing Systems*, F. Pereira, C.J. Burges, L. Bottou und K.Q.

- Weinberger, Hrsg., Bd. 25, Curran Associates, Inc, 2012. Adresse: [https://proceedings.neurips.cc/paper\\_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf).
- [8] A. Krizhevsky, I. Sutskever und G. E. Hinton, „ImageNet classification with deep convolutional neural networks“, *Communications of the ACM*, Jg. 60, Nr. 6, S. 84–90, 24. Mai 2017, ISSN: 0001-0782, 1557-7317. DOI: 10.1145/3065386. Adresse: <https://dl.acm.org/doi/10.1145/3065386> (besucht am 06.06.2024).
- [9] K. He, X. Zhang, S. Ren und J. Sun, *Deep Residual Learning for Image Recognition*, 10. Dez. 2015. arXiv: 1512.03385[cs]. Adresse: <http://arxiv.org/abs/1512.03385> (besucht am 25.10.2024).
- [10] C. Szegedy, W. Liu, Y. Jia u. a., *Going Deeper with Convolutions*, 17. Sep. 2014. Adresse: <http://arxiv.org/pdf/1409.4842v1>.
- [11] M. Tan, R. Pang und Q. V. Le, *EfficientDet: Scalable and Efficient Object Detection*, 27. Juli 2020. arXiv: 1911.09070[cs, eess]. Adresse: <http://arxiv.org/abs/1911.09070> (besucht am 24.05.2024).
- [12] J. Redmon, S. Divvala, R. Girshick und A. Farhadi, *You Only Look Once: Unified, Real-Time Object Detection*, 6. Aug. 2015. Adresse: <http://arxiv.org/pdf/1506.02640v5>.
- [13] J. Redmon und A. Farhadi, *YOLO9000: Better, Faster, Stronger*, 25. Dez. 2016. Adresse: <http://arxiv.org/pdf/1612.08242v1>.
- [14] J. Redmon und A. Farhadi, *YOLOv3: An Incremental Improvement*, 4. Sep. 2018. Adresse: <http://arxiv.org/pdf/1804.02767v1>.
- [15] A. Bochkovskiy, C.-Y. Wang und H.-Y. M. Liao, *YOLOv4: Optimal speed and accuracy of object detection*, 22. Apr. 2020. arXiv: 2004.10934[cs, eess]. Adresse: <http://arxiv.org/abs/2004.10934> (besucht am 06.06.2024).
- [16] W. Liu, D. Anguelov, D. Erhan u. a., „SSD: Single Shot MultiBox Detector“, in Bd. 9905, 2016, S. 21–37. DOI: 10.1007/978-3-319-46448-0\_2. arXiv: 1512.02325[cs]. Adresse: <http://arxiv.org/abs/1512.02325> (besucht am 24.05.2024).

- [17] J. Sweller, „Cognitive Load Theory“, in Ser. Psychology of Learning and Motivation, Academic Press, Hrsg., Bd. 55, Elsevier, 2011, S. 37–76, ISBN: 978-0-12-387691-1. DOI: 10.1016/B978-0-12-387691-1.00002-8.
- [18] D. Kahneman, *Schnelles Denken, langsames Denken*, 25. Auflage. München: Siedler, 2019, ISBN: 3-88680-886-6.
- [19] AMG Arzneimittelgesetz, *Gesetz über den Verkehr mit Arzneimitteln*, 1976. Adresse: [https://www.gesetze-im-internet.de/amg\\_1976/](https://www.gesetze-im-internet.de/amg_1976/).
- [20] N. Menche, C. Keller und B. Teigeler, Hrsg., *PflegeHeute: Lehrbuch für Pflegeberufe*, 8. Auflage, München: Elsevier, 2023, 1602 S., ISBN: 978-3-437-26779-6.
- [21] Omnicell, *Omnicell Produkte*, 29. Aug. 2024. Adresse: <https://www.omnicell.de/produkte/blisterkartenautomat-vbm-200f> (besucht am 29.08.2024).
- [22] PrimeBlister, *PrimeBlister Produkte*, 29. Aug. 2024. Adresse: <https://www.primeblister.eu/index.php/blisterpack98/> (besucht am 29.08.2024).
- [23] D. Rowa, *DB Rowa Produkte*, 29. Aug. 2024. Adresse: <https://rowa.de/de/page/automatisierung-zum-einstiegspreis/> (besucht am 29.08.2024).
- [24] L. Tan, T. Huangfu, L. Wu und W. Chen, „Comparison of RetinaNet, SSD, and YOLO v3 for real-time pill identification“, *BMC Medical Informatics and Decision Making*, Jg. 21, Nr. 1, S. 324, Dez. 2021, ISSN: 1472-6947. DOI: 10.1186/s12911-021-01691-8. Adresse: <https://bmcmidinformedecismak.biomedcentral.com/articles/10.1186/s12911-021-01691-8> (besucht am 02.08.2024).
- [25] A. Dhivya und M. Sundaresan, „Tablet identification using support vector machine based text recognition and error correction by enhanced  $n$ -grams algorithm“, *IET Image Processing*, Jg. 14, Nr. 7, S. 1366–1372, Mai 2020, ISSN: 1751-9667, 1751-9667. DOI: 10.1049/iet-ipr.2019.0993. Adresse: <https://onlinelibrary.wiley.com/doi/10.1049/iet-ipr.2019.0993> (besucht am 05.09.2024).

- [26] E. Elyan, P. Vuttipittayamongkol, P. Johnston u. a., „Computer vision and machine learning for medical image analysis: recent advances, challenges, and way forward“, *Artificial Intelligence Surgery*, 2022, ISSN: 27710408. DOI: 10.20517/ais.2021.15. Adresse: <https://www.oaepublish.com/articles/ais.2021.15> (besucht am 24. 10. 2024).
- [27] T. T. Nguyen, P. Le Nguyen, Y. Kawanishi, T. Komamizu und I. Ide, „Zero-Shot Pill-Prescription Matching With Graph Convolutional Network and Contrastive Learning“, *IEEE Access*, Jg. 12, S. 55 889–55 904, 2024, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2024.3390153. Adresse: <https://ieeexplore.ieee.org/document/10504270/> (besucht am 24. 10. 2024).
- [28] C. Gansohr, M. B. Thelen, S. Geisler und S. C. Eimler, „Supporting Hospital Nurses During Medication Dispensation with Augmented Reality – A Participatory Approach“, in *HCI International 2022 Posters: 24th International Conference on Human-Computer Interaction, HCII 2022, Virtual Event, June 26 – July 1, 2022, Proceedings, Part IV*, Ser. Springer eBook Collection, C. Stephanidis, M. Antona und S. Ntoa, Hrsg., Bd. 1583, Cham: Springer International Publishing und Imprint Springer, 2022, S. 35–42, ISBN: 978-3-031-06393-0. DOI: 10.1007/978-3-031-06394-7\backslashbackslashtextunderscore.
- [29] S. Aicher, K. Klinker, M. Wiesche und H. Krcmar, „Augmented Reality im Gesundheitswesen – Entwurf und Auswertung einer HoloLens-Anwendung zur Verteilung von Medikamenten“, in *Systematische Entwicklung von Dienstleistungsinnovationen: Augmented Reality für Pflege und industrielle Wartung*, Ser. Research, M. Wiesche, I. M. Welp, H. Remmers und H. Krcmar, Hrsg., Wiesbaden, Germany und Heidelberg: Springer Gabler, 2021, S. 287–306, ISBN: 978-3-658-31767-6. DOI: 10.1007/978-3-658-31768-3\backslashbackslashtextunderscore.
- [30] W.-J. Chang, L.-B. Chen, C.-H. Hsu, J.-H. Chen, T.-C. Yang und C.-P. Lin, „MedGlasses: A Wearable Smart-Glasses-Based Drug Pill Recognition System Using Deep Learning for Visually Impaired Chronic Patients“,

- IEEE Access*, Jg. 8, S. 17 013–17 024, 2020. DOI: 10 . 1109 / ACCESS . 2020 . 2967400.
- [31] Statista, *Anzahl zugelassener Arzneimittel in Deutschland nach Verschreibungs-/Abgabestatus im Jahr 2023*. Adresse: <https://de.statista.com/statistik/daten/studie/513971/umfrage/anzahl-zugelassener-arzneimittel-in-deutschland-nach-verschreibungs-abgabestatus/> (besucht am 12.09.2024).
- [32] Sozialgesetzbuch Fünf, *Gesetzliche Krankenversicherung: SGB V*, 3. Juni 2021.
- [33] Kassenärztliche Bundesvereinigung, *Medikationsplan*, 26. Feb. 2024. Adresse: [https://www.kbv.de/media/sp/Medikationsplan\\_Beispiel\\_2023.pdf](https://www.kbv.de/media/sp/Medikationsplan_Beispiel_2023.pdf).
- [34] T.-Y. Lin, M. Maire, S. Belongie u. a., *Microsoft COCO: Common Objects in Context*, 5. Jan. 2014. Adresse: <http://arxiv.org/pdf/1405.0312v3>.
- [35] G. Bradski, „The OpenCV library“, *Dr. Dobbs’s Journal of Software Tools*, 2000, tex.citeulike-article-id: 2236121 tex.posted-at: 2008-01-15 19:21:54 tex.priority: 4.
- [36] O. Russakovsky, J. Deng, H. Su u. a., „ImageNet large scale visual recognition challenge“, *International Journal of Computer Vision*, Jg. 115, Nr. 3, S. 211–252, Dez. 2015, ISSN: 0920-5691, 1573-1405. DOI: 10 . 1007 / s11263 - 015 - 0816 - y. Adresse: <http://link.springer.com/10.1007/s11263-015-0816-y> (besucht am 22.10.2024).
- [37] B. Geisendrees, „Methoden zur Erkennung von Hautfarbe in Farbbildern“, Sep. 2001.
- [38] M. L. Turner und R. W. Engle, „Is working memory capacity task dependent?“, *Journal of Memory and Language*, Jg. 28, Nr. 2, S. 127–154, 1989, ISSN: 0749596X. DOI: 10 . 1016 / 0749 - 596X (89) 90040 - 5.

- [39] J. Brooke, „SUS: A 'Quick and Dirty' Usability Scale“, in *Usability Evaluation In Industry*, CRC Press, Hrsg., CRC Press, 1996, S. 207–212, ISBN: 978-0-429-15701-1. DOI: 10.1201/9781498710411-35. Adresse: <https://www.taylorfrancis.com/chapters/edit/10.1201/9781498710411-35/sus-quick-dirty-usability-scale-john-brooke>.
- [40] J. Brooke, „SUS: a retrospective“, *Journal of Usability Studies*, Jg. 8, S. 29–40, 2013.
- [41] S. G. Hart, *NASA Task Load Index (TLX)*, 1. Jan. 1986. Adresse: <https://ntrs.nasa.gov/citations/20000021487>.
- [42] S. G. Hart und L. E. Staveland, „Development of NASA-TLX (task load index): Results of empirical and theoretical research“, in *Advances in Psychology*, Bd. 52, Elsevier, 1988, S. 139–183, ISBN: 978-0-444-70388-0. DOI: 10.1016/S0166-4115(08)62386-9. Adresse: <https://linkinghub.elsevier.com/retrieve/pii/S0166411508623869> (besucht am 11.06.2024).
- [43] BMAS und BMG, „Höhere Mindestlöhne in der Altenpflege“, Pressemitteilung, 29. Aug. 2023. Adresse: <https://www.bmas.de/DE/Service/Presse/Pressemitteilungen/2023/empfehlung-anhebung-der-mindestloehne-altenpflege.html> (besucht am 16.07.2024).



## Eigene Veröffentlichungen

- [1] J. Juhl, K. F. Rudolf, L. Kohout u. a., „Using augmented reality and artificial intelligence for an efficient and safe preparation of individual drug assortments in nursing homes“, in *2022 international conference on electrical, computer, communications and mechatronics engineering (ICECCME)*, IEEE, 2022, S. 1–6.
- [2] J. Juhl, T. Rudolf, L. Kohout u. a., „Automated Anomaly Detection in Data Streams on Engine Test Benches Using Multi-Scale Convolutional Recurrent Encoder-Decoder“, *Proceedings of the International Vienna Motor Symposium 2021*. Hrsg.: B. Geringer, 2021.
- [3] G. Biri, A. Vasilache, Tao Hu u. a., „Sleep Stage and Apnea Classification from Single-Lead ECG Using Artificial and Spiking Neural Networks“, in *2024 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, in press, 2024.
- [4] L. Schweickart, C. Erler, J. Juhl, C. Zimmermann und W. Stork, „User-centered representation of data flows in mHealth applications“, in *2023 international conference on electrical, communication and computer engineering (ICECCE)*, IEEE, 2023, S. 1–6.