

Kontaktlose Messung von Respirationsparametern mittels Künstlicher Intelligenz

Von der Fakultät für Ingenieurwissenschaften,
Abteilung Elektrotechnik und Informationstechnik
der Universität Duisburg-Essen

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften

genehmigte Dissertation

von

Felix Wichum

aus

Erfurt

Datum der Einreichung: 13.09.2023

Gutachter: Prof. Dr.-Ing. Karsten Seidl

Gutachter: Prof. Dr.-Ing. habil. Sebastian Zaunseder

Tag der mündlichen Prüfung: 19.04.2024

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

ub | universitäts
bibliothek

Diese Dissertation wird via DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt und liegt auch als Print-Version vor.

DOI: 10.17185/duepublico/82089

URN: urn:nbn:de:hbz:465-20240619-153227-4

Alle Rechte vorbehalten.

Danksagungen

Ich möchte mich herzlich bei meinem Doktorvater, Prof. Dr.-Ing. Karsten Seidl, sowie meinem direkten Betreuer Dr.-Ing. Christian Wiede für die Betreuung dieser Arbeit bedanken.

Ein besonderer Dank gebührt auch Prof. habil. Dr.-Ing. Sebastian Zaunseder für seine Rolle als Zweitgutachter sowie den Mitgliedern der Prüfungskommission, Prof. Dr. sc. techn. Erni, Prof. Dr.-Ing. Balzer und Prof. Dr. Weimann, für ihre anregenden Diskussionen, faire Kritiken und spannenden Anknüpfungspunkten.

Mein Dank gilt ebenso dem Fraunhofer IMS für die bereitgestellten Ressourcen und Mittel, die die Durchführung dieser Arbeit ermöglicht haben.

Für die medizinische Expertise und den unerschütterlichen Glauben an das Projekt danke ich Dr. med. Sebastian Sohrab und seinem Praxisteam um Jasmin von der Neudorfer Lungenpraxis. Weiterhin danke ich allen Probanden und Patienten, die an meiner Studie teilgenommen haben.

Ich danke Burkhard, Conny und Wolfgang für das eingeräumte Vertrauen und die gewährten Freiräume. Meinen Kollegen danke ich für zielführende Gespräche, aufmunternde Worte und abwechslungsreiche Ideenfindungen. Ein besonderer Dank geht an Carolin, Belmin, Dennis, Gabriel, Gongbo, Hüseyin, Johannes, Lars, Lukas, Maarten, Niels und Svenja.

Ich möchte mich auch bei den Studierenden bedanken, die ich auf ihrem akademischen Weg begleiten durfte und die ihren Teil zu dieser Arbeit beigetragen haben: Laura, Corinna, Hüseyin, Jacqueline, Emine, Julian, Navraj, Zhouyi und Smit.

Dank gebührt meinen Freunden für Ablenkung, Trost und Unterhaltung bei den gemeinsamen Abenteuern in fantasievollen Welten, bei schlechten Filmen oder einfach so.

Herzlichen Dank an meine Eltern, die unermüdlich an mich geglaubt und mich unterstützt haben, weit vor Beginn dieser Arbeit.

Vielen Dank auch an Dumbi und Luzi, für die eingeforderten Ablenkungen und aufgezwungenen Pausenzeiten.

Und nicht zuletzt vielen Dank an Evi, für ihre Liebe und Unterstützung. Ihr feines Auge für Details und die ermutigenden Worte waren für meine Arbeit unverzichtbar.

Danke.

Abstract

Depth-based Plethysmography (DPG) is an optical method for measuring the displacement of a subject's upper body from a depth camera. This potentially mobile and cost-effective approach allows visualizing the mechanics and contributions to respiration during natural breathing, enabling the inference of volume changes in the thorax and, consequently, respiratory volumes.

This dissertation explores the targeted use of Artificial Intelligence in DPG, focusing on three areas: the application of Machine Learning for regressive determination of respiratory volumes from extracted signal features, the development of a physiological Region of Interest for improved signal extraction, and the targeted application of DPG for the classification of respiratory diseases.

Using an existing dataset from the literature, it is shown that tidal volume can be determined with an accuracy of 11 ± 253 ml and vital capacity with 11 ± 624 ml, using Support Vector Regression and a Mutual Information-based feature selection from 379 extracted features. The achieved accuracy surpasses comparable models from the literature and further allows calibration-free measurements.

Another newly developed method utilizes the statistical relationships of the signals of individual pixels with the spirometer signal within a weighted, individual patient mask. The representation of a generalized, averaged mask for all subjects is based on detected body joints. The resulting improved signal extraction enables even more precise measurements of tidal volume (8 ± 206 ml) and vital capacity (-43 ± 599 ml). An exploratory analysis indicates the high potential benefits of individual masks in supporting diagnosis and therapy for respiratory diseases.

For the investigation of respiratory disease classification, a dedicated DPG dataset of asthmatics ($N = 26$), patients with chronic obstructive lung disease (18), and healthy subjects (9) was recorded. The developed classification models

using raw data, features, extracted signals, and masks demonstrate that the individual masks yield the best results. An accuracy of 57,7% and Cohen's Kappa of 0,28 is achieved. The individual masks were developed using correlation-based methods without the need for a reference spirometer signal.

In summary, this work demonstrates the feasibility of leveraging Artificial Intelligence in various aspects of DPG. Future developments on mobile devices can benefit from this and offer a portable and cost-effective DPG system.

Zusammenfassung

Die Tiefenplethysmographie (DPG) ist ein optisches Verfahren zur Messung der Abstandsänderung des Oberkörpers einer Testperson zu einer Tiefenkamera. Dieses potenziell mobile und kostengünstige Verfahren ermöglicht es, die Mechanik und die Beiträge bei natürlicher Atmung sichtbar zu machen, daraus auf die Volumenänderungen im Thorax und somit auf die Respirationsvolumina zu schließen.

Die vorliegende Dissertation untersucht den gezielten Einsatz von Künstlicher Intelligenz bei der DPG in drei Bereichen: Erstens die Anwendung von Maschinellem Lernen zur regressiven Bestimmung der Atemvolumina aus extrahierten Signalmerkmalen, zweitens die Entwicklung einer physiologischen Region of Interest zur verbesserten Signalextraktion und drittens die gezielte Anwendung der DPG für die Klassifikation von respiratorischen Erkrankungen.

Anhand eines bestehenden Datensatzes aus der Literatur wird gezeigt, dass das Tidalvolumen mit einer Genauigkeit von 11 ± 253 ml und die Vitalkapazität mit 11 ± 624 ml durch die Verwendung von Support Vector Regression und einer auf Mutual Information basierenden Merkmalsselektion aus 379 extrahierten Merkmalen bestimmt werden kann. Die erzielte Genauigkeit übertrifft vergleichbare Modelle aus der Literatur und ermöglicht zudem eine kalibrationsfreie Messung.

Eine weitere, neu entwickelte Methode nutzt die statistischen Zusammenhänge der Signale der einzelnen Pixel mit dem Spirometersignal in einer gewichteten, individuellen Patientenmaske. Die Darstellung einer verallgemeinerten, gemittelten Maske für alle Probanden erfolgt anhand der detektierten Körpergelenkspunkte. Die daraus resultierende, verbesserte Signalextraktion ermöglicht eine noch genauere Messung von Tidalvolumen (8 ± 206 ml) und Vitalkapazität (-43 ± 599 ml). Eine explorative Analyse zeigt einen hohen potentiellen Nutzen der individuellen Masken zur Unterstützung von Diagnose und Therapie bei respiratorischen Erkrankungen.

Für die Untersuchung der Klassifikation respiratorischer Erkrankungen wurde ein eigener DPG-Datensatz von Asthmatikern ($N = 26$), Patienten mit chronisch-obstruktiver Lungenerkrankung (18) und Gesunden (9) aufgezeichnet. Die entwickelten Modelle aus Rohdaten, Merkmalen, extrahierten Signalen und Masken zeigen, dass die individuellen Masken die besten Klassifikationsergebnisse ermöglichen. Eine Klassifikationsgenauigkeit von 57,7% und Cohens Kappa von 0,28 wird erreicht. Die individuellen Masken wurden durch korrelationsbasierte Methoden entwickelt, ohne den Einsatz eines Referenz-Spirometersignals.

Zusammenfassend zeigt diese Arbeit, dass der Einsatz von Künstlicher Intelligenz in den verschiedenen Bereichen der DPG möglich ist. Zukünftige Entwicklungen auf Mobilgeräten können davon profitieren und eine portable und kostengünstige DPG anbieten.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	1
1.2	Arbeitsumfang (Scope of Work)	4
1.3	Aufbau dieser Arbeit	5
2	Grundlagen	7
2.1	Anatomie und Physiologie der menschlichen Atmung	7
2.1.1	Aufbau des respiratorischen Systems	8
2.1.2	Funktionsweise der Atemmechanik	11
2.2	Messverfahren und Parameter der Respiration	12
2.2.1	Messparameter	13
2.2.2	Veränderungen in der Respiration	15
2.2.3	Klinische Goldstandards	18
2.2.4	Tiefenplethysmographie	19
2.3	Tiefenkameras	22
2.4	Methoden der Künstlichen Intelligenz	24
2.4.1	Merkmalsselektion	24
2.4.2	Modelle des Maschinellen Lernens	25
3	Stand der Forschung	29
3.1	Auswahl einer Region of Interest	29
3.1.1	Auswahlverfahren	30
3.1.2	Ausgewählte Regionen	33
3.2	Modelle zur Volumen-Zeit-Bestimmung	35
3.3	Genauigkeit und Evaluierung der Respirationsparametermessung	37
3.4	Klassifikation von Respiratorischen Erkrankungen	41

3.5	Zusammenfassung und Bewertung des Stands der Forschung . . .	42
3.6	Abgeleitete Arbeitshypothesen für diese Arbeit	43
4	Maschinelles Lernen zur Bestimmung von Atemvolumina	45
4.1	Methodik	46
4.1.1	ROI, Signalextraktion und Signalvorverarbeitung	47
4.1.2	Merkmalsextraktion	48
4.1.3	Merkmalsselektion	54
4.1.4	Regressionsmodelle	56
4.1.5	Evaluierungsstrategie	59
4.2	Datensatz	61
4.3	Ergebnisse und Diskussion	63
4.3.1	Vergleich der Methoden des Maschinellen Lernens	63
4.3.2	Relevante Merkmale für die Regression	68
4.3.3	Messgenauigkeit im Vergleich zur Literatur	72
4.4	Zusammenfassung	74
5	Verbesserung der Signalextraktion	77
5.1	Methodik	79
5.1.1	Untersuchung der Region of Interest	80
5.1.2	Long-Short-Term-Memory-Modelle	84
5.1.3	Merkmalsbasierte Modell	86
5.1.4	Evaluierungsstrategie	87
5.2	Ergebnisse und Diskussion	89
5.2.1	Patientenmasken	90
5.2.2	Auswirkungen auf die Modelle des Maschinellen Lernens	97
5.3	Zusammenfassung	100
6	Klassifikation von respiratorischen Erkrankungen	103
6.1	Methodik	104
6.1.1	Datensatzgenerierung	106
6.1.2	Referenzmodell	108
6.1.3	Merkmalsbasierte Klassifikation	109
6.1.4	Signalbasierte Klassifikation	111
6.1.5	Klassifikation anhand der physiologischen Maske	114

6.1.6	End-to-End Maschine Learning Modell Ansatz	118
6.2	Ergebnisse und Diskussion	120
6.2.1	Aufbau des Datensatzes	120
6.2.2	Klassifikationsgenauigkeiten	121
6.2.3	Diskussion der Klassifikationsgenauigkeiten	125
6.2.4	Korrelationsbasierte Patientenmasken	131
6.3	Zusammenfassung	135
7	Zusammenfassung und Ausblick	139
7.1	Zusammenfassung der Forschungsergebnisse	139
7.2	Wissenschaftliche Schlussfolgerungen und Contributions	142
7.3	Ausblick und zukünftige Forschungsfelder	145
A	Berechnungen	147
A.1	Extrahierte Merkmale	147
B	Zur Person	151
B.1	Veröffentlichungen	151
	Literatur	155

Abbildungsverzeichnis

2.1	Atemwege des respiratorischen Systems	8
2.2	Thoraxanatomie	10
2.3	Statische Lungenvolumina	14
2.4	Ablauf Tiefenplethysmographie	21
2.5	Kameraprinzipien für Tiefenmessungen	23
3.1	Auswahl einer ROI durch Amplituden	31
3.2	Auswahl einer Region-of-Interest durch Level-Set-Segmentierung .	32
3.3	Auswahl von Subregionen in der Literatur	34
3.4	Möglichkeiten zur Kalibrierung	38
4.1	Vorgehensweise bei der merkmalsbasierten Bestimmung von Re- spirationsvolumina	47
4.2	Festlegung der Region of Interest	48
4.3	Aufteilung des Messsignals	50
4.4	Geschachtelte Kreuzvalidierung zur Evaluierung	60
4.5	Verteilung der Grundwahrheit	62
4.6	Bland-Altman-Darstellung zur Abweichung der Messverfahren . .	67
4.7	Relevanz der Merkmale	69
5.1	Vorgehen bei der Erstellung von individuellen Patientenmasken .	79
5.2	Vorgehen bei der Entwicklung einer generalisierten Maske	83
5.3	Signalunterteilung für rekurrente Neuronale Netze	85
5.4	Validierung der Signalextraktion und ML-Modelle	87
5.5	Beispielhafte Maske mittels Mutual Information	91
5.6	Patientenmasken bei thorakale Atmung	92
5.7	Patientenmasken bei abdominaler Atmung	93
5.8	Patientenmasken mit ausgeprägter Kante	94

5.9	Patientenmasken bei faltenschlagender Kleidung	94
5.10	Verallgemeinerte Patientenmaske	96
6.1	Vorgehen zur Klassifikation	105
6.2	Messaufbau in der Lungenarztpraxis	107
6.3	Messabschnitt des Messsignals	113
6.4	ROI der Maske	117
6.5	Konfusionsmatrizen	124
6.6	Summierte Korrelationsmaske	132
6.7	Korrelationsmasken Brustkante	133
6.8	Korrelationsmasken Kleidungsfalten	133
6.9	Korrelationsmasken Segmentierungslücken	134
6.10	Korrelationsmasken Isolinien	134

Tabellenverzeichnis

2.1	Vergleich restriktiver und obstruktiver Ventilationsstörungen . . .	17
3.1	Messgenauigkeiten und Kalibrierungsstrategien in der Literatur .	39
4.1	Hyperparameter der Maschine Learning Modelle	58
4.2	Modellgenauigkeiten mit persönlichen Merkmalen	64
4.3	Modellgenauigkeiten ohne persönliche Merkmale	71
4.4	Vergleich der entwickelten Modelle mit der Literatur	72
5.1	Hyperparameter des Long Short-Term Memory Modells	86
5.2	Support Vector Regression mit generalisierten Masken	99
5.3	Auswirkungen auf die Modelle durch generalisierte Masken. . . .	100
6.1	Hyperparameter der Klassifikation für das Referenzmodell.	109
6.2	Hyperparameter der Signal-Klassifikation.	114
6.3	Hyperparameter der Masken-Klassifikation.	117
6.4	Hyperparameter der End-to-End-Klassifikation.	119
6.5	Beschreibung des Datensatz.	120
6.6	Klassifikationsgenauigkeiten der Modelle.	122
A.1	Übersicht der Merkmalsextraktion	147

Abkürzungsverzeichnis und Symbole

Abkürzungsverzeichnis

BMI	Körpermasseindex
CNN	neuronales Faltungsnetzwerk
COPD	chronisch obstruktive Lungenerkrankung
DPG	Tiefenplethysmographie
DTW	dynamische Zeitnormierung
EL	Ensemble-Lernen
EMD	empirische Zerlegung in Moden
FR	F-Regression
KI	Künstliche Intelligenz
KNN	Künstliche Neuronale Netze
LOOCV	Kreuzvalidierung mit Auslassung
LOPCV	Kreuzvalidierung mit Patientenauslassung
LSTM	langes Kurzzeitgedächtnis
MARM	manuelle Beurteilung der Atembewegung
MB	modellbasierte Merkmalsselektion
MI	Transinformation
ML	Maschinelles Lernen
OEP	opto-elektronische Plethysmographie
PCA	Hauptkomponentenanalyse
RF	Random Forest
RFE	Rekursive Merkmalselimination

RMSE	mittlere quadratische Fehler
ROI	Bereich von Interesse
SC	Stereoskopie
SF	Skalierungsfaktor
SL	Strukturiertes Licht
SNR	Signal-Rausch-Verhältnis
sROI	Teilbereich von Interesse
SVM	Support Vector Machine
SVR	Support Vector Regression
TOF	Laufzeitverfahren
TSFEL	Time Series Feature Extraction Library
ULoA	Untergrenze der Zustimmung
ULoA	Obergrenze der Zustimmung

Symbole

Symbol	Bezeichnung	Einheit
AF	Atemfrequenz	min^{-1}
AMV	Atemminutenvolumen	l min^{-1}
FEV ₁ /FVC	Tiffeneau-Index	%
PEF	Maximale Atemstromstärke bei forcierter Expiration	l s^{-1}
MEF ₂₅	mittlere Atemstromstärke bei 25 % der VC	l s^{-1}
MEF ₅₀	mittlere Atemstromstärke bei 50 % der VC	l s^{-1}
MEF ₇₅	mittlere Atemstromstärke bei 75 % der VC	l s^{-1}
sG _{Aw,eff}	spezifische Atemwegsleitfähigkeit effektiv	$\text{kPa}^{-1} \text{s}$
R _{Aw,eff}	Atemwegswiderstand effektiv	kPa s l^{-1}
sR _{Aw,eff}	spezifischer Atemwegswiderstand effektiv	kPa s
sR _{Aw,tot}	spezifischer Atemwegswiderstand total	kPa s
FEV ₁	Einsekundenkapazität	l
TGV	thorakales Gasvolumen	l
IC	Inspirationskapazität	l
s	Messsignal der tiefenbasierten Messung	mm
ERV	expiratorisches Reservevolumen	l
IRV	inspiratorisches Reservevolumen	l
FRC	funktionelle Residualkapazität	l
RV	Residualvolumen	l
α	Skalierungsfaktor	l mm^{-1}
TV	Tidalvolumen	l
TC	Totalkapazität	l
VC	Vitalkapazität	l
VC _{in}	inspiratorische Vitalkapazität	l
VC _{ex}	expiratorische Vitalkapazität	l
FVC	forcierte Vitalkapazität	l
SVC	langsame Vitalkapazität	l
V	Volumen	l

Kapitel 1

Einleitung

1.1 Motivation

Allein in Deutschland sind Erkrankungen des Atmungssystems für 20,2% aller Ursachen für Arbeitsunfähigkeit und 10,1% der Krankheitstage im Jahr 2021 zuständig [1] und belasten somit das Gesundheitssystem schwer. Weltweit litten 2017 544,9 Millionen Menschen an chronischen Atemwegserkrankungen. 3,9 Millionen Menschen starben in Folge einer solchen Erkrankung [2]. Die Todesopfer durch die noch immer anhaltende SARS-Cov-2 Pandemie werden auf 6,9 Millionen geschätzt (Stand Frühjahr 2023) [3]. Somit gehören chronische Atemwegserkrankungen zu den häufigsten Todesursachen weltweit. Die Gründe für die Letalität sind in der Lebensnotwendigkeit der Atmung begründet.

Die Atmung, auch Respiration genannt, übernimmt elementare Vorgänge im menschlichen Körper. Durch die Anhebung der Brustkorbs wächst das Volumen im Oberkörper. Der entstehende Unterdruck lässt Umgebungsluft in die Lungen strömen. Der Gasaustausch in den Lungen führt dazu, dass Sauerstoff aus der Luft in das Blut und von dort in alle Zellen des Körpers gelangt. Kohlenstoffdioxid als Stoffwechselprodukt wird über die Atmung aus dem Körper geführt. Durch den erzeugten Luftstrom übernimmt die Atmung auch eine wichtige Aufgabe bei der Tonerzeugung durch die schwingenden Stimmbänder. Ebenso ist die vorbeiziehende Luft eine Voraussetzung für die Stimulation des Riechepithels in den Nasenhöhlen. Die Atmung ist somit an der Sauerstoff-Versorgung, dem Abtransport von Kohlenstoffdioxid sowie beim Sprechen und Riechen beteiligt [4].

Störungen des respiratorischen Systems können demnach verheerende Auswirkungen auf die genannten Funktionen haben. Obstruktive Erkrankungen wie zum Beispiel die chronisch obstruktive Lungenerkrankung (COPD, chronic obstructive pulmonary disease) oder Asthma bronchiale führen zu einer Verengung der Atemwege und somit zu einem erhöhten Atemwegswiderstand [5]. Restriktive Erkrankungen führen zu einer verringerten Ausdehnbarkeit der Lunge, wie zum Beispiel im Krankheitsbild der Lungenfibrose [6].

Bei einer frühzeitige Diagnose kann jedoch eine entsprechend rechtzeitig beginnende Therapie helfen, eine klinische Verschlechterung zu vermeiden [7]. Für eine solche Diagnose werden Lungenfunktionstests mit Spirometern und Bodyplethysmographen, dem klinischen Goldstandard, bei Patienten durchgeführt. Bei der Spirometrie atmet der Patient durch ein Mundstück in ein Spirometer, welches den Volumenstrom misst und somit auf Respirationsparameter schließen kann [8, 9]. Bei der Bodyplethysmographie sitzt der Patient in einer geschlossenen Glaskammer. Der Begriff Plethysmographie beschreibt ein Messverfahren zur Bestimmung von Volumenschwankungen eines Körperteils oder Organs. In diesem Fall ist der Druck oder die Volumenänderungen in der Kammer proportional zum Atemvolumen der Patienten [10]. Abweichungen der gemessenen Parameter von den Normwerten weisen auf eine vorliegende Erkrankung hin und müssen in der Differenzialdiagnostik weiter untersucht werden.

Diese kontaktbasierten Messverfahren besitzen jedoch entscheidende Nachteile, welche durch kamerabasierte Ansätze gelöst werden können. Weder Spirometrie noch Bodyplethysmographie können einzelne Beiträge der Atemmechanik aus verschiedenen Thoraxregionen differenzieren - gemessen wird lediglich eine Volumenänderung des gesamten Körpers. Dafür müssen die Patienten durch ein Mundstück atmen und/oder sich in einer Glaskammer befinden. Unter anderem Klaustrophobie und Gesichtsfeldlähmungen führen dazu, dass eine solche Kammer nicht betreten werden kann oder das Mundstück nicht mit den Gesichtsmuskeln gehalten werden kann. Spirometrie und Bodyplethysmographie sind somit nur bedingt für solche Patientengruppen einsetzbar. Des Weiteren sind die klinischen Goldstandards in der Anschaffung teuer oder benötigen Ersatzmundstücke. Die Bodyplethysmographie ist ein gänzlich orts-gebundenes Verfahren, welches erfordert, dass die Patienten sich immer in einer Praxis einfinden. Im Gegensatz dazu, zeigen kontaktlose, kamerabasierte Verfahren bereits in anderen Bereichen

neue und moderne Möglichkeiten zur Messung von Vitalparametern auf [11–13]. Im Bereich der Respirationsparametermessung stellt insbesondere die kontaktlose Tiefenplethysmographie (DPG, depth-based plethysmography) eine Alternative zu den bisherigen klinischen Goldstands dar [14]. Bei der DPG wird die Abstandsänderung des Oberkörpers zu einer Kamera gemessen. Aus den zeitlichen und räumlichen Information kann über regressive Modelle auf die Atemvolumina geschlossen werden. Mittels DPG lässt sich die Atemmechanik sowie die Beiträge unterschiedlicher Thoraxregionen visualisieren. Des Weiteren wird kein Mundstück benötigt, was eine natürliche Atmung für die Patienten erlaubt. Tiefenkameras sind potentiell leicht und kostengünstig, wie zum Beispiel der Einsatz in modernen Smartphones zeigt [15]. Somit eignet sich DPG auch für mobile Einsatzbereiche im täglichen Leben der Patienten. Die DPG bietet somit gegenüber der Bodyplethysmographie und Spirometrie neue Vorteile für die Messung von Respirationsparametern.

Dass die DPG bisher keine weite Verbreitung gefunden hat, ist auf zum Teil ungelöste Probleme zurückzuführen. Die wesentlichen Herausforderungen stellen die Kalibration des Systems und die Superposition der Thoraxbewegungen der Atmung mit willkürlichen Bewegungen dar. Bisher häufig eingesetzte, lineare Modelle stoßen bei den komplexen Zusammenhängen an ihre Grenzen. Lösungen werden über komplexe und rechenaufwendige Verfahren oder die Verwendung von zusätzlichen Kameras an mehreren Seiten der Probanden gesucht. Diese widersprechen jedoch den eigentlichen Vorteilen der DPG Messung.

Künstliche Intelligenz (KI) beziehungsweise Maschinelles Lernen (ML) bieten das Potential, neues Wissen aus großen Datenmengen zu generieren. Künstliche Neuronale Netze (KNN) sind universelle Funktionsapproximatoren [16]. Ein solches Modell kann anschließend auch auf neue, unbekannte Daten gleicher Art angewendet werden. Ohne das festgelegte Regeln existieren, lassen sich sich Vorhersagen ableiten oder Entscheidungen und Empfehlungen generieren [17]. Verfahren des ML sind bereits in vielen Bereichen der menschlichen Expertise überlegen, insbesondere bei der Ergründung von Zusammenhängen aus komplexen Daten wie zum Beispiel bei der Hautkrebsdetektion [18] oder in der Radiologie [19]. Aus diesen Gründen untersucht diese Arbeit den Einsatz von KI zur kontaktlosen Bestimmung von Respirationsparametern bei DPG Messungen mit einer Tiefenkamera.

Ein kalibrationsfreies System, welches mit nur einer Kamera arbeitet und natürliche Atmung erlaubt, ohne Mundstück und ohne Einschränkungen in der Bewegung, bietet einen enormen Mehrwert für die Gesellschaft und Wissenschaft. Von einer einfach zu bedienenden, mobilen Lösung zur Messung der Atemvolumina profitieren Patienten sowie Ärzte. Ein solches System kann, in Smartphones integriert, in Krisengebieten eingesetzt werden, in denen keine Hardwareaufbauten möglich sind, oder in Szenarien, in denen eine aktive Mitarbeit der Patienten nicht möglich ist. Für den privaten Gebrauch ermöglicht sich eine regelmäßige Überwachung im Alltag und somit eine bessere Kontrolle der Therapie. Davon können auch wissenschaftliche Studien profitieren, wenn zukünftig durch eine vereinfachte Teilnahme mehr Probanden akquiriert werden können und Messungen im echten Leben, außerhalb der Praxis, von ungeschulten Probanden durchgeführt werden können. Solch klinische Studien mit erhöhter Aussagekraft haben das Potential, weitere Fortschritte für Diagnostik und Therapie von respiratorischen Erkrankungen zu erlauben. Für den Routinebetrieb von DPG-Verfahren im Vergleich zur Spirometrie sprechen Hygiene- und Nachhaltigkeitsaspekte. Da die Kamera in keinem direkten Kontakt zu den Patienten steht, werden keine Verbrauchsmaterialien benötigt und auch eine separate Reinigung ist nicht notwendig.

Kontaktlose Messverfahren bieten Vorteile gegenüber den bisherigen klinischen Goldstandards. Jedoch verhindern die bisherigen Limitationen linearer Modelle eine weite Verbreitung. KI bieten das Potential zur vollen Ausschöpfung der Möglichkeiten von DPG. Aus diesem Grund widmet sich diese Arbeit der Untersuchung des Einsatzes von KI in der DPG.

1.2 Arbeitsumfang (Scope of Work)

Diese Arbeit beschäftigt sich mit der kontaktlosen Messung von Atemvolumina über DPG unter Einsatz von Methoden der KI. Die Vision ist es, DPG in Richtung mobiler und leicht anwendbarer Methoden weiterzuentwickeln. Dafür wird lediglich eine Kamera in dieser Arbeit verwendet. Die entwickelten Verfahren sollen ohne vorherige Kalibrierung direkt einsatzbereit sein.

Die Methoden dieser Arbeit werden mit zwei Datensätzen evaluiert. Zum einen dient ein freier, öffentlich zugänglicher Datensatz von Soleimani u. a. [20] als

Grundlage. Anhand dieses Datensatzes werden KI-basierte Methoden mit den konventionellen Signalverarbeitungsmethoden aus der Literatur verglichen. Der Transfer der Ergebnisse wird auf einen neuen Datensatz vorgenommen. Dafür wurden in dieser Arbeit erstmals DPG-Daten von gesunden Probanden und kranken Patienten während der Bodyplethysmographie aufgenommen. Das Ziel ist die direkte Klassifizierung von Krankheiten durch Atem- und Bewegungsmuster.

Diese Studie befasst sich somit mit den Fragestellungen:

- Ist durch den Einsatz von KI eine kontaktlose und kalibrationsfreie Messung von Respirationsparametern möglich?
- Lassen sich Lungenerkrankungen mittels DPG feststellen und klassifizieren?

Diese Studie beschäftigt sich jedoch nicht mit einer Implementierung der DPG Algorithmen auf Smartphones, sondern legt wesentliche Grundsteine für diesen Einsatzbereich.

Als Entwicklungen dieser Arbeit geht ein Vorgehen zur ML-Modellentwicklung hervor, um die Respirationsparameter Tidalvolumen ($[TV] = l$) und langsame Vitalkapazität ($[SVC] = l$) aus den Aufnahmen einer Tiefenkamera zu extrahieren. Dies basiert auf einer gezielten Merkmalsextraktion des aufgenommenen Signals. Des Weiteren wird eine physiologische Maske entwickelt, die als Grundlage für die Signalextraktion dient und Fehlerquellen durch Bewegung unterdrückt. Durch diese Vorarbeit ist ein rekurrentes Modell, ein langes Kurzzeitgedächtnis (LSTM, Long Short-Term Memory), in der Lage, das komplette Volumen-Zeit-Diagramm zu rekonstruieren. Die abschließende technische Neuerung dieser Arbeit ist ein Modell, um zwischen Gesunden und Kranken mittels DPG zu klassifizieren.

1.3 Aufbau dieser Arbeit

Das Kapitel 2 hat das Ziel, die notwendigen Hintergründe für diese Arbeit kompakt darzulegen. Dabei wird auf die Anatomie und Physiologie der menschlichen Atmung eingegangen. Darauf aufbauend werden die Messverfahren und -parameter der Respirationsmessung beleuchtet, darunter die klinischen Goldstandards Spirometrie und Bodyplethysmographie, kontaktlose Messverfahren im Allgemeinen und DPG im Speziellen. Die technischen Messprinzipien von Tiefenkameras werden anschließend erklärt. Abschließend werden Methoden der Künstlichen Intelligenz erläutert.

Im Stand der Forschung in Kapitel 3 werden wissenschaftliche Veröffentlichungen der letzten Jahre analysiert. Dabei wird auf die wesentlichen Aspekte der Signalverarbeitungskette der DPG eingegangen: Messszenarien, Auswahl einer Region von Interesse, Modelle zur Transformation der Signale, Genauigkeit und Evaluierung der Verfahren sowie Klassifikation von respiratorischen Erkrankungen. Nach einer Zusammenfassung und Bewertung des Stands der Forschung erfolgt eine abgeleitete Zielstellung für diese Arbeit.

Die Darstellung der neuen Aspekte aus dieser Arbeit folgt in den nächsten drei Kapiteln. Dabei wird jeweils die Methodik, die verwendeten Daten und Ergebnisse sowie Diskussion beschrieben. Als erste Methode wird die direkte Bestimmung der Respirationsparameter in Kapitel 4 vorgestellt. Dafür werden basierend auf einer Signalvorverarbeitung Merkmale aus dem Tiefen-Zeit-Signal extrahiert. Eine Merkmalsselektion wählt die geeignetsten Merkmale aus und abschließend erfolgt ein Vergleich von geeigneten Regressionsverfahren des Maschinellen Lernens. Zusätzlich wird darauf eingegangen, welchen Einfluss die Merkmale, insbesondere der Einsatz von phänotypischen Merkmalen der Probanden, auf die Regression aufweisen.

Der zweite neue Aspekt dieser Arbeit in Kapitel 5 ist die Untersuchung einer geeigneten Region zur Extraktion des Messsignals. Dafür wird der statistische Zusammenhang mit einem Referenz-Spirometersignal genutzt, um entsprechende Regionen für die individuellen Patienten auszuwählen. Der Einfluss einer Generalisierung dieser einzelnen Masken wird dabei für die merkmalsbasierte Bestimmung der Respirationsparameter und eine gesamtheitliche Transformation des Messsignals mittels rekurrenter neuronaler Netze analysiert.

Die vorausgehenden Erkenntnisse werden in Kapitel 6 genutzt, um eine Klassifikation von respiratorischen Erkrankungen mittels DPG und KI zu untersuchen. Es werden verschiedene Modalitäten der DPG verwendet und entsprechende Modelle miteinander verglichen.

Im Kapitel 7 erfolgt eine abschließende Zusammenfassung der Ergebnisse dieser Arbeit. Die wissenschaftlichen Schlussfolgerungen und Contributions dieser Arbeit werden dargestellt und eine Bewertung wird vorgenommen. Daraus schließt sich ein Ausblick auf zukünftige Forschungsfelder an. Es werden zukünftige Herausforderungen, mögliche Strategien und der langfristige Nutzen dieser Arbeit dargestellt.

Kapitel 2

Grundlagen

Zum Verständnis der Neuerungen im Bereich der Atemphysiologie und Modellentwicklung dieser Arbeit werden in diesem Kapitel die entsprechenden Grundlagen dargelegt. Dafür wird die Anatomie und Physiologie des menschlichen Atemvorgangs dargestellt sowie dazugehörige Messverfahren und Parameter beschrieben. Abschließend werden die in dieser Arbeit verwendeten KI-Verfahren betrachtet.

2.1 Anatomie und Physiologie der menschlichen Atmung

Die Atmung, auch Respiration genannt, ist eine lebenswichtige Funktion des menschlichen Organismus. Es wird in innere und äußere Atmung unterschieden. Bei dem biochemischen Vorgang der inneren Atmung erfolgt auf zellulärer Ebene eine Redoxreaktion von Sauerstoff und Zucker zu Kohlenstoffdioxid und Wasser. Die bei diesem Stoffwechselfvorgang freigesetzte Energie wird für die Funktionsfähigkeit der Organe verwendet. Die äußere Atmung hingegen beschreibt den physiologischen Prozess der Sauerstoffaufnahme in der Lunge. Dies macht mit mehr als 99 % den größten Teil der Sauerstoffaufnahme im menschlichen Körper aus [21]. In dieser Arbeit bezieht sich der Begriff Respiration stets auf die äußere Atmung. In den folgenden Abschnitten wird die Anatomie und Physiologie der Respiration grundlegend dargestellt und die Auswirkungen von Erkrankungen erläutert.

2.1.1 Aufbau des respiratorischen Systems

Der Weg der eingeatmeten Luft folgt den räumlich unterteilten oberen und unteren Atemwegen und ist in Abbildung 2.1 dargestellt. Dabei kann zwischen luftleitenden (konduktiven) und gasaustauschenden (respiratorischen) Abschnitten differenziert werden. Zu letzteren gehören Bronchien und Lunge. Der Brustkorb schützt die inneren Organe und wirkt aktiv an der Atmung mit.

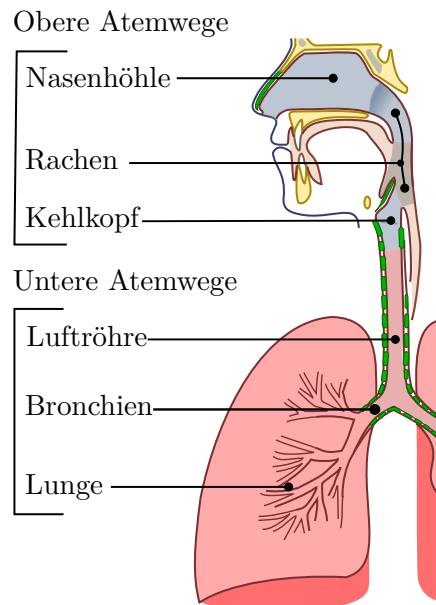


Abbildung 2.1: Atemwege des menschlichen respiratorischen Systems ¹.

Obere und untere Atemwege

Zu den oberen Atemwegen gehören die Nasenhöhle (Cavitas nasi), Mundhöhle, Rachen (Pharynx) und Kehlkopf (Larynx). Durch die Nasenlöcher wird Luft in den menschlichen Körper eingesogen, angewärmt und befeuchtet. Feine Härchen filtern größere Partikel aus der Luft. Die Luft gelangt anschließend in die Mundhöhle. Dorsal schließt sich der Rachen an, welcher den Zugang von Luft und Nahrung in die Luft-, respektive Speiseröhre darstellt. Der Kehildeckel (Epiglottis) im Kehlkopf hat die Funktion zur Verschiebung der Luftöhre. Somit wird verhindert, dass Nahrung und Flüssigkeiten in diese gelangen.

¹Veröffentlicht in die Public Domain von Lord Akryl, Jmarchn und J.Ammon am 24. Oktober 2020. Zugriff online über https://commons.wikimedia.org/wiki/File:Illu_conducting_passages_de.svg.

Kaudal der oberen Atemwege schließen sich die unteren Atemwege an. Diese bestehen aus Luftröhre (Trachea) und Lunge (Pulmo). Die Luftröhre verbindet den Kehlkopf und verzweigt sich in die Hauptbronchien (Bronchus principalis dexter und sinister). Sie ist mit einer Länge von 10 cm bis 12 cm und einem Durchmesser von 1,5 cm bis 2 cm ein elastisches Rohr. Die Aufteilungsstelle (Bifurkation) in die beiden Hauptbronchien ist altersabhängig. Bei älteren Menschen liegt sie tiefer als bei Kindern [22].

Der Aufbau der Lunge wird durch das Verzweigungsmuster der Bronchien diktiert und erfolgt hierarchisch in Lappen (Lobus pulmonalis), Segmente (Segmentum bronchopulmonale), Läppchen (Loboulus pulmonalis) und Azini. Rechte und linke Lunge, respektive Hauptbronchien, weisen unterschiedliche anatomische Ausprägungen auf. Diese Unterschiede sind auf die stärkere Linkslage des Herzens zurückzuführen. Die Hauptbronchien verzweigen sich im Bronchialsystem noch 20 bis 25 mal in kleinere Atemwege und schlussendlich in die Bronchiolen mit einem Durchmesser von circa 0,5 mm [23]. Die Wände der Bronchiolen sind mit glatten Muskeln ausgestattet und erlauben durch Kontraktion und Relaxation eine Veränderung des Volumens. Im Bronchialbaum übernehmen die proximalen Anteile konduktive Aufgaben. Der Gasaustausch findet zwischen den distalen Verzweigungen in den über 300 Millionen Lungenbläschen (Alveolen) und den Kapillaren stattfindet [22].

Aus den nicht am Gasaustausch beteiligten, konduktiven Strukturen bedingt sich das Totraumvolumen. Das Totraumvolumen beschreibt Luft, die nicht am Gasaustausch in den Alveolen teilnimmt. Der anatomische Totraum ergibt sich somit aus den oben genannten Atemwegsstrukturen: Nase, Mundraum, Rachen, Kehlkopf, Luftröhre und Bronchien. Circa 30 % der eingeatmeten Luft pro Atemzug verbleiben in diesen Strukturen [23]. Als alveolärer Totraum werden Regionen in der Lunge bezeichnet, die durch eine Schädigung oder fehlende Durchblutung der Lungenbläschen nicht am Gasaustausch teilnehmen. Im gesunden Gewebe nehmen alle Bereiche der Lunge am Gasaustausch teil. Der funktionelle Totraum als Summe aus anatomischen und alveolären Totraum stimmt somit beim Gesunden mit dem anatomischen Totraum überein.

Brustkorb und Brustwand

Neben den inneren Organen übernimmt der Brustkorb (Thorax) weitere wichtige Funktionen für die Atmung. Der Thorax bildet dabei den Schutz der lebensnotwendigen Organe. Dabei muss er Über- und Unterdruck aushalten und Volumenschwankungen aktiv unterstützen. Der Aufbau des Thorax ergibt sich aus Längselementen der Brustwirbelsäule und dem Brustbein (Sternum), welche durch zwölf Rippenpaaren (Costae) miteinander verbunden sind, vergleiche Abbildung 2.2. Muskeln dichten die Räume zwischen den Rippen (Interkostalräume) ab und halten den wechselnden Druckverhältnissen stand.

Der Aufbau einer Rippe gliedert sich dabei von dorsal nach ventral in:

- Rippenkopf (Caput costae) mit einer Gelenkflächen für die Interaktion mit den Brustwirbelkörpern,
- Rippenhals (Collum costae) mit einer Gelenkfläche für die Brustwirbelquerfortsätze,
- Rippenkörper (Corpus costae) welcher den Thorax halbkreisförmig umgibt und
- Rippenknorpel (Cartilago costalis) als medialer Abschluss in Richtung Sternum.

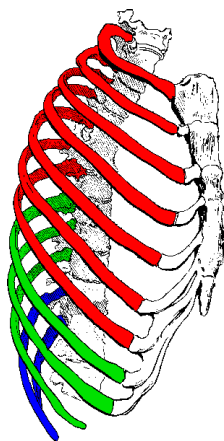


Abbildung 2.2: Thoraxanatomie mit Brustwirbelsäule, Brustbein und Rippenpaaren ohne (blau), mit indirekter (grün) oder direkter (rot) Verbindung zum Brustbein ².

²Veröffentlicht in die Public Domain von Cristobal carrasco am 06. Januar 2021. Zugriff online über <https://en.wikipedia.org/wiki/File:Costillas.png>.

Die ersten sieben Rippenpaare weisen eine direkte Verbindung mit dem Sternum auf. Die Rippenpaare acht bis zehn erreichen das Sternum indirekt und das elfte und zwölfte Rippenpaar endet frei in der Rumpfwand. Kranial ist der Thorax durch den Hals begrenzt, während kaudal das Zwerchfell (Diaphragma) eine Abtrennung zum Abdomen darstellt. Das Zwerchfell ist dabei eine Platte bestehend aus Muskeln und Sehnen, aufgehängt an Lendenwirbeln, Rippen und Brustbein. Diese Thoraxanatomie ist in Abbildung 2.2 dargestellt.

2.1.2 Funktionsweise der Atemmechanik

Unter der Ventilation wird die Belüftung des Respirationstrakts verstanden. Druckdifferenzen sind dabei die wesentliche Antriebskraft für den Gasaustausch zwischen Organismus und Umwelt [4]. In der Einatmung (Inspiration) muss der Druck in den Alveolen niedriger sein als der atmosphärische Umgebungsdruck. Bei der Ausatmung (Expiration) muss das Druckgefälle entsprechend invertiert sein. Zur Herstellung einer negativen Druckdifferenz (Inspiration) wirken:

- Anspannung des Zwerchfells,
- Hebung des Brustkorbs durch Hals- und äußere Zwischenrippenmuskeln und
- sonstige Atemhilfsmuskeln.

An der Expiration sind beteiligt:

- Muskeln, die das Zwerchfell nach oben drängen,
- Senkung des Brustkorbs durch Schwere und Eigenelastizität und
- die Anspannung der inneren Zwischenrippenmuskeln.

Die Atmung lässt sich in die sogenannte Brustatmung (Thorakalatmung) und Bauchatmung (Abdominalatmung) weiter unterscheiden. Durch den höheren Energieverbrauch wird die Brustatmung bei stärkerer Belastung eingesetzt, um die Atemvolumina bei Stress zu erhöhen [24]. Die Kontraktion der Zwischenrippenmuskulatur führt zu einem Anheben der Rippen. Dabei werden die Rippen längs nach außen gedreht, was zu einer Thoraxvergrößerung in sagittaler und transversaler Richtung führt. Die Vergrößerung des Lungenvolumens führt nun zu einer Reduktion des interpulmonalen Drucks. Der Druckgradient gegenüber des Umgebungsdrucks führt zu einem Einströmen der Luft. Bei der Expiration erschlaffen die Zwischenrippenmuskeln entsprechend und die Eigenelastizität der Lunge führt zu einer Kompression des Lungenvolumens, gefolgt von einem Ausströmen der Luft.

Bei der Bauchatmung hingegen erfolgt die Vergrößerung des Lungenvolumens wesentlich in longitudinale (kaudale) Richtung. Die Kontraktion des Zwerchfells führt zu einer Verschiebung der Verdauungsorgane. Der Bauch wölbt sich nach außen vor. Gleichzeitig hebt sich der Brustkorb durch die unteren Rippen an. Durch den entstehenden Unterdruck im Pleuraspalt folgt die Lunge ebenso der Bewegung des Brustkorbs und vergrößert ihr Volumen.

Der Atemrhythmus wird in einem komplexen System unterbewusst geregelt, kann jedoch auch bewusst kontrolliert werden [4, 23]. Die biologischen Regelgrößen der Blutgaskonzentration (Sauerstoff-, Wasserstoff- und Kohlenstoffdioxidgehalt im Blut) werden über Chemorezeptoren gemessen. Über motorische Efferenzen erfolgt anschließend die Regelung der Atmung. Übergeordnete Faktoren können dabei als Störgrößen wirken, wie zum Beispiel sensomotorische Afferenzen bei körperlicher Arbeit, Temperatur oder bewusste Einflussnahme.

2.2 Messverfahren und Parameter der Respiration

Krankhafte Veränderungen der in der Lungenfunktion werden in der klinischen Praxis durch Untersuchungen mit Spirometern oder Bodyplethysmographen festgestellt. Dabei werden Messparameter erhoben, deren Abweichungen zu einer Norm auf das Vorhandensein von Erkrankungen hinweist. Abschnitt 2.2 widmet sich der Darstellung der gemessenen Parameter, sowie den dazugehörigen Messverfahren. Ausgehend von den klinischen Goldstandards, wie Spirometer und Bodyplethysmographen, werden alternative, kontaktlose Messverfahren dargestellt. Die Grundlagen der DPG werden erläutert.

Eine quantitative Beurteilung der Lungenfunktion wird unter anderem benötigt für: präoperative Diagnostik, Prognose für Lungentransplantationen, Kontrolle therapeutischer Interventionen, Verlaufskontrolle von Erkrankungen und Beschwerden, arbeitsmedizinische Tauglichkeitsuntersuchungen und forensische und sozialmedizinische Gutachten [25].

2.2.1 Messparameter

Zur Diagnostik und Überwachung werden verschiedene Messparameter erhoben. Die Überwachung der Atmung umfasst dabei hauptsächlich die Beurteilung der Brustwandbewegung, die Messung physiologischer Parameter wie Luftstrom und Atemvolumen, Atemfrequenz sowie Atem- und Blutgaskonzentrationen [26]. Für spirometrische und bodyplethysmographische Messverfahren stehen dabei die Lungenvolumen im Vordergrund, welche in diesem Abschnitt genauer beschrieben werden. Erfassbare Messparameter für die Untersuchung der Lungenvolumen werden in statische und dynamische Kenngrößen unterteilt. Die folgenden Beschreibungen basieren auf dem Lehrbuch von Oczenski [27].

Statische Lungenvolumen

Statische Lungenvolumen sind nicht zeit-abhängige Messparameter der Lungenfunktion. Abbildung 2.3 stellt sie grafisch mit entsprechenden Referenzwerten dar. Dabei beschreibt das Tidalvolumen ($[TV] = l$), auch Atemzugvolumen genannt, das Atemvolumen, welches bei normaler Ruheatmung ein- und ausgeatmet wird. Das inspiratorische Reservevolumen ($[IRV] = l$) beschreibt das Volumen, welches nach Inspiration in Ruhe zusätzlich maximal eingeatmet werden kann. Das expiratorische Reservevolumen ($[ERV] = l$) hingegen ist definiert als das Volumen, welches nach Expiration in Ruhe zusätzlich maximal ausgeatmet werden kann. Mit dem Begriff Residualvolumen ($[RV] = l$) wird das Volumen beschrieben, welches stets in der Lunge verbleibt und nicht ausgeatmet werden kann.

Weitere statische Lungenparameter ergeben sich aus einer Kombination der einzelnen Lungenvolumen. So ergibt sich die Inspirationskapazität ($[IC] = l$) als Summe von TV und IRV (Gleichung (2.1a)) und beschreibt das Volumen, welches nach ruhiger Expiration maximal eingeatmet werden kann. Das Volumen welches maximal durch forcierte Atmung mobilisierbar ist, wird als Vitalkapazität ($[VC] = l$) bezeichnet. Die VC ergibt sich somit als Summe von TV, IRV und ERV, siehe Gleichung (2.1b). Die Bezeichnung inspiratorische Vitalkapazität ($[VC_{in}] = l$) erfolgt für das größtmögliche Volumen, welches nach maximaler Expiration eingeatmet werden kann. Die expiratorische Vitalkapazität ($[VC_{ex}] = l$) hingegen beschreibt das größtmögliche Volumen, welches nach maximaler Inspiration ausgeatmet werden kann. Das Volumen, welches die Lunge maximal

aufnehmen kann wird als Totalkapazität ($[TC] = 1$) beschreiben. Die TC ergibt sich in Gleichung (2.1c) aus der Summe der maximal mobilisierbaren Luft (VC) und dem stets in der Lunge verbleibenden Luftvolumen (RV). Die funktionelle Residualkapazität ($[FRC] = 1$) ist definiert als das Volumen, welches sich am Ende der ruhigen Ausatmung in der Lunge befindet, siehe Gleichung (2.1d).

$$IC = IRV + TV \tag{2.1a}$$

$$VC = IRV + TV + ERV \tag{2.1b}$$

$$TC = IRV + TV + ERV + RV \tag{2.1c}$$

$$FRC = ERV + RV \tag{2.1d}$$

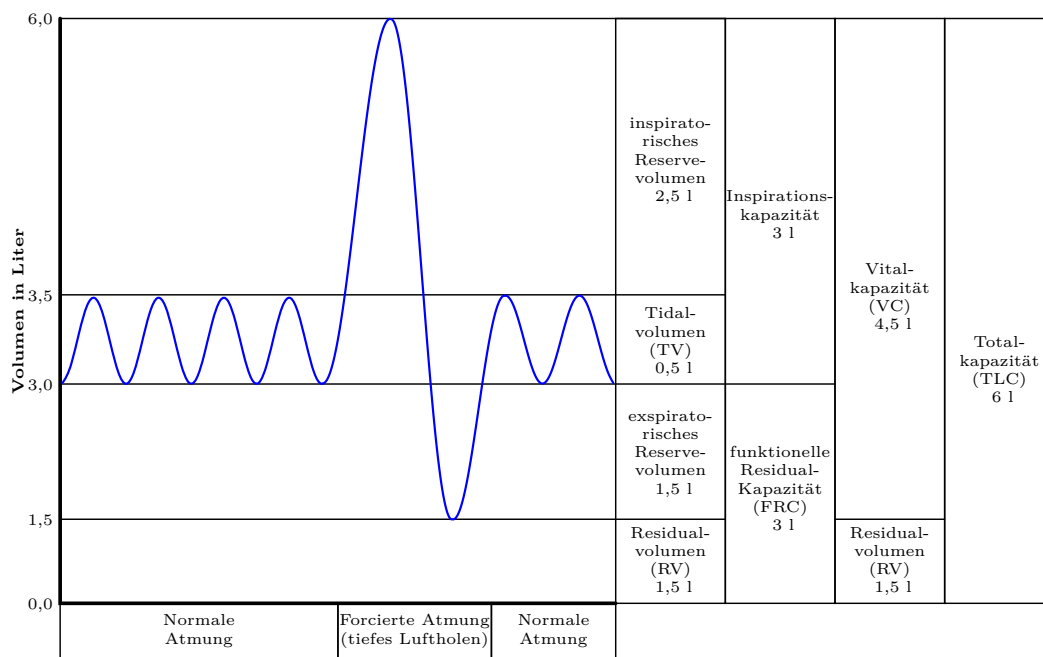


Abbildung 2.3: Statische Lungenvolumina mit Angaben für Normwerte [27]³. Dargestellt ist ein Volumen-Zeit-Diagramm für drei aufeinanderfolgende Phasen: normale Atmung, forcierte Atmung und normale Atmung. Die daraus abgeleiteten Respirationparameter sind auf der rechten Seite abgebildet.

³Veröffentlicht unter der Creative Commons Attribution-ShareAlike 3.0 Unported Lizenz von Vihsadas am 19. August 2020 und adaptiert in die deutsche Version von Dietzel65. Zugriff online über <https://commons.wikimedia.org/wiki/File:Lungenvolumen.svg>. Veränderung der Achsenbeschriftung und Bezeichnungen der statischen Respirationparameter für die Darstellung in dieser Arbeit.

Dynamische Lungenvolumen

Zeitabhängige, dynamische Messgrößen der Lungenfunktion werden aus einem Volumen-Zeit-Diagramm abgeleitet. Sie betrachten den Zeitraum, in welchem der Patient ein Volumen ein- oder ausatmen kann. Das Atemminutenvolumen ($[AMV] = l \text{ min}^{-1}$) beschreibt das Produkt aus TV und Atemfrequenz ($[AF] = \text{min}^{-1}$). Die forcierte Vitalkapazität ($[FVC] = l$) ist definiert als größtmögliches Volumen, welches nach maximaler Inspiration unter maximaler Anstrengung ausgeatmet werden kann. Der Anteil der FVC, welcher nach vollständiger Inspiration maximal innerhalb einer Sekunde ausgeatmet werden kann, wird als Einsekundenkapazität ($[FEV_1] = l$) bezeichnet. Der Tiffeneau-Index ($[FEV_1/FVC] = \%$) setzt FEV_1 und FVC in ein Verhältnis und gibt somit die relative Einsekundenkapazität an. Weitere Kenngrößen sind die maximale Atemstromstärke bei forcierter Expiration (engl. peak expiratory flow, $[PEF] = l \text{ s}^{-1}$) und die mittlere Atemstromstärke bei noch 25% VC in der Lunge (engl. mean expiratory flow, $[MEF_{25}] = l \text{ s}^{-1}$) respektive MEF_{50} sowie MEF_{75} .

Weitere Parameter

Die Aufzählung zusätzlicher Lungenfunktionsparameter kann nahezu unbegrenzt erweitert werden. An dieser Stelle sei noch die Bedeutung von Compliance und Resistance hervorgehoben. Die Compliance ist ein Maß für die Dehnbarkeit der Lunge und die elastischen Eigenschaften des respiratorischen Systems. Sie wird definiert als Quotient der Volumenänderung zur Druckänderung in der Lunge. Eine elektrische Kapazität kann als Analogie zur Compliance angesehen werden. Die Resistance als Maß für den Strömungswiderstand des respiratorischen Systems wird definiert als Produkt der Druckdifferenz zwischen Umgebung und Alveolen und dem Volumenstrom. Ein elektrischer Widerstand bildet die Analogie zur physiologischen Größe.

2.2.2 Veränderungen in der Respiration

Der Vergleich der in Abschnitt 2.2.1 vorgestellten Messgrößen mit einer gesunden Probandenpopulation wird zur Differentialdiagnose durchgeführt. Abweichungen zur Norm können auf pathologische Veränderungen hindeuten. Gleichzeitig haben auch weitere Faktoren einen Einfluss auf die Lungenfunktion.

Pathologische Veränderungen

Störungen des Gasaustausches können die folgenden Bereiche oder eine Kombination aus ihnen befallen [28]:

- Atempumpe,
- Gasaustauschstrecke zwischen den Lungengefäßen und Alveolarraum und
- Atemregulation über das zentrale Nervensystem.

Für diese Arbeit relevant sind dabei die Störungen der Atempumpe, welche sich in obstruktive und restriktive Lungenfunktionsstörungen klassifizieren lassen. Diese werden innerhalb einer eigens angefertigten Studie untersucht, siehe Kapitel 6.

Obstruktive Lungenerkrankungen gehen mit einem erhöhten Atemwegwiderstand und verringerter Absolutwerte forcierter Atmung einher. Sie stellen die häufigsten chronischen respiratorischen Funktionsschwächen dar [6]. Die namensgebende Einengung (Obstruktion) der Luftwege kann zum Beispiel durch Schleim, Tumore oder Fremdkörper entstehen. Bei einer Bronchialobstruktion ist der Tonus der glatten Muskulatur erhöht. Dies kann anfallsweise (Asthma Bronchiale) oder bei Rauchern in Rahmen von COPD zu dauerhaften Folgen führen. Diese Verengung der Atemwege hat einen gegenüber der Norm bis zu fünffach erhöhten Atemwegwiderstand zur Folge und führt zu klinischer Atemnot (Dyspnoe), besonders deutlich bei forcierter Expiration [5, 28]. FEV_1/FVC kann auf unter 30 % sinken [28]. Als weitere Folge kann auch das Residualvolumen erhöht sein.

Bei restriktiven Erkrankungen hingegen ist die Expansionsfähigkeit der Lunge eingeschränkt und es liegt eine verminderte Dehnbarkeit (Compliance) von Lunge und/oder Thorax vor. Daraus resultiert eine verringerte totale Lungkapazität und Vitalkapazität [6]. Pulmonale Restriktionen können zum Beispiel durch Bindegewebsvermehrung der Lunge (zum Beispiel idiopathische Lungenfibrose) hervorgerufen werden. Versteifungen des Brustkorbs oder eine Kompression von Lungengewebe (Pneumothorax) können eine extrapulmonale Ursache für eine Restriktion darstellen.

Es handelt sich an dieser Stelle um eine Auswahl an dargestellten Krankheitsbildern, welche keinen Anspruch auf Vollständigkeit erhebt. Die vorgestellten Pathologien orientieren sich an den in dieser Arbeit behandelten, respiratorischen Erkrankungen. Für die messbaren Auswirkungen von restriktiven und obstruktiven Krankheiten auf Lungenfunktionsparameter sei auf Tabelle 2.1 verwiesen.

Tabelle 2.1: Vergleich restriktiver und obstruktiver Ventilationsstörungen hinsichtlich Ursachen und Auswirkungen auf die Messparameter VC, FEV₁ und FVC. ↑ symbolisiert dabei einen Anstieg und ↓ eine Verminderung der Parameter.

	Obstruktiv	Restriktiv
Ursachen	Strömungsquerschnitt ↓	diverse
Krankheitsbilder	COPD, Asthma bronchiale	Zwerchfelllähmung, Skoliose
Auswirkungen	Resistance ↑ FEV ₁ ↓ , FEV ₁ /FVC ↓	Compliance ↓ ¹ VC ↓ , FEV ₁ ↓, RV ↑, TC ↑

¹ nicht bei Störungen der atemmuskulären Pumpe.

Weitere Einflussfaktoren

Neben pathologischen Änderungen sind Variationen in den Messgrößen der Lungenfunktion auch auf phänotypische Einflussfaktoren zurückzuführen. So kommt es mit zunehmender Lebenszeit zu einer altersbedingten Osteoporose und einer Versteifung des Brustkorbs. Dies hat eine verringerte Beweglichkeit der Rippen und somit ein erhöhtes Residualvolumen zur Folge [29]. Auch die Körpergröße steht im direkten Zusammenhang mit dem Tidalvolumen, da ein größerer Körper über eine größere Lunge verfügt [30]. Das Gewicht und die Fettverteilung haben einen weiteren Einfluss auf die Lungenfunktion. Bei adipösen Patienten ist die Atemfrequenz erhöht und Tidalvolumen sowie weitere Lungenvolumina verringert [31]. Der Einfluss des Geschlechts ist hingegen umstritten und möglicherweise auf die anderen genannten Einflussfaktoren zurückzuführen [32]. Die Lungenfunktionsparameter stehen des weiteren im Zusammenhang mit der sportlichen Leistungsfähigkeit. So sind diverse Lungenvolumen bei zum Beispiel Profi-Sportlern im Vergleich zu einer gesunden Kontrollgruppe erhöht [33]. Bei der Messung dieser Parameter ist jedoch auch Wert auf die aktuelle Körperposition zu legen, da diese die Funktionsfähigkeit der Lunge beeinflusst. So erzielen Messungen im Stehen höhere Lungenvolumina, als Messungen in Rückenlage [34].

2.2.3 Klinische Goldstandards

Die verbreitetsten medizinischen Untersuchungen zur Bewertung der Lungenfunktion sind Spirometrie und Ganzkörperplethysmographie. Beide Verfahren werden von Lungenspezialisten durchgeführt, um Abweichungen der Messparameter von Normwerten festzustellen und somit Krankheiten zu diagnostizieren. In den folgenden Abschnitten werden diese Verfahren vorgestellt, da sie als klinischer Goldstandard in der späteren Arbeit zur Evaluierung der entwickelten Algorithmen dienen.

Spirometrie

Spirometrie ist das verbreitetste Verfahren zur pulmonalen Routineuntersuchung. Dabei werden Lungenvolumina und Volumenfluss gemessen. In der modernen Spirometrie werden sogenannte offene Systeme verwendet, welche den Volumenfluss der Gase im zeitlichen Verlauf messen. Dieses Verfahren wird auch als Pneumotachographie bezeichnet. Über ein Mundstück atmet der Patient Umgebungsluft ein und aus. Ein spirometrischer Sensor misst dabei den mittleren Volumenfluss $[\dot{V}] = \frac{1}{s}$ im Atemrohr [35].

Aus dem Surrogat-Signal des zeitlichen Volumenflusses lassen sich über Integration die Atemvolumina $[V] = l$ bestimmen:

$$V = \int \dot{V} dt \quad (2.2)$$

Für die Messung des Flusses können unterschiedliche Sensoren und Messprinzipien verwendet werden [35]. Erwähnenswert sei an dieser Stelle das Verfahren nach Fleisch 1925. Im Strömungsrohr befinden sich parallel angeordnete Kapillaren als Strömungswiderstände. Diese sorgen für eine laminare Strömung. Die Bestimmung des Flusses erfolgt über das Gesetz nach Hagen-Poiseuille, in dem der Volumenfluss proportional zum gemessenen Druckabfall ΔP über dem bekannten Widerstand R ist:

$$\dot{V} = \frac{\Delta P}{R} \quad (2.3)$$

Lamellen, Düsen oder Gitter können ebenso genutzt werden, um eine laminare Strömung zu erzeugen. Alternativ lassen sich auch über Turbinen- oder Ultraschallbasierende Messverfahren Fluss und Volumen messen [35].

Offene Systeme zeichnen sich durch eine geringe Gerätegröße aus, weisen aber nichtlineares Verhalten auf. So sind die Flusssensorprinzipien unter anderem temperatur- und gasartsabhängig. Es wird eine regelmäßige Kalibration und Reinigung benötigt, häufig kompensiert durch Einwegmundstücke zu Lasten der Umwelt [35].

Bodyplethysmographie

In der erweiterten Lungenfunktionsdiagnostik werden Bodyplethysmographen, auch als Ganzkörperplethysmographen bezeichnet, eingesetzt zur Differentialdiagnostik und Bestimmung der nicht ventilierbaren Atemvolumina, wie zum Beispiel dem Residualvolumen. Der Patient begibt sich dafür in eine geschlossene Kammer. Die Atmung wird anschließend über Volumen- und Druckänderungen im Kamerinneren gemessen. Heutige Bodyplethysmographen setzen dabei hauptsächlich auf die volumenkonstante Methode, in welcher der Patient die Kammerluft durch ein Flusssensor atmet [35]. Ein Verschlussmechanismus (Shutter) kann die Atmung kurzzeitig unterbrechen.

Das Messprinzip basiert auf dem Gesetz von Boyle und Mariotte: Bei Zustandsänderungen eines idealen Gases ist das Volumen umgekehrt proportional zum Druck bei fester Teilchenzahl und konstanter Temperatur. So bewirkt eine Vergrößerung des Thoraxvolumens, durch den sich hebenden Brustkorb bei der Inspiration, eine Volumenverkleinerung in der Kammer mit gleichzeitig steigenden Kammerdruck. Über den Shutter wird die Ventilation bei der Ausatmung unterbrochen. Somit kann der Alveolardruck am Mundstück gemessen werden. Die FRC kann anschließend über das Produkt einer Eich-Konstante mit dem Verhältnis aus den Messgrößen Kammerdruck und Munddruck bestimmt werden.

2.2.4 Tiefenplethysmographie

Tiefenplethysmographie (DPG, depth-based plethysmography) ist ein kontaktloses Verfahren, bei dem Abstandsinformationen genutzt werden, um auf die Volumenschwankungen im Oberkörper bei der Atmung zu schließen. Der räumliche Abstand des Probanden zur Kamera kann für die Koordinaten x (Breite) und y (Höhe) gemessen werden. Im zeitlichen Verlauf wird die Änderung der Oberkörpers analysiert und ein Volumen-Zeit-Signal extrahiert.

Die wesentlichen Schritte der DPG sind in Abbildung 2.4 dargestellt. Zuerst wird ein Bereich von Interesse (ROI, Region of Interest) aus dem Tiefenbild I ausgewählt und ein Messwert s zum Zeitpunkt t aus der ROI über eine Extraktionsfunktion f extrahiert, mathematisch dargestellt in Gleichung (2.4). I besitzt Pixel in horizontaler Ausprägung x und in vertikaler Ausrichtung y .

$$s[t] = f[I[x, y, t]] \quad (2.4)$$

Nach einer möglichen Vorverarbeitung wird das Signal ($[s] = \text{mm}$) nun über eine Modellfunktion g auf das Zielsignal (zeitliches Volumensignal $[v] = 1$) abgebildet, siehe Gleichung (2.5a). Zusätzlich kann ein Referenzsignal r aus einem Referenzsystem wie zum Beispiel einem Spirometer vorliegen. Für die Optimierung des Modells wird die Modellfunktion f so gewählt, dass die Differenz aus f und r minimal wird, siehe Gleichung (2.5b).

$$v[t] = g[s[t]] \quad (2.5a)$$

$$r[t] - g[s[t]] \rightarrow \min \quad (2.5b)$$

In das Modell können optional noch weitere phänotypische Merkmale der Testperson einfließen. Abschließend können die Atemparameter aus dem Modell extrahiert werden. Im Gegensatz zur traditionellen Spirometrie hat die DPG mehrere Vorteile:

- Mechanik und Beiträge der Atembewegung werden sichtbar gemacht
- DPG entspricht der natürlichen Atmung und
- DPG ist eine potentiell mobile, leichte und kostengünstige Methode.

Die Beiträge der einzelnen Thoraxregionen zur Atembewegung können bei der DPG visualisiert und bewertet werden. Dazu gehört zum Beispiel die unterschiedliche Atemmechanik bei Schwimmern [36], Tänzern [37] oder Säuglingen [38]. Darüber hinaus ist es denkbar, dass asynchrone Muskelschwächen oder sogar der Ausfall eines Lungenlappens durch eine gezielte Volumenbestimmung einzelner Regionen sichtbar gemacht werden können. Dies ist mit der herkömmlichen Spirometrie nicht möglich [39, 40]. Mit Atemfrequenz, Atemvolumen und Brustkorbbewegungen ermöglicht die DPG die Messung von drei der vier Klassen der Atmungsbeurteilung [26]. Nur die Gaskonzentration kann mit der DPG nicht gemessen werden.

Für die spirometrische Messung wird ein Mundstück benötigt. Ein solches Mundstück kann aber nicht von allen Patientengruppen verwendet werden. Insbesondere bei einer Gesichtsmuskelschwäche kann es zu Abweichungen bei den Messungen kommen [41]. Andere Patientengruppen, zum Beispiel mit Tracheostomie, können ein solches Mundstück gar nicht erst verwenden [42]. Die DPG kann ohne aktive Beteiligung des Patienten am Tidalvolumen durchgeführt werden, da kein Mundstück erforderlich ist. Eine berührungslose Messung in Ruhe kann problemlos durchgeführt werden, insbesondere bei Kindern, Hörgeschädigten, Lernbehinderten oder Menschen mit Sprachbarrieren. Die Atmung wird dabei nicht durch weitere Randbedingungen beeinflusst.

Abgesehen vom Entwicklungsstand und der verwendeten Technologie bieten DPG-Verfahren die Möglichkeit, einfach und überall eingesetzt zu werden, ohne dass geschultes Personal benötigt wird. Dies gilt nicht für Mehrkamerasysteme, die eine weitere Kalibrierung oder die Anbringung von Markern, erfordern. Einzelkamerasysteme mit Tiefensensoren, wie von Takamoto u. a. [43] beschrieben, die das Atemvolumen ohne Kalibrierung bestimmen können, bieten die oben beschriebenen Vorteile. Mit der Verbreitung von Tiefensensoren in mobilen Smartphone-Kameras [15] können solche Technologien potentiell in Zukunft eine einfache Messung von Atemungsparametern im Alltag von Patienten ermöglichen. Im Vergleich zum Ganzkörperplethysmographen, bei dem es zu Problemen aufgrund von Klaustrophobie kommen kann [44], ist die DPG nicht durch räumliche Anforderungen eingeschränkt und kann mobil eingesetzt werden. Die verwendeten Tiefensensoren zeigen im Vergleich zur Gold-Standard-Technologie eine niedrigere Kostenstruktur auf. Keine weiteren Verbrauchsmaterialien werden benötigt.

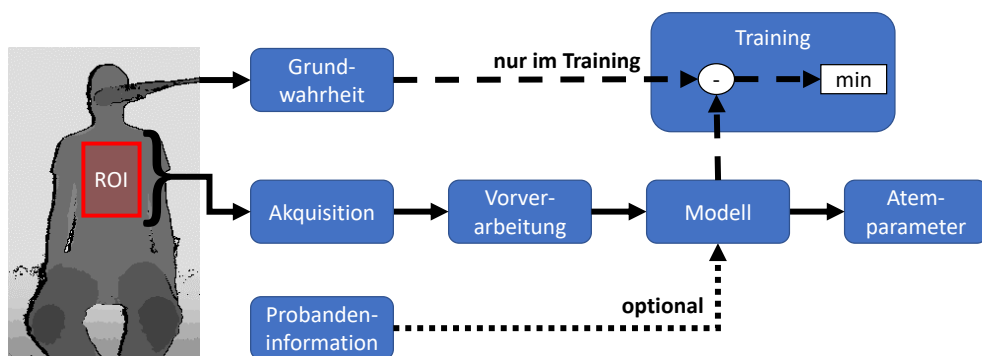


Abbildung 2.4: Allgemeiner Ablauf der Tiefenplethysmographie.

Aus diesen Vorteilen leiten sich die Anwendungsgebiete für die DPG nach Rehouma u. a. [26] ab. Im klinischen Bereich ist die Diagnose von Atemwegserkrankungen wie COPD oder Asthma von besonderer Bedeutung. Die Anwendungsszenarien erstrecken sich daher auf den klinischen Bereich, als Ersatz, wenn andere Geräte zum Beispiel aufgrund von Klaustrophobie oder Gesichtsmuskelschwäche nicht eingesetzt werden können; als mobile Anwendung im Heim- und Pflegebereich zur spontanen und mobilen Überwachung von Atmungsparametern ohne geschultes Personal; zur Beurteilung der Atmungsmechanik; in schwer zugänglichen Bereichen oder zur berührungslosen Überwachung zum Beispiel in Pflegeheimen, Krankenhäusern oder Gefängnissen.

2.3 Tiefenkameras

Im Vergleich zu konventionellen Videokameras mit Farbkanälen, sind Tiefenkameras aufgrund ihres Messprinzips in der Lage, Abstände und Distanzen zu bestimmen. Bei den verwendeten Messprinzipien handelt es sich um Strukturier-tes Licht (SL, Structured Light), Laufzeitverfahren (TOF, Time Of Flight) und Stereoskopie (SC, Stereoscopia). Diese Messprinzipien werden in den folgenden Abschnitten grundlegend erläutert. Für eine detailliertere Darstellung sei auf die Ausführungen von Giancola, Valenti und Sala [45] verwiesen. Die Messung der Tiefeninformation ist dabei die Grundlage für die Bestimmung der respirationsbedingten Thoraxausdehnung in der DPG.

Stereoskopie

Um 3D-Informationen zu extrahieren wird über SC das Ziel aus verschiedenen Blickwinkeln betrachtet. Ähnlich wie bei der Verarbeitung des Seheindrucks der menschlichen Augen im Gehirn, wird die Disparität der Kamerabilder ausgenutzt, um die Tiefe zu rekonstruieren. In Abbildung 2.5 ist zu sehen, wie ein Proband im Bild beider Kameras zu sehen ist. Über trigonometrische Verhältnisse und den bekannten Abstand und Winkel beider Kameras lassen sich für die registrierten Bildpunkte die Abstände zu den Kameras bestimmen [45].

Strukturiertes Licht

Ähnlich wie auch bei der SC basiert SL auf der Triangulation. Beim SL Verfahren wird eine Kamera, in Kombination mit einem strukturiertem Muster eingesetzt, welches in die Szenerie projiziert wird. Das kodifizierte Muster enthält Strukturen, die eindeutige Korrespondenzen für die Triangulation liefern. Das ursprüngliche Muster ist der Kamera a priori bekannt, sodass aus der gemessenen Formänderung die Tiefeninformationen extrahiert werden können [46]. Abbildung 2.5 stellt das Messprinzip schematisch dar.

Laufzeitmessverfahren

TOF ist Verfahren, welches seit mehr als einem Jahrhundert für die Entfernungsmessung eingesetzt wird. Während früher meist Schall- oder Funksignale genutzt wurden, ist es durch den technischen Fortschritt in modernen Messgeräten möglich, Lichtsignale für TOF zu nutzen. Ein solches TOF-System besteht aus Empfänger und Sender. Der Sender sendet ein Lichtsignal, welches von Objekten in der Szene zum Empfänger reflektiert wird. Aus der Laufzeit von Sender zum Empfänger kann über die Lichtgeschwindigkeit die Distanz zum Objekt bestimmt werden [47]. Ein solcher Vorgang ist in Abbildung 2.5 dargestellt.

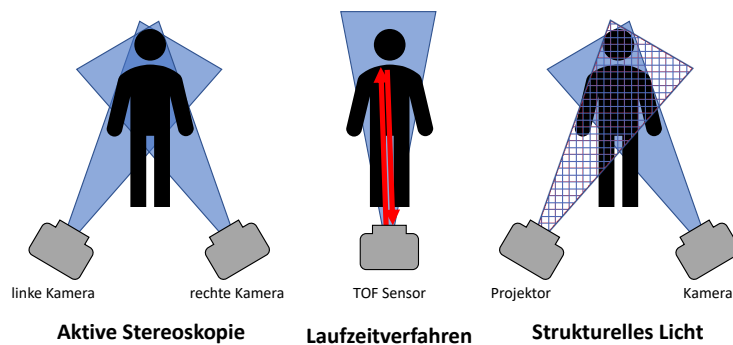


Abbildung 2.5: Kameraprinzipien für Tiefenmessungen [14] ⁴.

⁴Veröffentlicht unter der Creative Commons Attribution Lizenz. Abbildungen 2 (a) der Originalveröffentlichung für diese Arbeit übersetzt.

2.4 Methoden der Künstlichen Intelligenz

Der Begriff KI bezieht sich auf die Abbildung menschlicher Intelligenz durch technische Geräte, mit dem Ziel, Aufgaben zu bewältigen, wie zum Beispiel in der Robotik oder Entscheidungsfindung. Das ML ist eine spezifische Teildisziplin der KI, bei der Maschinen ohne explizite Programmierung eigenständig aus Daten lernen, Muster zu erkennen und Vorhersagen zu treffen.

2.4.1 Merkmalsselektion

Ein Verfahren zur Dimensionsreduktion ist die Extraktion von Merkmalen aus zum Beispiel Zeitsignalen oder Bildern. Eine anschließende Merkmalsauswahl hat die Ziele [48]:

1. Datenreduktion, zur Begrenzung des Speicherbedarfs und Erhöhung der Geschwindigkeit der Algorithmen,
2. Reduzierung der Merkmalsanzahl, um Ressourcen bei der Datensammlung oder Nutzung zu sparen,
3. Leistungsverbesserung, um die Vorhersagegenauigkeit zu erhöhen und
4. Datenverständnis, um Kenntnisse über den Prozess zu gewinnen oder Daten zu visualisieren.

Methoden der Merkmalsselektion lassen sich dabei den Kategorien Filter-, Wrapper- und Embedded-Methoden zuordnen, mit den folgenden Eigenschaften nach [48]. Filter-Methoden erstellen ein Ranking der Merkmale anhand von relevanter Kennzahlen, wie zum Beispiel Korrelationskoeffizienten. Das Verfahren ist unabhängig vom gewählten ML-Modell und ist schnell und einfach anzuwenden. Die Wechselwirkung zwischen den Merkmalen wird jedoch nicht berücksichtigt. Wrapper-Methoden bewerten Merkmale anhand ihrer Leistung für spezifische Modelle. Verschiedene Merkmals-Kombinationen werden am Black Box Modell ausprobiert und die beste ausgewählt. So können auf Kosten der Rechenzeit die Wechselwirkung zwischen den Merkmalen berücksichtigt werden. Embedded-Methoden kombinieren den Vorteil von Filtern und Wrappern, indem die Merkmalsauswahl direkt in den Trainingsprozess der ML-Modelle integriert wird. Die ausgewählten Merkmale gelten für das ausgewählte Modell.

2.4.2 Modelle des Maschinellen Lernens

Zu den bekanntesten Verfahren des ML gehören KNN [49]. Die Verarbeitung ist dabei an menschliche Neuronen angelehnt. Die Daten einer Eingangssequenz x_1, \dots, x_N werden mit entsprechenden Gewichten w_i gewichtet. Die Summe aus gewichteten Eingangsdaten und einem Bias b wird als a bezeichnet und entspricht der Aktivierung des Zellkerns, siehe Gleichung (2.6a). In Gleichung (2.6b) bestimmt die Aktivierungsfunktion f_{KNN} aus a den Ausgang y , welcher in einem Netzwerk von Neuronen der nächsten Schicht als Eingang dient. So können KNN mit Eingangsschicht, Zwischenschichten und einer Ausgabeschicht erstellt werden. Das Training von KNN ist ein Optimierungsproblem der gegebenen Grundwahrheit \tilde{y} und dem Ausgang \hat{y} . Alle Gewichte werden über Fehler-Backpropagation angepasst. Der Fehler wird dabei über eine sogenannte Loss-Funktion g bestimmt, siehe Gleichung (2.6c). KNN sind universelle Funktionsapproximatoren und somit in der Lage, jede mathematische Lage approximieren zu können.

$$a = b + \sum_{i=1}^N w_i x_i \quad (2.6a)$$

$$y = f_{\text{KNN}}(a) \quad (2.6b)$$

$$\text{Loss} = g(\tilde{y}, \hat{y}) \quad (2.6c)$$

Ein neuronales Faltungsnetzwerk (CNN, convolutional neural network) ist ein neuronales Netz, welches speziell für die Mustererkennung in Eingabedaten wie Bildern entwickelt wurde [50]. Die Convolutional Layers sind spezielle Schichten, die Filter auf die Eingabedaten anwenden, um Merkmale zu extrahieren, siehe Gleichung (2.7a). Jede Aktivierung der Filter wird berechnet als das Skalarprodukt zwischen dem Filter und einem Ausschnitt aus den Eingabedaten. Die Pooling-Schichten reduzieren die Größe der Aktivierungen und machen sie invariant gegenüber kleinen Translationen im Eingabebild. Eine gängige Pooling-Funktion ist das Max-Pooling, vergleiche Gleichung (2.7b). Die Aktivierungen der letzten Convolutional Layer werden in eine Fully-Connected Layer überführt, in der die Entscheidungen getroffen werden.

$$a_{i,j}^l = \sigma\left(\sum_{m=1}^{M_{l-1}} \sum_{u=1}^k \sum_{v=1}^k w_{u,v,m}^l a_{i+u-1,j+v-1}^{l-1} + b_m^l\right) \quad (2.7a)$$

$$h_{i,j}^l = \max_{u=1}^p \max_{v=1}^p a_{(i-1)p+u,(j-1)p+v}^{l-1} \quad (2.7b)$$

Hierbei ist $a_{i,j}^l$ die Aktivierung des i -ten und j -ten Features in der l -ten Schicht, M_{l-1} die Anzahl der Features in der vorherigen Schicht, k die Größe des Filters, $w_{u,v,m}^l$ das Gewicht des Filters an der Position (u, v) für das m -te Feature in der vorherigen Schicht, $a_{i+u-1,j+v-1}^{l-1}$ die Ausgabe des u -ten und v -ten Features in der vorherigen Schicht für das (i, j) -te Feature in der aktuellen Schicht und b_m^l der Bias-Term des m -ten Filters in der l -ten Schicht. p entspricht der Größe des Pooling-Fensters.

Zur Verarbeitung von Sequenzen, in denen die zeitliche Reihenfolge von Bedeutung ist, werden rekurrente neuronale Netze eingesetzt. LSTM-Modelle als spezielle Art von rekurrenten Netzen sind in der Lage, Informationen über längere Zeiträume zu speichern und somit Zusammenhänge in den Eingangsdaten zu erkennen [51]. Durch Schleifen sind sie in der Lage, zurückliegende Zustände zu speichern und in die Berechnung der aktuellen Ausgabe einzubeziehen. Dafür werden sogenannte Gates verwendet, welche als Steuergrößen den Informationsfluss regulieren [52].

Support Vector Machine (SVM) sind eine Methode zur Trennung von Daten anhand eines hyperplanaren Entscheidungsraums [53]. Sie besitzen eine hohe Generalisierungsfähigkeit auch bei höher dimensional Daten. Die Entscheidungsgrenze wird durch eine Maximierung des Abstands zu den nächsten Datenpunkten definiert, die Support-Vektoren genannt werden. Die finale Entscheidungsfunktion f_{SVM} in Abhängigkeit vom Eingang x in Gleichung (2.8) ergibt sich mithilfe der Kernfunktion K , einem Bias b und α dem Variablen gewicht.

$$f_{\text{SVM}}(x) = \text{sign}\left(\sum_{i=1}^N \alpha_i y_i K(x_i, x) + b\right) \quad (2.8)$$

Die Signum-Funktion gibt eine Ausgabe von $+1$ oder -1 zurück, je nachdem, auf welcher Seite der Entscheidungsgrenze sich der Eingabevektor befindet. Die Wahl des Kerns beeinflusst die Wahl der Entscheidungsgrenze, wobei der lineare Kern eine lineare Grenze erzeugt und der radiale Basisfunktions (RBF)-Kern eine

nicht-lineare Grenze erzeugen kann. Der Fehlerparameter C stellt die Kompromisslösung zwischen der Maximierung des Abstands zwischen den Klassen und der Minimierung von Klassifizierungsfehlern dar. Es gilt $0 \leq \alpha_i \leq C$ mit C als Fehlerparameter.

Ensemble-Lernen (EL, ensemble learning) ist eine Methode, um die Ausgaben mehrerer Modelle zu kombinieren [54]. Das Ziel ist, die Stärken und Schwächen der einzelnen Modelle zu kombinieren und eine robustere Vorhersage zu treffen. Eine Möglichkeit des EL ist somit eine Gewichtung der einzelnen Modelle, siehe Gleichung (2.9). Dabei sind f_i die einzelnen Modelle in Abhängigkeit vom Eingang x gewichtet mit den Faktoren ω_i .

$$f_{\text{EL}}(x) = \sum_{i=1}^N \omega_i f_i(x) \quad (2.9)$$

Random Forest (RF, Random Forest) ist eine Art des EL, bestehend aus einer großen Anzahl an Entscheidungsbäumen [55]. Entscheidungsbäume teilen sich rekursiv in zwei Untergruppen anhand der verfügbaren Merkmale. Die einzelnen Bäume werden zufälligen Teilmengen der Eingangsdaten trainiert. Die Vorhersage ergibt sich aus dem Durchschnitt der Vorhersagen aller Bäume. RF zeichnet sich dadurch aus, mit vielen Merkmalen und mehrdimensionalen Eingangsdaten, ohne Optimierung von Hyperparametern, gute Klassifikations- und Regressionsergebnisse zu erzielen.

Kapitel 3

Stand der Forschung

In diesem Kapitel erfolgt eine wissenschaftliche Darstellung des aktuellen Forschungsstands bezüglich der kontaktlosen Messung von Respirationsparametern durch die DPG. Es werden sowohl die bestehenden Entwicklungen als auch die bisher weniger beachteten Forschungsrichtungen herausgearbeitet. Daraus ergeben sich Anknüpfungspunkte und die Forschungsfragen für diese Arbeit. Die Ausführung in diesem Kapitel basiert auf der eigenen Veröffentlichung [14]. Da DPG im Allgemeinen bereits in Abschnitt 2.2.4 beschrieben wurde, erfolgt hier eine detaillierte Betrachtung der einzelnen Komponenten, wie in Abbildung 2.4 dargestellt.

3.1 Auswahl einer Region of Interest

Ausgehend von Bilddaten mit Tiefeninformationen wird eine ROI definiert. Eine ROI beschreibt den relevanten Messbereich. Diese Teilauswahl ermöglicht es, die folgende Auswertung auf dem ausgewählten Bereich effizient und zielgerichtet auszuführen, indem Bereiche ohne Relevanz nicht berücksichtigt werden. Die Verwendung einer ROI reduziert unter anderem die Rechenkapazität und verbessert das Signal-Rausch-Verhältnis (SNR, Signal-to-Noise Ratio). Für die DPG Anwendung kann eine ROI auf dem Oberkörper ausgewählt werden, um zum Beispiel den Thorax auszuwählen. Eine Teilmenge der ROI wird im Folgenden durch den Begriff Teilbereich von Interesse (sROI, sub Region of Interest) beschrieben. Mehrere sROIs können verschiedene Bereiche des Thorax darstellen und separat verarbeitet werden.

3.1.1 Auswahlverfahren

Eine ROI kann manuell, halbautomatisch oder vollautomatisch ausgewählt werden. Bei der manuellen Auswahl der ROI ist der Anwender selbst direkt beteiligt und wählt den Bereich aus [38]. In der halbautomatischen Auswahl wird der Anwender durch die Auswertesoftware unterstützt. Markerbasierte Verfahren können diesen halbautomatischen Methoden zugeordnet werden. Hierbei werden Marker auf den Probanden geklebt, welche im Anschluss automatisch erkannt und als Bereichsauswahl fungieren [36, 37, 42, 56–59]. Vollautomatische ROI-Auswahlverfahren benötigen keinerlei Eingriffe durch den Anwender [60–63].

Eine manuelle Auswahl einer ROI ist ein simples Verfahren mit Nachteilen hinsichtlich der Reproduzierbarkeit und automatisierten Anwendung. So erfolgt eine manuelle, rechteckige Auswahl des Thorax durch Wiegandt u. a. [38] und Harte u. a. [64]. Aufgrund der hohen Anzahl an Bildern, welche pro Sekunde aufgenommen werden, ist eine Wiederholung der ROI-Auswahl nicht praktikabel und wird einmalig zu Beginn der Messung vorgenommen. Dies kann jedoch dazu führen, dass es durch Atembewegungen zu Verschiebungen im Bildbereich kommt. Relevante Bereiche können dabei aus der ROI fallen und dafür Störquellen hinein gelangen. Aufgrund der Individualität in der Auswahl ist eine ROI nicht vollständig reproduzierbar.

Markerbasierte Verfahren bieten eine genauere Präzision, dafür jedoch auch einen erhöhten Aufwand zur Vorbereitung. Für DPG werden zwischen fünf und 89 Marker auf den unbedeckten Oberkörper der Probanden geklebt [56, 59]. Die genaue Zahl der Marker ist dabei jedoch nicht begrenzt und hängt auch von der zu untersuchenden Region ab [65]. Die Anbringung der Marker benötigt direkten Probandenkontakt und kann durch die Hautverschiebung auf den Knochen erschwert sein.

Beim Flutfüllung-Algorithmus (engl. Floodfill) durch Addison u. a. [66] werden, ausgehend von einem manuell ausgewählten Startpunkt, iterativ benachbarte Pixel mit ähnlichen Eigenschaften ausgewählt. So wird halbautomatisch eine ROI ausgewählt.

Eine weitere Möglichkeit der ROI-Bestimmung bietet sich durch eine vorherige Skelettierung. Softwarebibliotheken wie OpenPiPaf [67], MMPose [68], leighweight-humanpose-estimation von PyTorch [69], PoseNet [70] und MediaPipe [67] detektieren markante Körperpunkte (engl. skeletal joints), wie zum

Beispiel Augen, Schultern, Hüfte und weitere. Diese können dann wiederum zur Berechnung einer ROI genutzt werden. Soleimani u. a. [71] nutzen detektierte Körperpunkte durch eine Kinect v2 zur Auswahl der ROI. Die detektierbaren Punkte unterscheiden sich dabei auch zum Teil zwischen den oben genannten Skelettierungsbibliotheken hinsichtlich physiologischer Zuordnung und somit Positionierung. Alternativ verwenden Imano u. a. [63] eine automatische Oberkörperdetektion durch das OpenCV haar-cascade Modell [72] zur Bestimmung einer vorläufigen ROI. Eine weitere zeilenweise Aufteilung erfolgt, um anschließend die Regionen auszuwählen, die über 90 % der Gesamtamplitude verfügen. Die Auswahl ist in Abbildung 3.1 veranschaulicht.

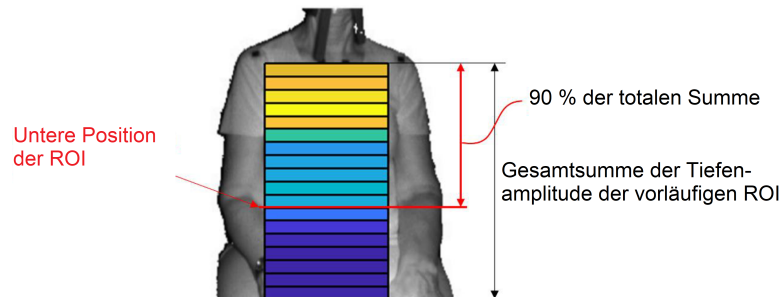


Abbildung 3.1: Auswahl einer Region von Interesse (ROI) durch Amplituden [63]¹. Die vorläufige ROI wird in Streifen unterteilt. Von oben nach unten werden Streifen ausgewählt, bis 90 % der Gesamtamplitude erreicht ist.

Neben der Körperposition können auch spatiale Nachbarschafts- und temporale Informationen zur Bestimmung einer ROI genutzt werden. Abbildung 3.2 stellt die Level-Set Segmentierung nach Oh u. a. [60] dar. Dabei werden a priori Informationen bezüglich eines manuell definierten Formmodells [73] des Brustkorbs zur Segmentierung genutzt. Über Differenzen sukzessiver Tiefenbilder werden zeitliche Veränderungen in der ROI bestimmt (Abbildung 3.2a). Änderungen des Tiefenwerts weisen auf einen Beitrag zur Atmung hin. Diese Änderungen können durch Rauschen jedoch auch zufällig auftreten (Abbildung 3.2b), sodass sie mit dem Abstand zur vorhergehenden ROI gewichtet zur ROI-Detektion im folgenden Bild beitragen (Abbildung 3.2c). Die vorher segmentierte ROI wirkt abschließend adaptiv als Filter für die neue ROI (Abbildung 3.2d). Für eine Berechnung von 4 s Bildaufnahmen werden 10 min benötigt, was einen praktischen Einsatz im realen Alltag unmöglich erscheinen lässt.

¹Veröffentlicht unter der Creative Commons Attribution (CC BY) Lizenz. Die Abbildung wurde für diese Arbeit ins Deutsche übersetzt.

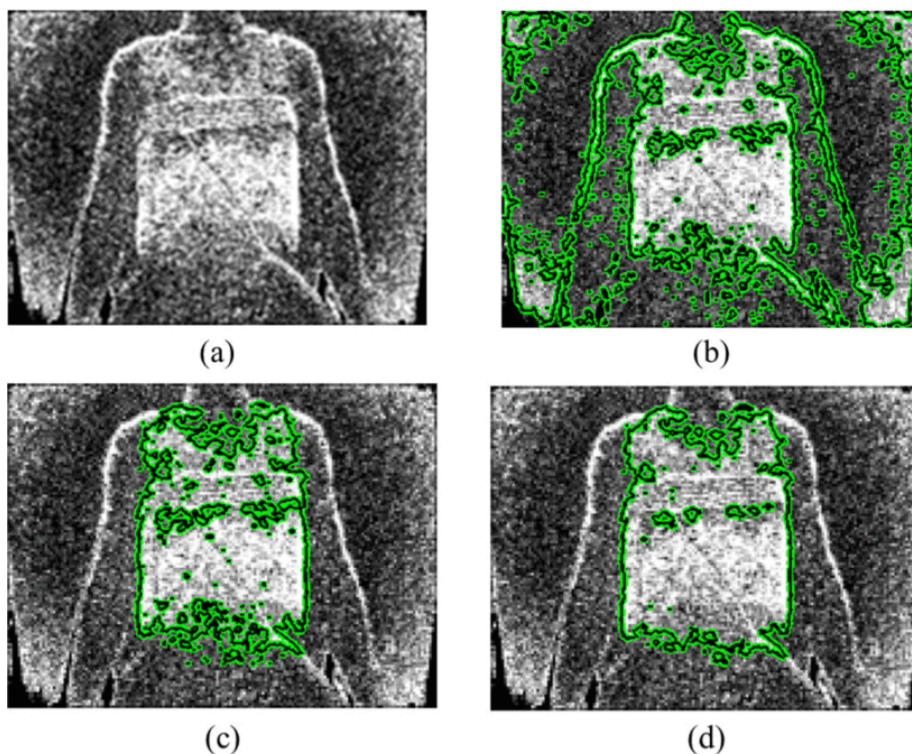


Abbildung 3.2: Auswahl einer Region-of-Interest durch Level-Set-Segmentierung [60] ©2019 IEEE ²: (a) Differenzbild aufeinanderfolgender Tiefenbilder, (b) Ergebnis der Level-Set-Segmentierung, (c) Einbeziehung spatialer Zusammenhänge, (d) weitere Einbeziehung temporaler Informationen durch adaptive Filterung.

Eine weitere automatische Segmentierung wird von Ostadabbas u. a. [61] verwendet. Aus dem Mittelwertbild der gesamten Messung werden über Thresholding Hintergrund- und Vordergrundpixel entfernt. Zusammenhängende Segmente werden zeilen- und spaltenweise gebildet. Abschließend erfolgt eine anatomische Anpassung um die ROI an die Torsohöhe der Probanden anzupassen.

²Abbildungen (a) und (b) aus der Abbildung 5 und Abbildungen (c) und (d) aus Abbildung 6 der Originalveröffentlichung für diese Arbeit fusioniert.

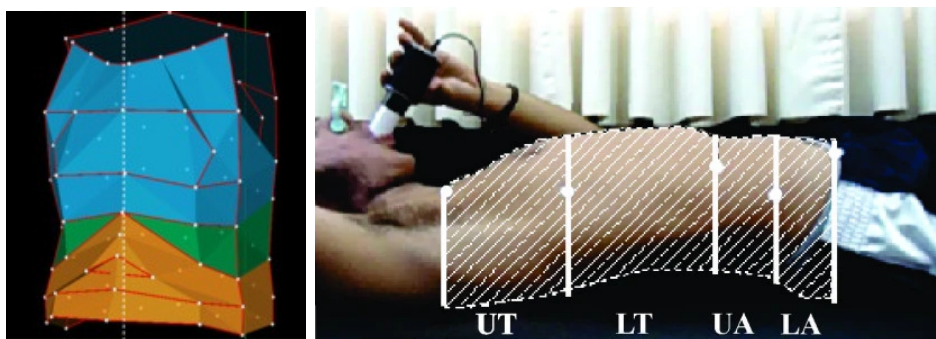
3.1.2 Ausgewählte Regionen

ROIs werden verwendet, um Regionen auszuwählen, die zur Atmung beitragen. Aus diesem Grund ist die Region von Oberkörper bis Kopf von besonderem Interesse. sROIs können eine feinere Unterteilung darstellen und beispielsweise den Beitrag verschiedener Thoraxregionen differenzieren. Die Atemmechanik kann mit der DPG über die zeitliche und örtliche Darstellung von Teilvolumen einzelner sROIs, wie zum Beispiel Thorax und Abdomen, visualisiert werden. sROIs werden vor allem für die opto-elektronische Plethysmographie (OEP, Opto-Electronic Plethysmography) verwendet, wo die Marker an den Regionsgrenzen platziert sind. In der OEP können drei sROIs unterschieden werden: pulmonaler Brustkorb, abdominaler Brustkorb und Bauchraum, siehe Abbildung 3.3a.

Eine andere Unterteilung wird hingegen durch Ripka, Ulbricht und Gewehr [59] vorgenommen. In der lateralen Betrachtung des Thorax erfolgt die Unterteilung nach: oberer Thorax, unterer Thorax, oberes Abdomen und unteres Abdomen, siehe Abbildung 3.3b. Die gleichen sROIs werden in Silvatti u. a. [36] aus einer Ganzkörperansicht verwendet. Seppanen u. a. [74] geben an, zwei sROIs zu verwenden, die als horizontale Streifen am Schwertfortsatz in der Nähe des Nabels definiert sind. Mit Hilfe einer Hauptkomponentenanalyse (PCA, Principle Component Analysis) werden in Soleimani [75] viele gleichmäßig verteilte sROIs zu einem Atmungssignal zusammengefasst, um den Einfluss von Körperbewegungen zu reduzieren. In Arbeiten mit automatischer ROI-Auswahl ohne Marker werden nach derzeitigem Stand keine anderen sROIs verwendet, um die Beteiligung dieser Regionen darzustellen.

Bei der Betrachtung der ROI spielt die Atemmechanik eine entscheidende Rolle. Die Ausdehnung des Thorax erfolgt in vertikaler, transversaler und sagitaler Richtung [76]. Dabei ist die Ausdehnung in anteroposteriorer Richtung am größten, und der Beitrag zur Gesamtänderung des Thoraxvolumens ist am stärksten [22]. Neben der thorakalen Bewegung kann die abdominale Bewegung von größerer Bedeutung sein, insbesondere bei der abdominalen Atmung. Unterschiede in der Interaktion der Kompartimente können auf unterschiedliche Bedingungen zurückzuführen sein. So wurden beispielsweise Unterschiede bei Schwimmern [36] und Tänzern [37] im Vergleich zu untrainierten Vergleichspersonen festgestellt. Es gibt unterschiedliche Daten zur signifikanten Atmungsmechanik zwischen den Geschlechtern. So wurden signifikante Unterschiede in den thorakoab-

dominalen Bewegungen gefunden [77–79], während andere Studien keine derartigen Unterschiede fanden [34, 37]. Auch pathologische Veränderungen können zu Veränderungen der Atmungsmechanik führen, wie zum Beispiel bei Wirbelsäulen-anomalien [80] oder partiellen Lungenfehlern [81], die sich auf die Symmetrie und das Atemmuster auswirken können. Ein Unterschied in der Lungenfunktion zwischen der erkrankten und der kontralateralen Seite des Thorax nach einer Thorakotomie kann mit DPG sichtbar gemacht werden [82]. Zur Untersuchung der Atemfrequenz mit einer Tiefenkamera zeigen Kempfle und van Laerhoven [83], dass der Brustkorb die beste Region zur Extraktion von Atmungssignalen ist. Die Bauchregion hat die geringste Signalqualität.



(a) Zoumot u. a. [84] ³

(b) Ripka, Ulbricht und Gewehr [59] ⁴

Abbildung 3.3: Auswahl von Subregionen für tiefenbasierte Plethysmographie in der Literatur. (a) Markerbasierte Verfahren mit pulmonalen Brustkorb (blau), abdominaler Brustkorb (grün) und Bauchraum (orange). (b) Unterregionen in lateraler Ansicht: Oberer Thorax (UT), unterer Thorax (LT), oberes Abdomen (UA) und unteres Abdomen (LA).

Es werden in der Literatur verschiedene Verfahren zur Auswahl einer ROI eingesetzt. Dabei sind automatische Verfahren den manuellen oder halbautomatischen im realen Einsatz vorzuziehen, da sie kein Mitwirken von Proband oder Bediener erfordern. Häufig werden rechteckige Regionen ausgewählt [61, 63, 71]. Komplexere Formen [60, 66] und sROI [59, 84, 85] finden ihre Begründung in den unterschiedlichen physiologischen Beiträgen der Thoraxregionen. Eine universelle, komplexe Maske, welche auf physiologischen Zusammenhängen basiert und au-

³Veröffentlicht unter der Creative Commons Lizenz. Die Abbildung wurde für diese Arbeit abgeschnitten.

⁴Veröffentlicht unter der Creative Commons Attribution Lizenz. Die Abbildung wurde für diese Arbeit abgeschnitten.

tomatisch ausgewählt werden kann, existiert derzeit nicht und hat das Potential die Qualität des extrahierten Signals weiter zu verbessern. Während eine automatische Auswahl von sROI durch Amplituden [63], vergleiche Abbildung 3.1, die Signalform vernachlässigt, könnten korrelationsbasierte Auswahlverfahren dazu beitragen, auch die Signalform aus der ROI zu erhalten.

3.2 Modelle zur Volumen-Zeit-Bestimmung

Das Messsignal der tiefenbasierten Messung s wird aus der ROI extrahiert. Daher beeinflusst die ROI erheblich die Signalqualität. Je nach Extraktionsart kann s in unterschiedlichen Einheiten vorliegen: als Abstandsänderung des Oberkörpers zur Kamera in mm oder als Volumenänderung des Thorax in l. Ein Modell f , bildet s auf einer Zielfunktion V_{Ziel} ab. Das Ziel ist es, den Fehler zwischen dieser Zielfunktion und einer Grundwahrheit V_{Wahr} durch ein geeignetes Modell und eine geeignete Extraktionsart zu minimieren. Eine solche Grundwahrheit kann beispielsweise durch eine parallele Messung mittels eines Spirometers gewonnen werden.

$$|V_{Wahr} - V_{Ziel}| \rightarrow \min \text{ mit } V_{Ziel} = f(s) \quad (3.1)$$

Von Boudarham u. a. [42] wurde das Messsignal mithilfe eines Bewegungsanalyse-Systems (Motion Analysis, Santa Rosa, Kalifornien) ermittelt. Eine lineare Funktion mit Steigung und Achsenabschnitt wird verwendet, um die Vitalkapazitäten in das Zielsignal zu transformieren. Die Parameter werden auf der Basis aller Messsignale erhoben. Ein alternatives OEP-System wurde von Feitosa u. a. [56] mit BTS Bioengineering (Italien) ohne weitere Skalierung oder mit OptiTrack Prime 17W (2017 NaturalPoint, Inc. USA) [37] eingesetzt.

Von Oh u. a. [60] ergibt sich das Zielsignal aus der Summierung der Differenzen aufeinanderfolgender Bilder. Die Pixeleinheit wird mit Hilfe einer linearen Funktion in die tatsächliche Pixellänge umgerechnet. Die sich daraus ergebenden Koeffizienten beschreiben das Verhältnis zwischen Entfernung und tatsächlicher Pixellänge. Ostadabbas u. a. [61] wird ebenfalls eine lineare Skalierung zwischen dem Messsignal und dem Zielterminal verwendet. Die Koeffizienten ergeben sich aus einer Referenzmessung mit bekanntem Volumen. Es wird angenommen, dass die Koeffizienten für ein bestimmtes Subjekt konstant sind. Eine andere Methode

zur Kalibrierung einer linearen Transformation wird in [86] verwendet. Für die Kalibrierung werden zufällige 50 % der Testpunkte einer Messung ausgewählt. Die Kalibration erfolgt der Grundwahrheit des Spirometers mit dem Messsignal.

Der Ansatz von Soleimani u. a. [71], Sharp u. a. [80] und Soleimani u. a. [87] basiert auf mehreren Skalierungsfaktoren, die zum einen für das Tidalvolumen, aber auch separat für die Vitalkapazitätsbestimmung angewendet werden. Aus der 3D-Punktwolke des Oberkörpers wird ein Netz erzeugt und das eingeschlossene Volumen in Bezug auf eine ebene Referenzfläche bestimmt. Aufgrund der Skalierungsfaktoren ist die Lage der Referenzfläche unerheblich. Ein Butterworth-Filter vierter Ordnung wird verwendet, um die Überglättung eines Filters mit gleitendem Mittelwert zu reduzieren. Eine weitere Filterung mit einer Grenzfrequenz von 1 Hz für das Tidalvolumen und 3 Hz für die Vitalkapazität wird angewendet. Aus dem Messsignal werden Keypoints extrahiert, die auf lokale Minima und Maxima des Messsignals ausgerichtet sind. In einer Trainingsphase werden die idealen Skalierungsfaktoren anhand eines Spirometersignals ermittelt. In der Testphase werden die Skalierungsfaktoren so gewählt, dass aus den Trainingsdaten die am besten übereinstimmenden Skalierungsfaktoren mit dem geringsten Fehler ausgewählt werden. Als Fehlermaß dienen die Differenzen der Keypoints. Training und Test der Skalierungsfaktoren werden patientenindividuell durchgeführt. Es werden bisherige Aufnahmen eines Patienten zur Bewertung neuer Aufnahmen berücksichtigt. Bei Verwendung einer zweiten Kamera wird das Messsignal als Differenz zwischen dem Messsignal der Frontkamera und dem Messsignal der Rückkamera berechnet. So wird der gesamte Oberkörper erfasst und Artefakte, die durch Bewegungen entstehen, werden reduziert [20]. In einer Weiterentwicklung verwendet die Gruppe die PCA zur Reduzierung von Bewegungsartefakten [85]. Die Hauptkomponente entspricht einem bewegungsbereinigten Signal. Die empirische Zerlegung in Moden (EMD, Empirical Mode Decomposition) [88] kann zur Detrendierung von Daten verwendet werden [20].

Über einen multiplen Regressionstest fanden Ripka, Ulbricht und Gewehr [59] zusätzliche Prädiktoren zur Transformation der Atemvolumina. Für die Vitalkapazität wird die Körpergröße des Patienten ebenfalls mit einem linearen Skalierungsfaktor versehen. Takamoto u. a. [43] verwenden ebenfalls eine multiple lineare Regression und mit der Körpergröße und dem Körpermasseindex (BMI, Body-Mass-Index) auch somatotypische Faktoren.

Der Einsatz von linearen Modellen ist zurückzuführen auf den Zusammenhang von Messsignal und physiologischer Bedeutung. Bewegungen, die während eines forcierten Atemmanövers durchgeführt werden, beeinflussen jedoch die Signalqualität, sodass für Ruheatmung und forciertes Atmung mit unterschiedlichen Skalierungsfaktoren gearbeitet werden muss [87]. Alternativ verwenden weitere Modelle bereits somatotypische Merkmale zur Bestimmung der Atemvolumina [43, 59]. KNN als universelle Funktionsapproximatoren, sowie weitere ML Verfahren, könnten als nicht-lineare Modelle potentiell diese Zusammenhänge besser berücksichtigen und zu genaueren Messergebnissen führen.

3.3 Genauigkeit und Evaluierung der Respirationsparametermessung

Für eine Bewertung der DPG-Verfahren sind auf der einen Seite die Genauigkeit der Methoden, aber auch die dabei verwendete Evaluationsmethode von Bedeutung. Die Genauigkeit von DPG-Messungen wird mit Goldstandard-Referenzgeräten verglichen. Zu diesem Zweck atmen Patienten zusätzlich zur kontaktlosen Messung durch zum Beispiel ein Spirometer. Abweichungen können hinsichtlich der Signalunterschiede, oder anhand der daraus abgeleiteten Respirationsparameter angegeben werden.

Es ist jedoch zu unterscheiden, mit welchen Daten die Evaluation erfolgte. Wie in Abschnitt 3.2 dargestellt, werden Modelle verwendet, um eine Transformation der Messdaten in Volumen-Zeit-Signale vorzunehmen. Die Kalibration dieser Modelle kann wie folgt differenziert werden, mit entsprechenden Beispielen in Tabelle 3.1 und dargestellt in Abbildung 3.4:

- keine: es wird keine Transformation der Messdaten vorgenommen,
- Gesamt: das Modell wurde mit dem vollständigen Datensatz erstellt,
- Proband: das Modell verwendet Messdaten desselben Probanden oder
- Messung: das Modell verwendet Testpunkte der gleichen Messung.

Bei der Kalibration anhand des Probanden werden zur Erstellung des Modells andere Messungen des selben Probanden zur Erstellung des Modells verwendet. Die Evaluation anhand des gesamten Datensatzes sieht vor, dass ein Modell anhand aller Messungen aufgestellt und anhand aller Messungen charakterisiert wird. Bei der Verwendung eines Test-Datensatzes wird ein oder mehrere Probanden dem

Modell bei der Erstellung vorenthalten. Ausschließlich diese vorenthaltenen Daten werden dann für die Evaluation verwendet. Für eine Differenzierung von der Probanden-Kalibrierung ist es notwendig, dass nicht nur einzelne Messungen eines Probanden, sondern gesamte Probanden als Test-Datensatz verwendet werden. In der Kreuzvalidierung mit Auslassung (LOOCV, Leave-One-Out Cross Validation) wird auch ein einzelner Proband als Test-Datensatz vorenthalten und das Modell bewertet. Anschließend wechselt jedoch der Proband und das Modell wird mit den selben Hyperparametern neu erstellt und für den nächsten Probanden als Test-Datensatz neu evaluiert. Dies wiederholt sich für alle Probanden.

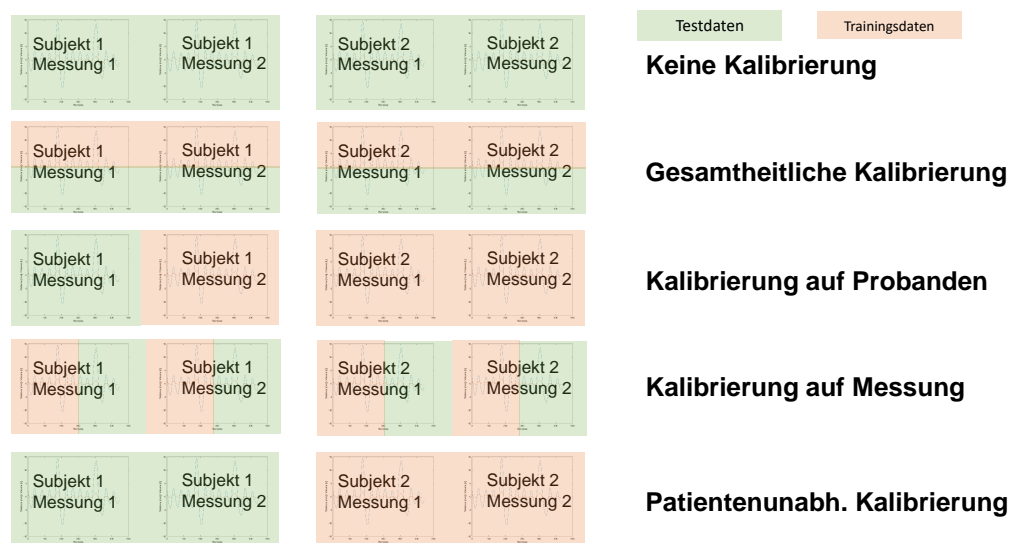


Abbildung 3.4: Möglichkeiten zur Kalibrierung von DPG. Aufteilung von Trainings- und Testdaten für diverse Kalibrierungstechniken. Horizontale Trennung einer Messung bedeutet, dass diese vollständig im Training- und Testdatensatz enthalten ist. Eine vertikale Teilung bedeutet, dass sich ein Teil dieser Messung im Trainings- und der andere Teil im Testdatensatz befindet.

Die Genauigkeit von Modellen hinsichtlich Mittelwert \pm Standardabweichung und der gewählten Validierungsmethode ist in Tabelle 3.1 dargestellt. Oh u. a. [60] stellen ein System vor, das keine eigene Anpassung an die 10 gesunden Probanden erfordert. Die Level-Set-Methode ergibt einen mittleren Fehler von $8,41 \pm 2,16$ % des Tidalvolumens, bereinigt um zwei Probanden mit stärkeren Ausreißern aufgrund von Beatmungslecks oder Körperbewegungen. Die Patienten erhielten das Luftvolumen während der Messungen über ein Beatmungsgerät.

Die Bestimmung des Atemwegswiderstandes von Ostadabbas u. a. [61] erfolgt mit einer Kincet-Kamera zur Bestimmung des Tidalvolumens. Es ergibt sich eine mittlere Abweichung von 70 ± 60 ml Tidalvolumen. Die 14 Probanden wurden gebeten, sich während der Atmung nicht zu bewegen. Die Referenzwerte sind dabei jedoch nicht simultan zur DPG aufgenommen wurden, sondern entstammen vorheriger Messungen von zum Teil über einem Jahr.

Tabelle 3.1: Messgenauigkeiten und Kalibrierungsstrategien in der Literatur für Tidalvolumen (TV) und Vitalkapazität (VC). Die Unterscheidung erfolgt von Modellen, Kalibrierung (Kalib.) und der Anzahl der Probanden (Prob.), respektive gesunden Probanden (ges.).

Quelle	Modell	Kalib.	Messabweichung	Prob. (ges.)
[60]	linear	keine	TV: mittel: $8,41 \pm 2,16\%$	10 (10)
[61]	linear	keine	TV: mittel: 70 ± 60 ml	14 (14)
[86]	linear	Proband	TV: RMSE: 182 ± 107 ml	50 (50)
[80]	linear	Gesamt	VC: mittel: 16 ± 51 ml	50 (50)
[62]	linear	Proband	VC: mittel: -150 ± 842 ml TV: mittel: 100 ± 255 ml	35 (35)
[71]	linear	Messung	VC: mittel: 9 ± 39 ml TV: mittel: 74 ± 88 ml	40 (0)
[43]	linear	keine	VC: mittel: 57 ± 716 ml	53 (21)
[74]	nicht linear	Proband	TV: $9,4 \pm 8,4\%$	8 (8)
[76]	nicht linear	Proband	TV: 5,8 % bis 7,8 %	4 (4)
[63]	linear	Gesamt	TV: mittel: 10,7 % bis 15,5 %	39 (39)

Die Methode von Reyes u. a. [86] unter Verwendung eines Smartphones resultiert in einen Root Mean Square Error (RMSE) für das Tidalvolumen von 182 ± 107 ml. Zur Kalibrierung des linearen Systems wurden die Daten mit einem Trainingsdatensatz derselben Versuchsperson trainiert.

Das System von Sharp u. a. [80] weist einen mittleren Fehler sowie den prozentualen Anteil der Messungen außerhalb eines Bereichs von ± 150 ml für die

forcierte Vitalkapazität (19 ml, 0% außerhalb ± 150 ml), das forcierte expiratorische Volumen in einer Sekunde (82 ml, 61,9%), die Vitalkapazität (16 ml, 4,8%) und die Inspirationskapazität (23 ml, 6%) auf. Zu diesem Zweck wurden 100 Probanden, inklusive COPD- und Asthma-Patienten, mit dem Ein-Kamera-System getestet. Das Volumen-Zeit-Signal wurde anhand von charakteristischen Merkmalen des Signals linear regressiv approximiert. Zu diesen Signalmerkmalen gehören die Extrempunkte des Messsignals. Die Patienten wurden aufgefordert, sich nicht zu bewegen. Soleimani u. a. [20] vergleichen einen Ansatz mit mehreren Kameras mit einer Methode mit nur einer Kamera. Mittelwert und Standardabweichung werden für forcierte Vitalkapazitätsmessungen und langsame Vitalkapazitätsmessungen angegeben. So werden Mittelwert \pm Standardabweichung u. a. für die Vitalkapazität (-150 ± 842 ml) und das Tidalvolumen (100 ± 255 ml) angegeben, wobei die Hinzunahme einer weiteren Kamera den Fehler in jedem Fall signifikant reduziert. Wie bei anderen Messungen von Soleimani erfolgt die Kalibrierung des Systems durch LOOCV anhand von Messungen derselben Person. Durch die zusätzliche Unterdrückung aktiver Körperbewegungen während der Atmung kann der Fehler weiter reduziert werden [85]. Die angegebenen Fehlerwerte sind jedoch normalisiert und daher nicht direkt mit anderen Methoden vergleichbar. Die anderen Datensätze [62, 71] dieser Arbeitsgruppe schränken die aktive Bewegung des Patienten während der Atemaufzeichnungen ein, so dass weniger Fehler auftreten [20].

Das System zur Bestimmung des Atemvolumens von Takamoto u. a. [43] berücksichtigt im realen Einsatz die Körpergröße und den BMI des Patienten. Das Ziel, COPD über VC und FEV₁ zu erkennen, wird mit 81% Sensitivität und 90% Spezifität erreicht. Die Fehler bei der Abgrenzung von VC und FEV₁ wurden in einer weiteren Messung wiederholt, um die Reproduzierbarkeit zu überprüfen.

Durch Seppanen u. a. [74] werden zwei sROIs miteinander kombiniert. Über kleinste Quadrate werden in einer Testmessung FIR-Filterkoeffizienten bestimmt, die das Messsignal der sROIs nichtlinear zum Zielsignal weiterverarbeiten. Es wird ein absoluter Messfehler von $9,4 \pm 8,4\%$ angegeben.

Transue u. a. [76] schlagen ein Modell vor, das auf einer Iso-Oberflächen-Rekonstruktion basiert. Die Bewegung des Brustkorbs wird omnidirektional berücksichtigt und ein 3D-Modell mittels Surface-Hole-Filling erstellt. Mit Hilfe eines Bayes'schen neuronalen Netzes und der Volumenänderung des 3D-Modells

als Input wird das Atemvolumen approximiert. Die Abweichungen vom Spirometersignal werden mit 5,8 % bis 7,8 % angegeben.

Ein System, welches von Imano u. a. [63] vorgeschlagen wurde, wurde an 39 älteren Menschen getestet. Es wurde ein Vergleich zwischen bekleideten und unbekleideten Probanden sowie Männern und Frauen durchgeführt, wobei die mittleren absoluten relativen Fehler zwischen 10,7 % bis 15,5 % lagen. Nach einer individuellen, automatischen ROI-Auswahl wird das Tidalvolumen mittels linearer Regression bestimmt.

Die Bewertung der Modelle auf Daten der selben Messung oder des gesamten Datensatzes lassen keine Rückschlüsse auf die Anwendung auf neuen Daten. Auch eine Kalibration auf bestehende Probanden ist nicht alltagstauglich, da hierfür erneut ein Spirometer benötigt wird. Modelle, die diese Anforderungen berücksichtigen und mit gleichzeitig aufgenommen Spirometerdaten validieren, weisen erwartungsgemäß höhere Messfehler auf [43, 60]. Gleichzeitig ist eine solche Evaluation notwendig, für die Etablierung eines kalibrationsfreien Modells zur Bestimmung der Respirationsparameter.

3.4 Klassifikation von Respiratorischen Erkrankungen

In der klinischen Praxis kann eine Bewertung der Atemdynamik manuell durch Abtastung mittels manuelle Beurteilung der Atembewegung (MARM, manual assessment of respiratory motion) erfolgen [89]. Die Anwendung von DPG zur Messung der thorakalen und abdominalen Atemmechanik wurde bereits in der Literatur untersucht [82, 90, 91]. So sind in freier Ruheatmung Inspirationszeit und thorakale Asynchronität signifikant verschieden zwischen COPD-Patienten und einer nach Alter, Größe, Geschlecht und Gewicht gematchten gesunden Kontrollgruppe [92]. Studien mit OEP an COPD-Patienten im Vergleich zu gesunden Kontrollgruppen zeigen keine signifikante Änderung des Tidalvolumens, aber unterschiedliche Beiträge von Rippen und Abdomen [91]. Dies ist abhängig von der Körperlage der Patienten. Eine stärkere Asynchronität liegt im Sitzen vor. COPD-Patienten weisen im Vergleich zu Gesunden einen größeren Beitrag des abdominalen Kompartiments zum Tidalvolumen auf. Diese paradoxe Atemmechanik tritt jedoch nicht bei allen COPD Patienten auf [93].

3.5 Zusammenfassung und Bewertung des Stands der Forschung

DPG ist eine neuartige Methode zur Bestimmung von respiratorischen Parametern. Der erste Schritt besteht darin, eine ROI auszuwählen. Die Auswahl einer ROI erfolgt dabei meist statisch als konstante Position über die gesamte Messung. Eine solche rechteckige Auswahl lässt sich jedoch nicht physiologisch oder signaltechnisch begründen. Ansätze, die eine ROI aus Signalamplituden ableiten, basieren ebenfalls auf rechteckigen Regionen und vernachlässigen die Signalform. Komplexere Ansätze, welche unter hohem Rechenaufwand spatiale Zusammenhänge einbeziehen, müssen individuell für jede Messung angefertigt und wurden nicht hinsichtlich des Informationsgehalt der ausgewählten Regionen untersucht. Da jedoch gezeigt wurde, dass nicht alle Thoraxregionen im gleichen Maß zum Respirationssignal beitragen, lassen sich diese Unterschiede durch statistische Zusammenhänge mit dem Spirometersignal genauer untersuchen. Es lässt sich somit postulieren, dass diese Erkenntnisse zu einer auswählbaren ROI führen, welche einer rechteckigen ROI hinsichtlich der extrahierten Signaleigenschaften überlegen ist.

Die Weiterverarbeitung der Messsignale kann mit rechenintensiven Modellen erfolgen, die jedoch in der Praxis einer einfachen Mittelwert- oder Medianberechnung nicht überlegen sind. Insbesondere für die Filterung von überlagerten Bewegungsdaten sind weitere Signalverarbeitungsschritte notwendig. Möglichkeiten sind die Verwendung von Referenzbereichen, die von diesen Bewegungen nicht betroffen sind, oder die Hauptkomponenten-Zerlegung und die generelle Überwachung mit mehreren Kameras.

Um schließlich aus dem gemessenen Signal auf das Atemvolumen schließen zu können, ist eine Transformation des gemessenen Signals erforderlich. Dies geschieht in der Regel mit Hilfe linearer Skalierungsfaktoren. Es hat sich jedoch gezeigt, dass insbesondere der Übergang zu forcierter Atmung nichtlinear ist und auch somatotypische Faktoren einen Einfluss aufweisen. Dennoch können lineare Skalierungsfaktoren über Signalmerkmale ausgewählt werden. In der Zukunft können nichtlineare Modelle und maschinelles Lernen eingesetzt werden, um die Atmungsparameter anhand von extrahierten Merkmalen des Messsignals und anderen patientenspezifischen Merkmalen zu approximieren. Eine Evaluierung der

auszuwählender Merkmale und der Modelle sollte mittels LOOCV erfolgen, so dass die Unabhängigkeit der Kalibrierung des Modells für den einzelnen Probanden nachgewiesen werden kann.

Krankheiten wie COPD bewirken Veränderungen der Atemmechanik. Über die Modellierung von Beiträgen unterschiedlicher Thoraxregionen zum Gesamtvolumen kann die DPG die Atemmechanik darstellen. Es wird hypothetisiert, dass diese spatio-temporalen Informationen eine Klassifikation von respiratorischen Erkrankungen ermöglichen. Die im Speziellen zu verwendenden Datenmodalitäten und Darstellungen, sind genauer zu untersuchen.

Heutige Systeme zeigen, dass sie bereits mit einer solchen Kalibrierung oder mehreren Kameras akzeptable Ergebnisse erzielen [43, 60], auch wenn die geforderten Genauigkeiten eines klinischen Spirometers nicht erreicht werden. Spirometer selbst dürfen nach den Richtlinien der American Thoracic Society und der European Respiratory Society [9] einen maximalen Fehler von 2,5 % aufweisen. Eine solche Genauigkeit kann jedoch für ein berührungsloses System nicht erwartet werden. Um jedoch die Möglichkeiten und Vorteile von DPG voll auszuschöpfen, müssen zukünftige Modelle als kalibrierungsfreie Einzelkamarasysteme konzipiert werden.

3.6 Abgeleitete Arbeitshypothesen für diese Arbeit

Aus diesen bisherigen Entwicklungen ergeben sich offene Fragestellungen, die in dieser Arbeit beantwortet werden sollen. Diese basieren auf den folgenden Hypothesen:

1. Statistische Ansätze führen zu einer überlegenen Maske für die Messsignal-
extraktion in der DPG im Vergleich zu einer konventionellen, rechteckigen
ROI.
2. Der Einsatz von KI mit der Berücksichtigung von somatotypischen Merk-
malen eignet sich dazu, eine kalibrationsfreie DPG zu etablieren.
3. Die durch die DPG verfügbaren spatio-temporalen Daten sind ausreichend,
um eine Klassifikation von respiratorischen Erkrankungen vorzunehmen.

Kapitel 4

Maschinelles Lernen zur Bestimmung von Atemvolumina

Wie im vorherigen Kapitel 3 gezeigt, bieten nicht-lineare Modelle ein weitgehend ungenutztes Potential in der Bestimmung der Respirationsparameter. Ohne Fixation des Rückens der Probanden, kann die willkürliche Bewegung des Oberkörpers die Thoraxbewegung superpositionieren. Dieser Störanteil muss für eine genauere Bestimmung der Respirationsparameter unterdrückt werden. Für die Aufteilung in TV und VC zeigt sich, dass eine unabhängige Betrachtung sinnvoll ist. In der bisherigen Literatur werden zur Transformation der Messsignale sogenannte Skalierungsfaktor (SF) genutzt. Eine große Varianz in diesen SF kann über persönliche Merkmale erklärt werden. Die Modellierung des Skalierungsfaktors kann dabei über Signalmerkmale erfolgen, welche anhand bestehender Messung desselben Probanden kalibriert werden.

Basierend auf diesen Betrachtungen des bisherigen Stands der Forschung, wird in diesem Kapitel der Einsatz von nicht-linearen Maschinelles Lernen (ML) Verfahren zur Bestimmung der Atemvolumina bewertet. Es werden Merkmale aus dem Signal extrahiert und anschließend auf ihre Eignung mit Merkmalsselektionsverfahren getestet. Der Einsatz von verschiedenen ML-Modellen wird analysiert.

Die Untersuchungen in diesem Kapitel dienen zur Beantwortung der Fragestellung: Wie genau können Respirationsparameter aus Merkmalen des Tiefen-Zeit-Signals bestimmt werden? Darüber hinaus ergeben sich zwei Nebenfragestellungen:

- Kann EL bestehend aus den einzelnen Modellen die Genauigkeit der Respirationsparameterbestimmung erhöhen?
- Welchen Einfluss haben phänotypische Merkmale der Probanden auf die Genauigkeit der Respirationsparameterbestimmung?

4.1 Methodik

In diesem Abschnitt wird die Vorgehensweise zur direkten Bestimmung von TV und VC über ML vorgestellt. Aus dem gesamten Prozess bei der Anwendung von DPG, siehe oben Abbildung 2.4, bestehend aus: Bildaufnahme, Auswahl einer ROI, Signalextraktion, Signalvorverarbeitung und Transformations-Modell, wird in diesem Abschnitt der letztgenannte Aspekt genauer untersucht. Im Vergleich zur Literatur auf selber Datenbasis [75, 94] werden die anderen Verarbeitungsschritte konstant gehalten, sodass die Auswirkungen der Modellveränderungen untersucht werden können.

Als nicht-lineare Modelle werden in dieser Arbeit Verfahren des überwachten ML eingesetzt. Nach der Signalverarbeitung werden Zeitbereichsmerkmale, morphologische Merkmale, statistische Merkmale, Frequenzmerkmale, Zeit-Frequenzbereichsmerkmale, bildbasierte Merkmale und persönliche Merkmale extrahiert. So ist ein großes Spektrum abgedeckt und sichergestellt, dass alle relevanten Domänen adressiert sind. Aus dieser Vielzahl an Merkmalen werden über Merkmalsselektionsverfahren die geeignetsten ausgewählt. Dabei werden Embedded, Wrapper sowie Filter-Methoden zur Selektion eingesetzt. Die Reduktion der Merkmale bietet den Vorteil von Verallgemeinerungsfähigkeit, erhöhter Lerngeschwindigkeit und reduzierter Komplexität der Modelle [48, 95].

Die somit ausgewählten Merkmale bieten den Input für die ML-Modelle. Zur regressiven Approximation der Respirationsparameter werden RF, Support Vector Regression (SVR), KNN und EL eingesetzt und miteinander verglichen. Die Hyperparameter werden optimiert und die Ergebnisse anhand des öffentlichen Datensatzes von Soleimani, Mirmehdi und Damen [94] mit LOOCV gegenübergestellt. Iterativ dient ein Proband als Testdatensatz, während die anderen Probanden den Trainingsdatensatz für die ML-Verfahren darstellen. Das beschriebene Vorgehen der DPG in diesem Kapitel ist übersichtsweise in Abbildung 4.1 dargestellt.

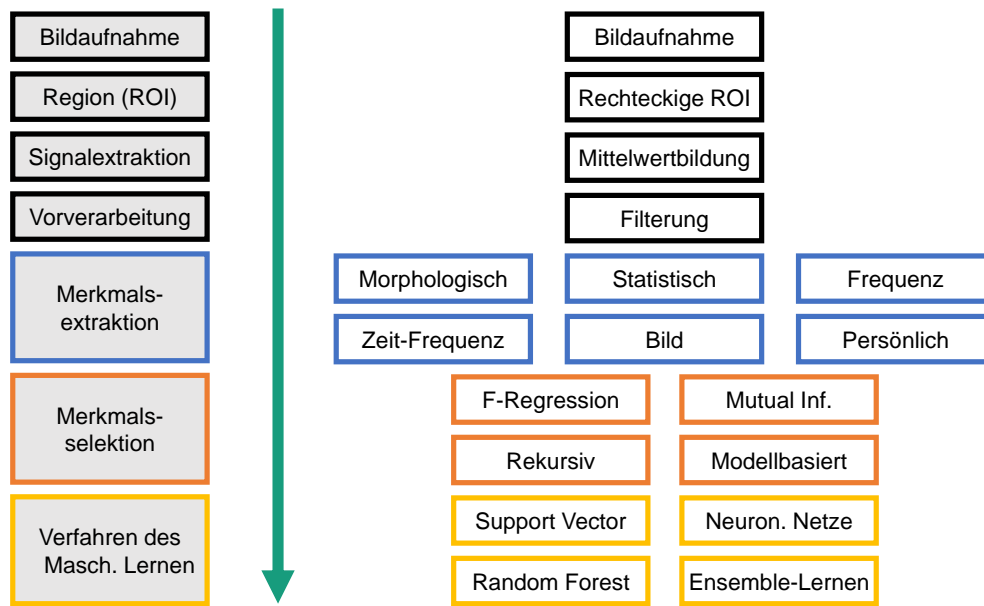


Abbildung 4.1: Vorgehensweise zur Untersuchung der merkmalsbasierten Bestimmung von Respirationsvolumina mittels Maschinellem Lernen. Links befinden sich in sequenzieller Reihenfolge die allgemeinen Schritte der Tiefenplethysmographie mit den speziellen Schritten in dieser Arbeit rechts. Farblich hervorgehoben sind die Veränderungen im Vorgehen im Vergleich zu Soleimani [75].

4.1.1 ROI, Signalextraktion und Signalvorverarbeitung

Die Signalextraktion basiert auf den detektierten Körperpunkten der Microsoft Kinect Kamera (TOF-Prinzip). Es wird eine ROI auf dem Thorax definiert, identisch mit der Arbeit von Soleimani [75]. Ausgehend von den vier Positionen rechte Schulter (rS), linke Schulter (lS), Wirbelsäule Mitte (WM) und Wirbelsäule Schulter (WS) werden die jeweiligen x und y Positionen für die ersten 100 Bilder der Messung gemittelt, um \bar{x} und \bar{y} zu erhalten. Breite b und Höhe h der ROI werden wie folgt definiert:

$$b = \frac{2}{3} \cdot (\bar{x}_{lS} - \bar{x}_{rS}) \quad (4.1a)$$

$$h = \frac{3}{2} \cdot (\bar{y}_{WS} - \bar{y}_{WM}) \quad (4.1b)$$

$$x_{ROI} = \bar{x}_{WS} - \frac{b}{2} \quad (4.1c)$$

$$y_{ROI} = \max(\bar{y}_{lS}, \bar{y}_{rS}) \quad (4.1d)$$

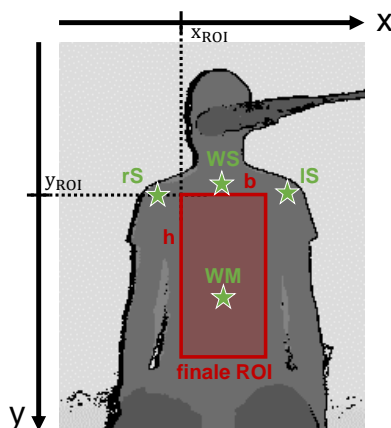


Abbildung 4.2: Festlegung anhand der Region of Interest (ROI) mithilfe der Körperpunkte rechte Schulter (rS), linke Schulter (lS), Wirbelsäule Schulter (WS) und Wirbelsäule Mitte (WM). Diese Punkte werden durch die Kinect Kamera identifiziert und daraufhin die finale ROI bestimmt.

Der Mittelwert der ROI wird zu jedem Zeitpunkt t_i für i von 0 bis N mit der Anzahl der Aufnahmen N gebildet. Somit entsteht das Messsignal s in Gleichung (4.2).

$$s[t] = \frac{1}{b \cdot h} \sum_{x=x_{ROI}}^{x_{ROI}+b} \sum_{y=y_{ROI}}^{y_{ROI}+h} I[x, y, t] \quad (4.2)$$

Das Messsignal wird mit einem Butterworth-Tiefpass-Filter dritter Ordnung gefiltert, sodass s_{filt} entsteht. Die Grenzfrequenz liegt bei 1.5 Hz in Anlehnung an bisherige Arbeiten [75]. Die Filterung bewirkt ein Unterdrücken des Rauschanteils.

4.1.2 Merkmalsextraktion

Aus dem vorliegenden Signal werden Charakteristika extrahiert, welche das Signal eindeutig beschreiben. Dedizierte Merkmale werden separat für die normale Atmung (TV) und maximale Atmung (SVC) extrahiert. Diese Differenzierung ergibt sich aus den Ergebnissen aus Soleimani u. a. [71]. Wie bereits in Abschnitt 3.2 beschrieben, werden bisher SF verwendet, welche das Messsignal auf das Spirometersignal transformieren. Dabei sind die Ergebnisse besser, wenn zwei Skalierungsfaktoren verwendet werden: für die normale Atmung und die maximale,

forcierte Atmung. Als Ursachen werden Oberkörperbewegungen benannt, die mit der maximalen Anstrengung bei der Atmung einhergehen und somit das Messsignal der reinen Thoraxbewegung superpositionieren. Für eine Übersicht der folgenden Beschreibung der Parameter sei auf Tabelle A.1 verwiesen.

Unterteilung des Messsignals

Über die Peak-Prominence Methode [96] werden die lokalen Extremstellen gefunden, welche die Umkehr des Volumenstroms kennzeichnen. Im Folgenden bezeichnet E_i die Punkte vollständiger Einatmung und A_i die Punkte vollständiger Ausatmung. Der Index i erstreckt sich von 1 bis zur Anzahl N der jeweiligen Punkte. Diese Extremstellen sind in Abbildung 4.3 dargestellt. Zur Abgrenzung des Signals ist der Index j gesucht, für den die Bewegung des Brustkorbs zwischen vorheriger Aus- und Einatmung maximal ist:

$$|E_j - A_j| \rightarrow \max, \quad \text{für } j \in [1, \dots, N] \quad (4.3)$$

Das gefilterte Signal s_{filt} wird in den Bereich der normalen Atmung s_{TV} und maximaler Atmung s_{SVC} separiert, anhand des Zeitpunkts der letzten normalen Einatmung $t_{E_{j-1}}$:

$$s_{\text{TV}}[t] = s_{\text{filt}}[t], \quad \text{für } t \leq t_{E_{j-1}} \quad (4.4a)$$

$$s_{\text{SVC}}[t] = s_{\text{filt}}[t], \quad \text{für } t \geq t_{E_{j-1}} \quad (4.4b)$$

Merkmale der Ruheatmung

Für die Ruheatmung werden insgesamt 52 Merkmale F_{TV} aus dem Signal s_{TV} extrahiert. Zur Unterdrückung von Bewegungsartefakten wird die EMD [88] auf das Signal angewendet. Die EMD ist eine Methode zur Signalzerlegung. Nicht-lineare und nicht-stationäre Signale können mit ihr in einzelne Komponenten zerlegt werden. Die Zerlegung in verschiedene Oszillationen des Signals entspricht dem gewünschten Einsatzbereich, der Entfernung der langsameren Oberkörperbewegung. Dementsprechend wird allein die erste Komponente (erste intrinsische Modenfunktion, IMF_1) berücksichtigt.

$$s_{\text{TV,EMD}}[t] = \text{IMF}_1 \quad (4.5)$$

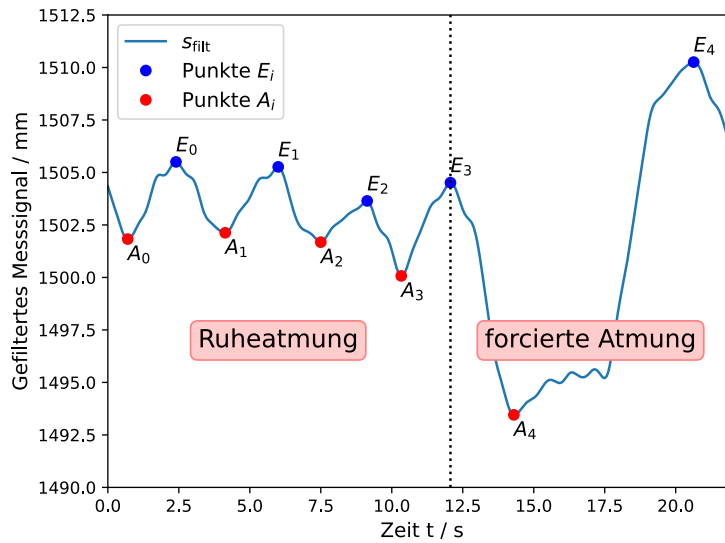


Abbildung 4.3: Aufteilung des Messsignals in die Bereiche der Ruheatmung und forciertes Atmung. Über die Peak-Prominence gefundenen Extremstellen für die Einatmung E und die Ausatmung A wird das Messsignal beim letzten ruhigen Atemzug E_3 untergliedert.

Bereits in der Literatur entwickelte Merkmale [62] werden ebenfalls für diese Arbeit berechnet. Diese Merkmale F ergeben sich aus den ersten sechs auftretenden Punkten der Volumenstromumkehr, siehe erneut Abbildung 4.3. Dies ist für die Punkte der Inspiration in den Gleichungen (4.6a) und (4.6b) dargestellt und gilt für die Expiration gleichermaßen.

$$F_{\text{TV},A_i} = A_i, \quad \text{für } i = 0, \dots, 2 \quad (4.6a)$$

$$F_{\text{TV},t_{A_i}} = t_{A_i}, \quad \text{für } i = 0, \dots, 2 \quad (4.6b)$$

Weitere Merkmale der Signalmorphologie ergeben sich aus dem Mittelwert der einzelnen Tidalvolumina (Gleichung (4.7a)) und der Umgebung der Inspirationspeaks, siehe Gleichungen (4.7b) bis (4.7c).

$$F_{\text{TV},\text{MW}} = \frac{1}{3} \sum_{i=1}^{N=3} s[t_{e_i}] - s[t_{a_i}] \quad (4.7a)$$

$$F_{\text{TV},U_{i,+}} = s_{\text{TV,EMD}}[t_{E_{N-i}} + 2], \quad \text{für } i = 0, \dots, 2 \quad (4.7b)$$

$$F_{\text{TV},U_{i,-}} = s_{\text{TV,EMD}}[t_{E_{N-i}} - 2], \quad \text{für } i = 0, \dots, 2 \quad (4.7c)$$

Neben dem Abstand von zwei Samples um die Inspirationspeaks, wird die weitere Umgebung berücksichtigt. So wird die euklidische Distanz summiert zwischen den Inspirationspeaks und dem Messsignal eine Sekunde zuvor $F_{\text{TV,euklid,-}}$, respektive danach $F_{\text{TV,euklid,+}}$. Das Verhältnis dieser Summen wird als weiteres Merkmal $F_{\text{TV,euklid,ratio}}$ gebildet. Diese Merkmale zur Charakterisierung der Peak-Umgebungen, sollen für eine genaue Anpassung sorgen. Ähnliche Merkmale sind bei anderen physiologischen Signalen wie der merkmalsbasierten Blutdruckbestimmung über Photoplethysmographie von Bedeutung [97].

Der größte positive und negative Anstieg im Signal wird zusammen mit dem Zeitpunkt als Merkmal $F_{\text{TV,Anstieg}}$ verwendet. Dies korreliert mit dem größten Fluss des Volumenstrom und kann hinsichtlich der Atemanstrengung von Bedeutung sein. Ähnliche Parameter werden in der klinischen Diagnostik verwendet, siehe zum Beispiel FEV_1 .

Ebenfalls auf den Volumenfluss beziehen sich die Merkmale der Ein- und Ausatmung. Hierfür wird der Mittelwert des Anstiegs von Ein- zur Ausatmung, respektive Aus- zur Einatmung- als Merkmal $F_{\text{TV,V,e}}$ verwendet, siehe Gleichung (4.8).

$$F_{\text{TV,V,e}} = \frac{1}{3} \sum_{i=1}^{N=3} \frac{s[t_{e_i}] - s[t_{a_i}]}{t_{e_i} - t_{a_i}} \quad (4.8)$$

Zusätzlich werden aus der partiellen Ableitung $\frac{\partial s_{\text{TV,EMD}}}{t}$ die stärksten positiven und negativen Anstiege sowie deren Position als Merkmal $F_{\text{TV},\delta}$ extrahiert.

Aus den Extremwerten E_i werden Mittelwert $F_{\text{TV,AF,MW,E}}$ und Standardabweichung $F_{\text{TV,AF,STD,E}}$ der Abstände aufeinanderfolgender Extrempunkte gebildet. Dieses Verhalten ist in den Gleichungen (4.9a) und (4.9b) für die Einatmung dargestellt und verhält sich für die Ausatmung mit den Extremstellen A_i gleich. Diese zeitlichen Abstände der Peaks können als AF aufgefasst werden, welche ebenfalls ein wichtiger Atemparameter im klinischen Alltag darstellt.

$$F_{\text{TV,AF,MW,E}} = \frac{1}{N-1} \sum_{i=2}^N (s[t_{E_i}] - s[t_{E_{i-1}}]) \quad (4.9a)$$

$$F_{\text{TV,AF,STD,E}} = \sqrt{\frac{1}{N-2} \sum_{i=2}^N ((s[t_{E_i}] - s[t_{E_{i-1}}]) - F_{\text{TV,AF,e,MW}})^2} \quad (4.9b)$$

Aus den Extremstellen E_i wird darüber hinaus das Merkmal $F_{\text{TV,Peak,STD}}$ als Standardabweichung der Höhe der lokalen Extremwerte berechnet.

Über die Methode der kleinsten Fehlerquadrate wird eine Sinus-Funktion als Approximation des $s_{TV,EMD}$ gebildet. Als Merkmale $F_{TV,Sinus_{Ampl}}$ und $F_{TV,Sinus_{\omega}}$ werden die Amplitude und die Frequenz dieser Sinusapproximation definiert. Des Weiteren wird die absolute Differenz der Flächeninhalte von Sinuskurve und Regressand als Merkmal $F_{TV,Sinus_{Fehler}}$ verwendet. Dies kann als ein Gütekriterium der Sinusregression angesehen werden.

Neben den genannten Merkmalen aus dem Zeitbereich, können die Frequenzen des Signals von besonderem Interesse sein. Aus diesem Grund werden die drei stärksten Frequenzen und die zugehörigen Amplituden aus dem Betragsspektrum $F_{TV,FFT,abs}$ und dem Leistungsdichtespektrum $F_{TV,FFT,pow}$ ausgewählt.

Merkmale der forcierten Atmung

Aus dem forcierten Teil des Atemmanövers s_{VC} werden 36 Merkmale extrahiert. Diese sind zum Teil kongruent mit den zuvor beschriebenen Merkmalen der Ruheatmung. Analog berechnete Merkmale sind $F_{VC,\delta}$, $F_{VC,FFT,abs}$, $F_{VC,FFT,pow}$, F_{VC,U_-} und $F_{VC,Anstieg}$.

Merkmale basierend auf den gefundenen Extrempunkten werden ebenso berücksichtigt. Dafür werden Signalwert und Zeitpunkt der globalen Extremstellen als Merkmale F_{VC,A_i} und F_{VC,E_i} definiert, vergleiche Gleichungen (4.6a) und (4.6b). Die Differenz und somit die Spannweite des Signals s_{VC} entspricht der maximalen Thoraxbewegung verursacht durch die VC. Dies wird als Merkmal $F_{VC,VC}$ bezeichnet.

Weitere Merkmale

Die ROI hat durch die Auswahl des Messbereichs direkten Einfluss auf das Messsignal und somit die Bestimmung der Respirationsparameter. Zur Beschreibung dieser ROI werden dementsprechend ebenfalls Merkmale eingeführt, welche Hinweise auf den Zusammenhang darstellen sollen. In diesem Szenario wird eine statische ROI verwendet. Daher erfolgt die Beschreibung der ROI durch Merkmale ebenfalls statisch anhand des ersten auszuwertenden Bildes der Aufnahme. Die Merkmale stellen eine deskriptive Beschreibung dar.

Die Merkmale sind wie folgt definiert:

- Höhe und Breite der ROI, siehe Gleichungen (4.1a) und (4.1b),
- Gesamtfläche aus Multiplikation von Höhe und Breite,
- Mittelwert, analog zu Gleichung (4.2),
- Standardabweichung sowie
- Median und
- Minimal- und Maximalwert der Pixel der ROI.

Die Auswahl von phänotypischen Merkmale zielt auf ihren direkten Einfluss auf die Atemparameter ab. Diese Zusammenhänge wurden in Abschnitt 2.2.2 ausführlich beschrieben. Das Alter korreliert mit der Thoraxsteifigkeit und beeinflusst somit das Residualvolumen [29]. Die Lungengröße sowie das Tidalvolumen hängen von der Körpergröße [30] und dem Körpergewicht [31] ab. Trotz des umstrittenen Einflusses des Geschlechts auf die Respirationsparameter [32] wird dieses für die Evaluation hinzugezogen. Eine Erfassung des Trainingszustands ist nicht direkt, sondern nur über weitere Parameter und Messungen möglich und wird daher nicht berücksichtigt. Als Kombination von Größe und Gewicht wird der BMI verwendet.

Neben den oben beschriebenen, manuell ausgewählten Merkmalen mit direktem Bezug auf die Physiologie und vorausgegangene Zusammenhänge in der Literatur werden weitere Merkmale berechnet, die in keine direkten Verbindung zu den gesuchten Respirationsparametern stehen. Dies erfolgt anhand der Time Series Feature Extraction Library (TSFEL) [98]. Der Vorteil einer solchen Merkmalsberechnung ist, dass ein großes Spektrum an deskriptiven Signalmerkmalen automatisch extrahiert wird. Dies umfasst zum Beispiel den Zeitbereich, den Frequenzbereich und den Zeit-Frequenz-Verbundbereich. Neben klassischen deskriptiven Maßen wie Kurtosis und Exzess werden darüber hinaus Schätzer der empirischen kumulativen Verteilungsfunktion ermittelt. Insgesamt werden 144 Merkmale jeweils für $s_{TV,EMD}$ und 144 s_{SVC} berechnet. Durch die später folgende Merkmalsselektion, bei welcher ausschließlich die relevantesten Merkmale zur finalen Regression ausgewählt werden, ist es hilfreich, vorab eine große Anzahl an Merkmalen zur Verfügung zu stellen. Der systematische Ansatz durch TSFEL bereichert somit die zuvor manuell extrahierten Merkmale.

Zusammenfassung der extrahierten Merkmale

Insgesamt werden 379 Merkmale extrahiert. 52 Merkmale werden aus s_{TV} berechnet und weisen einen direkten physiologischen Bezug auf. Weitere 144 Merkmale werden für die Ruheatmung aus der TSFEL ermittelt. Die gleiche Anzahl an Merkmalen wird automatisch für s_{SVC} extrahiert. Weitere 26 manuelle Merkmale werden zusätzlich für die VC extrahiert. Acht Merkmale des Phänotyps und fünf bildbasierte Merkmale komplettieren die extrahierten Merkmale. Für eine vollständige Übersicht der Merkmale sei auf Tabelle A.1 verwiesen.

4.1.3 Merkmalsselektion

Um ausschließlich mit relevanten Merkmalen fortzufahren wird eine Merkmalsreduktion vorgenommen. Diese ermöglicht es, irrelevante und verrauschte Merkmale zu entfernen, den Regressionsprozess zu beschleunigen und die Vorhersagegenauigkeit sowie die Verständlichkeit durch bessere Erklärbarkeit zu erhöhen. Die zu reduzierende Merkmalsmenge umfasst die 379 extrahierten Merkmale aus Abschnitt 4.1.2. Die ideale Zielgröße wird als Hyperparameter optimiert.

Zur Auswahl einer geeigneten Merkmalsmenge werden Methoden eingesetzt, die den drei Bereichen der Merkmalsselektion entstammen. Dies entspricht den Filter-basierten, Wrapper-basierten und Embedded Methoden, siehe Abschnitt 2.4.1. Die Ergebnisse dieser Methoden werden im Anschluss mit verschiedenen Modellen (Abschnitt 4.1.4) kombiniert und untereinander verglichen.

F-Regression

Bei der Filter-Methode F-Regression (FR) werden die f -Werte der unabhängigen Variablen mit den abhängigen Variablen über univariante lineare Regressions-test bestimmt und als Metrik für die Bedeutung des entsprechenden Merkmals verwendet. Initial wird die Korrelation zwischen Regressor und dem Regressand berechnet. Als Regressor dienen die einzelnen Merkmale, zusammengefasst als Merkmalsmatrix $\mathbf{X} \in \mathbb{R}^{M \times N}$ mit N Merkmalen und M Messwerten. Als $\mathbf{y} \in \mathbb{R}^{M \times 1}$ wird die Grundwahrheit der zu bestimmenden Respirationsparameter angesehen. Die Korrelationskoeffizienten nach Pearson r_i (Gleichung (4.10a)) werden berechnet und anschließend in f -Kriterien umgerechnet (Gleichung (4.10b)), welche den

statistischen Zusammenhang zweiter Ordnung messen [99].

$$r_i = \frac{(\mathbf{X}[:, i] - \text{mean}(\mathbf{X}[:, i]))^\top \times (\mathbf{y} - \text{mean}(\mathbf{y}))}{\text{std}(\mathbf{X}[:, i]) \cdot \text{std}(\mathbf{y})} \quad (4.10a)$$

$$f_i = \frac{r_i^2 \cdot (M - 2)}{1 - r_i^2} \quad (4.10b)$$

Mutual Information

Die Transinformation (MI, mutual information) beschreibt den statistischen Zusammenhang zwischen unabhängiger Variable mit der abhängigen Variable [99] und ist somit ebenfalls eine Filter-Methode. Im Gegensatz zur F-Regression können über MI auch nicht-lineare Beziehungen erkannt werden. Die MI beschreibt die Stärke des statistischen Zusammenhangs. Klassisch wird die MI über die Entropie H berechnet, siehe Gleichung (4.11).

$$\text{MI}(\mathbf{X}[:, i], \mathbf{y}) = H(\mathbf{X}[:, i]) - H(\mathbf{X}[:, i]|\mathbf{y}) \quad (4.11)$$

Die verwendete Implementierung von scikit-learn [99] basiert für die Regression auf nicht parametrischen Methoden basierend auf Entropieabschätzung der k -nächsten Nachbarn.

Modellbasierte Merkmalsauswahl

Der modellbasierte Merkmalsselektion (MB) beschreibt in dieser Arbeit eine Wrapper-Methode zur Merkmalsreduktion anhand des entsprechenden Regressionsmodells, siehe Abschnitt 4.1.4. Hierbei wird eine Bewertung der Merkmale mithilfe des Modells vorgenommen und die besten Merkmale ausgewählt. Die Wichtigkeit der Merkmale wird für RF über den Gini-Koeffizienten [100], für SVR über die Wichtungsfaktoren bei linearem Kernel und für KNN über die Feature Permutation [101] evaluiert.

Rekursive Merkmalsauswahl

Als weitere Wrapper-Methode wird in dieser Arbeit Rekursive Merkmalselimination (RFE, recursive feature elimination) eingesetzt. Ausgehend von der vollen Merkmalsmenge werden rekursiv die Wichtigkeit der Merkmale anhand des jeweiligen Modells bestimmt und die irrelevantesten Merkmale entfernt. Dies wird wiederholt, bis die angestrebte Größe der auszuwählenden Merkmalsmenge erreicht

ist. Im Gegensatz zum modellbasierten Ansatz wiederholt sich die Bestimmung der Merkmalswichtigkeit in jedem Schritt. Die rekursive Selektion zielt darauf ab, die Zusammenhänge der einzelnen Merkmale untereinander zu berücksichtigen.

Embedded Merkmalsauswahl

Der modellbasierte Ansatz in Verbindung mit einem RF Regressionsmodell dient ebenso als Embedded Merkmalsauswahl. Das RF Modell wird in Abschnitt 4.1.4 beschrieben.

4.1.4 Regressionsmodelle

Zur Bestimmung der Atemvolumina über Regression der ausgewählten Merkmale werden verschiedene ML-Modelle eingesetzt und miteinander verglichen. Basierend auf den unterschiedlichen Eigenschaften dieser Modelle, sind zueinander abweichende Ergebnisse zu erwarten. Es erfolgt eine Optimierung geeigneter Hyperparameter, siehe Tabelle 4.1.

Random Forest

RF [55] werden in dieser Arbeit verwendet, da sie sich dadurch auszeichnen, mit vielen Merkmalen und mehrdimensionalen Eingangsdaten, ohne Optimierung von Hyperparametern, gute Klassifikations- und Regressionsergebnisse zu erzielen. Des Weiteren sind die Ergebnisse eines solchen Modells eingängiger erklärbar als zum Beispiel bei einem KNN. Nach Probst, Wright und Boulesteix [102] bieten sich jedoch zwei Hyperparameter für ein weiteres Tuning an. Diese sind die Stichprobengröße je Baum und die Anzahl der zu berücksichtigenden Merkmale. Die Stichprobengröße je Baum beschreibt die Anzahl der insgesamt N Trainingsdaten, welche für das Training jedes Entscheidungsbaums verwendet werden. Der zweite zu optimierende Parameter ist die Anzahl der zu berücksichtigenden Merkmale für die Suche nach einer weiteren Knotenunterteilung. Eine maximale Tiefe wird für diese Arbeit daher nicht festgelegt. Die Anzahl der Bäume wird nicht optimiert, sondern mit 600 hinreichend groß gewählt.

Künstliche Neuronale Netze

Bekannt als universeller Funktionsapproximator sind KNN in der Lage automatisch jede mathematische Lage approximieren zu können und sollen daher auch in dieser Arbeit untersucht werden. Die Funktionalität der KNN spiegelt sich in den einstellbaren Hyperparametern wieder. In dieser Arbeit wird sich auf die wesentlichen Einstellungsmöglichkeiten beschränkt, welche die generelle Netzstruktur, bestehend aus der Anzahl an Knoten und Zwischenschichten, der Aktivierungsfunktionen, der Lernrate und den Wiederholungen umfassen. Empirisch wird diesen Parametern die größte Bedeutung zugemessen. Weitere Parameter werden nicht hinzugefügt, um die Optimierungszeit zu reduzieren.

Support Vector Regression

Aufgrund der Generalisierungsfähigkeit auch bei höher dimensional Daten wird der Einsatz von SVR in dieser Arbeit untersucht. SVR ist eine Erweiterung von SVM [53] für Regressionsprobleme. Neben dem Kernel sind C und ϵ bedeutende Hyperparameter für die Optimierung der SVR [103]. Der Regularisierungsparameter C dient zur Einstellung des Verhältnisses von korrekter Klassifizierung der Trainingsdaten und Maximierung des Spielraums der Entscheidungsfunktion. ϵ ist der Parameter für den Fehlertoleranzbereich.

Ensemble Learning

Zusätzlich zu den einzelnen Modellen mit den entsprechenden Hyperparametern wird ein EL Ansatz in dieser Arbeit entwickelt. Dabei werden die ermittelten Respirationparameter aller Modelle gewichtet verwendet und zu einem gemeinsamen Ergebnis summiert. Die Gewichtsadjustierungen erfolgen über SVR mit den oben genannten Parametern. Vorteile dieses EL Ansatzes sind, dass die Ergebnisse vermeintlicher schwacher Schätzer zu einem Resultat mit geringer Varianz fusioniert werden können [104].

Zusammenfassung der Hyperparameter

Eine Zusammenfassung der zu optimierenden Hyperparameter findet sich in Tabelle 4.1. Für die Merkmalsselektion werden als Hyperparameter jeweils die Anzahl der auszuwählenden Merkmale k bestimmt. Diese werden in einer Schrittweite von 10 bis 60 erhöht. Die Auswahl der besten Parameter erfolgt anhand einer halbierenden Gittersuche, bei der alle Möglichkeiten miteinander kombiniert werden [99]. Initial wird eine reduzierte Menge an Daten zur Verfügung gestellt. Iterativ werden anschließend die besten Modelle ausgewählt und mehr Daten verwendet [105]. Als Fehlermaß dient der RMSE, welcher höhere Abweichungen stärker gewichtet.

Tabelle 4.1: Verwendete Hyperparameter der Maschine Learning Modelle und Merkmalselektion (FS) für Random Forest (RF), Support Vector Regression (SVR) und Künstliche Neuronale Netze (KNN).

Typ	Parametername	Werte
FS	Anzahl Merkmale k	[10, 20, 30, 40, 50, 50]
RF	Anzahl Merkmale zu berücksichtigen	$\left[\frac{k}{3}, k\right]$
RF	Stichprobengröße je Baum	$\left[\frac{N}{4}, \frac{N}{2}, \frac{3N}{4}, N\right]$
SVR	Kernel	[linear, poly, rbf, sigmoid]
SVR	Gamma	$\left[\frac{1}{k}, \frac{1}{k \cdot \text{var}(\mathbf{X})}\right]$
SVR	C	[0.5, 1, 2]
SVR	ϵ	[0.05, 0.1, 0.15]
KNN	Netz	[(60, 30, 15), (30, 15), (30)]
KNN	Aktivierung	[Tanh, Sigmoid]
KNN	Lernrate	[0.005, 0.05]
KNN	Epochen	[100, 200]

Referenzmodelle

Zum Vergleich der ML-Modelle werden die Respirationsmerkmale gleichermaßen über Referenzmodelle berechnet. Dafür wird zum einen ein Mittelwert-Regressor und zum anderen ein Probanden-Regressor verwendet. Der Mittelwert-Regressor bildet den Mittelwert der Trainingsdaten und gibt ihn als Prädiktion aus. Der Probanden-Regressor verwendet ausschließlich die phänotypischen Merkmale (Geschlecht, Größe, Alter, BMI und Gewicht) der Probanden, um damit über die lineare Regression die abhängige Variable zu bestimmen. Beide Verfahren werden somit bei mehrmaligen Messungen eines Probanden stets die gleichen Werte produzieren. Als Vergleich dienen diese Regressionsmodelle zur Abgrenzung und zum Beweis des Lerneffekts der eigens entwickelten ML-Modelle.

Zum Vergleich der Messverfahren mit dem Referenzmodell des Mittelwert-Regressors werden die Messfehler auf Unterschiede getestet. Als Messfehler dient die absolute Abweichung. Das Signifikanzniveau wird auf 95 % festgelegt und mit einem einseitigen, verbundenen T-Test getestet. Die Nullhypothese H_0 gibt an, dass der absolute Fehler des Modells größer oder gleich des absoluten Fehlers des Mittelwert-Regressors beträgt. Entgegengesetzt besagt die Alternativhypothese H_1 , dass der Fehler des Modells kleiner als der Referenz ist, siehe Gleichungen (4.12a) und (4.12b).

$$H_0 : |\text{TV}_{\text{Spirometer}} - \text{TV}_{\text{Modell}}| \geq |\text{TV}_{\text{Spirometer}} - \text{TV}_{\text{Mittelwert-Regressor}}| \quad (4.12a)$$

$$H_1 : |\text{TV}_{\text{Spirometer}} - \text{TV}_{\text{Modell}}| < |\text{TV}_{\text{Spirometer}} - \text{TV}_{\text{Mittelwert-Regressor}}| \quad (4.12b)$$

4.1.5 Evaluierungsstrategie

Ziel ist es, eine kalibrationsfreie DPG zu ermöglichen. Demnach müssen die Modelle auf zuvor ungesehenen Datensätzen performen. Dies ermöglicht jedoch ausschließlich eine Bewertung der Genauigkeit des Modells für ebendiese Testdaten. Aus diesem Grund wird für diese Arbeit Kreuzvalidierung mit Patientenauslassung (LOPCV, Leave-One-Patient-Out Cross Validation) verwendet, wobei das Modell anhand jedes Probanden getestet wird, während alle anderen Probanden im Trainingsdatensatz enthalten sind, siehe Abbildung 4.4. Um Probleme bei der gleichzeitigen Anwendung der Merkmalsselektion und Hyperparameteroptimierung hinsichtlich eines Datenlecks zu vermeiden, werden Merkmalsauswahl und

Hyperparameteroptimierung ebenfalls innerhalb der LOPCV durch Kreuzvalidierung durchgeführt. In diesem Fall spricht man von einer geschachtelten Kreuzvalidierung (nested Cross Validation). Es wird somit vermieden, dass das Modell aus Daten oder optimierten Parametern aus dem Testdatensatz profitieren kann (Datenleck).

Dieses Vorgehen ist geeignet, um den gesamtheitlichen Modellentwicklungsprozess zu testen und zu evaluieren, welche Genauigkeit der Raum an Hyperparametern inklusiver aller Merkmale mit den entsprechenden Modellen für die Probanden erreichen kann. Eine gesamtheitliche Analyse der Hyperparameter und ausgewählten Merkmale wird so jedoch erschwert, da das Vorgehen für die Anzahl an Probanden N wiederholt wird. Die Merkmalssektion wird somit beispielsweise insgesamt $N \cdot M$ durchgeführt, mit M der Anzahl der inneren Folds. Für das final beste Modell mit höchster Genauigkeit, wird der Ablauf für den Gesamtdatensatz wiederholt. Dies entspricht der Messung eines neuen Probanden unter gleichen Aufnahmebedingungen. Aus den somit ermittelten Hyperparametern und Merkmalen erfolgt abschließend eine physiologische Interpretation.

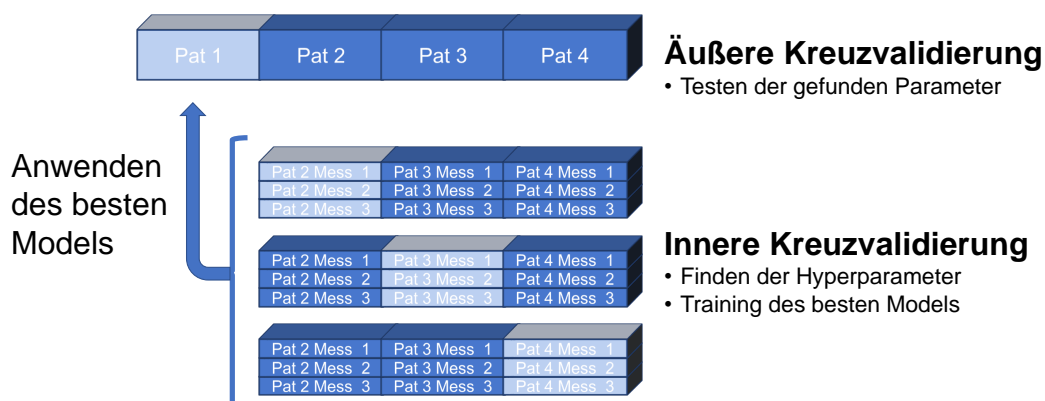


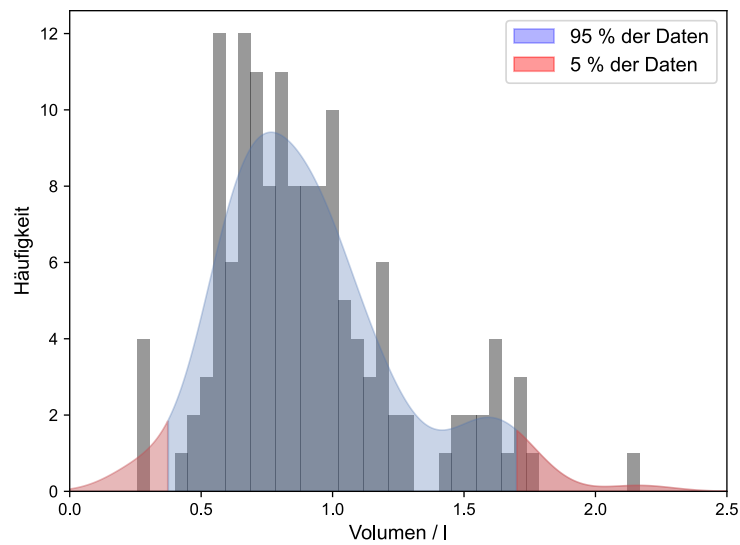
Abbildung 4.4: Geschachtelte Kreuzvalidierung zur Evaluierung der Ergebnisse und Vermeidung von Datenlecks am Beispiel von vier Patienten mit je drei Messungen.

4.2 Datensatz

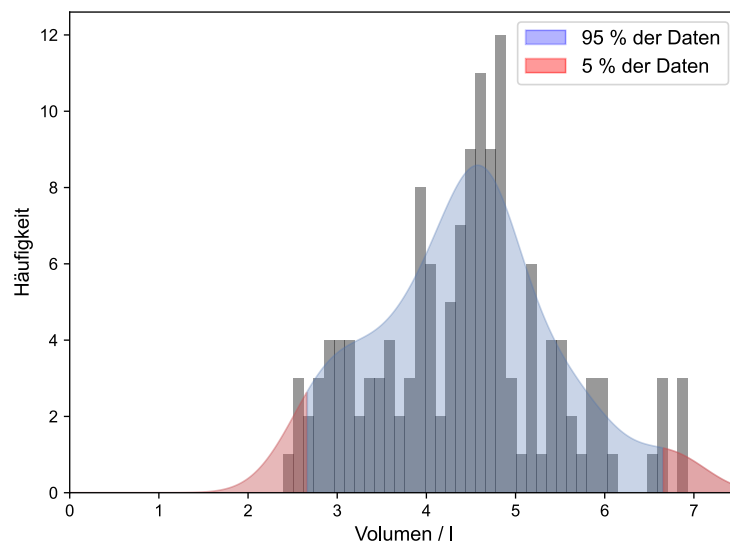
Die entwickelten ML-Modelle sollen mit dem bisherigen Stand der Technik verglichen werden. Aus diesem Grund wird ein frei zugänglicher Datensatz verwendet, um die in dieser Arbeit entstandenen Modelle den bestehenden, konventionellen, linearen Ansätzen der Literatur gegenüber zu stellen.

Für die Evaluation wird der Datensatz von Soleimani, Mirmehdi und Damen [94] verwendet. Darin enthalten sind insgesamt 298 Messungen von 35 gesunden Probanden. Diese Messungen enthalten ein Volumen-Zeit-Signal eines Referenz-Spirometers, zusammen mit den extrahierten Kenngrößen, unter anderem TV und VC. Die Hälfte der Aufnahmen erfolgte als Messung der SVC, während die andere Hälfte die FVC bestimmte. Da eine Messung der SVC einfacher von einem Probanden ohne Überwachung und Korrektur durch klinisches Personal durchgeführt werden kann und somit den gewünschten späteren Anwendungsbereich abdeckt, werden für diese Arbeit ausschließlich diese Aufnahmen verwendet. Des Weiteren ist erwähnenswert, dass die Probanden nicht hinsichtlich ihrer Bewegung eingeschränkt waren, in Vergleich zu Arbeiten mit zum Beispiel einer Rückenlehne [60, 61]. Für die Tiefeninformationen wurden zeitsynchron die Daten zweier Tiefenkameras aufgezeichnet. In dieser Arbeit werden jedoch ausschließlich die Aufnahmen der frontal ausgerichteten Kamera ausgewertet, da eine Verwendung multipler Kameras nicht von Laien im Alltag durchgeführt werden könnte, zusätzliches Material benötigt wird sowie die Transportfähigkeit des Systems eingeschränkt wird. Durch den Einsatz mehrerer Kameras lassen sich im Gegensatz die Oberkörperbewegungen präziser rekonstruieren und Bewegungsartefakte vermeiden.

Die Verteilung der Grundwahrheiten ist in Abbildung 4.5 dargestellt. Es werden in dieser Arbeit für TV und VC jeweils 95 % der Daten verwendet und die niedrigsten und höchsten 2,5 % verworfen. Dies erhöht die prozentuale Häufigkeit ähnlicher Messwerte für den Testdatensatz. Dieser Kompromiss verringert den Validierungsbereich des Modells, erhöht hingegen die Genauigkeit innerhalb desselben.



(a) Tidalvolumen



(b) Vitalkapazität

Abbildung 4.5: Verteilung der Grundwahrheit im verwendeten Datensatz [94] für Tidalvolumen (a) und Vitalkapazität (b) als Histogramm. Farblich in der Kernel-Dichte-Schätzung ist die Auswahl für die weiteren Modelle. 95 % der Daten werden weiter verwendet (blau) und 5 % werden verworfen.

4.3 Ergebnisse und Diskussion

Die Ergebnisse in diesem Kapitel analysieren das Vorgehen zur Entwicklung eines ML-Modells zur Regression der Respirationsparameter TV und SVC. Über den Vergleich mit der Literatur soll die Frage beantwortet werden, ob die nicht-linearen Modelle des ML zur Bestimmung der kontaktlosen Atemvolumina über eine Tiefenkamera geeignet sind. Darüber hinaus ist zu untersuchen, welche Merkmale zu der Regression am relevantesten beitragen und ob EL die Ergebnisse verbessern kann.

4.3.1 Vergleich der Methoden des Maschinellen Lernens

Tabelle 4.2a stellt die Ergebnisse der Modelle für die Bestimmung des TV dar. Das beste Modell ist hierbei SVR verbunden mit einer Merkmalsselektion über FR. Der durchschnittliche Fehler beträgt 11 ± 253 ml. Die Fehler dieser Kombination sind signifikant geringer als beim Mittelwert-Regressor. Am schlechtesten bezüglich der Standardabweichung schneidet das KNN mit der RFE und einer Messabweichung von -2 ± 330 ml ab. Das KNN weist von allen Modellen, abgesehen für die MB, die geringste mittlere Abweichung und die höchste Standardabweichung auf. Die Unterschiede zwischen RF und SVR liegen bei der Standardabweichung bei weniger als 15 ml. Diese stärkeren Abweichungen für die KNN sind möglicherweise auf die vermehrte Anzahl an Hyperparametern zurückzuführen. KNN benötigen eine größere Optimierung dieser Hyperparameter, wohingegen RF und SVR mit weniger Hyperparametern auskommen. Die Begrenzung auf vier verschiedene Hyperparameter mit zwei, beziehungsweise drei, Variationen in dieser Arbeit kann diese schlechtere Performance der KNN somit erklären.

Die FR führt, abgesehen vom KNN zu den niedrigsten Messfehlern, während die MI zu höheren mittleren Fehlern und einer höheren Standardabweichung führt. Modelle mit Wrapper basierten Merkmalsselektionen weisen mehrheitlich eine schlechtere Performance auf als die Filter-basierten Äquivalente. Auffällig ist, dass die MB über die Merkmalspermutation für das KNN mit -9 ± 260 ml als einzige Merkmalsselektion dafür sorgt, dass das KNN vergleichbare Messergebnisse mit den anderen Modellen aufweist. Diese Unterschiede in der Performance für verschiedene Merkmalsselektionsstrategien sind möglicherweise zurückzuführen auf die vorgegebene Größe der auszuwählenden Merkmale k . k wurde im Hyper-

parametertuning zwischen 10 und 60 in einer Schrittweite von 10 variiert. Eine optimale Anzahl von Merkmalen, die kein Vielfaches von 10 ist, kann somit dazu führen, dass nur eine hinreichend genaue Menge an Merkmalen auswählbar ist. Die Unterschiede in den Vorgehensweisen können potentiell auf diesen Einfluss zurückgeführt werden.

Tabelle 4.2: Genauigkeiten der Verfahren zur Bestimmung des Tidalvolumens (a) und Vitalkapazität (b). Als Maschine Learning Verfahren werden Künstliche Neuronale Netze (KNN), Support Vector Regression (SVR) und Random Forest (RF) eingesetzt. Zur Merkmalsselektion kommen F-Regression (FR), Mutual Information (MI), modellbasierte Merkmale (MB) und rekursive Merkmalselimination (RFE) zum Einsatz. Die Angabe der Genauigkeit erfolgt als mittlerer Fehler \pm Standardabweichung.

(a) Tidalvolumen (TV)

	FR	MI	MB	RFE
RF	-20 ± 264 ml	-33 ± 273 ml	-14 ± 280 ml	-25 ± 283 ml
SVR	11 ± 253 ml ***	37 ± 271 ml *	21 ± 283 ml	37 ± 269 ml **
KNN	-12 ± 292 ml	-26 ± 278 ml	-9 ± 260 ml **	-2 ± 330 ml

(b) Vitalkapazität (VC)

	FR	MI	MB	RFE
RF	-23 ± 720 ml **	-6 ± 726 ml **	-71 ± 713 ml **	-90 ± 690 ml ***
SVR	-43 ± 768 ml **	11 ± 624 ml ***	-88 ± 723 ml ***	-65 ± 647 ml ***
KNN	-7 ± 661 ml ***	-18 ± 676 ml ***	4 ± 653 ml ***	11 ± 661 ml ***

* Signifikant geringer ggü. Mittelwert-Regressor (TV: 0 ± 288 ml, VC: 2 ± 889 ml) $p < 0.05$.

** Signifikant geringer ggü. Mittelwert-Regressor (TV: 0 ± 288 ml, VC: 2 ± 889 ml) $p < 0.01$.

*** Signifikant geringer ggü. Mittelwert-Regressor (TV: 0 ± 288 ml, VC: 2 ± 889 ml) $p < 0.001$.

Das EL erzielt eine Abweichung von 18 ± 283 ml und somit eine der höchsten Standardabweichungen von RF und SVR. Der Mittelwert-Regressor weist mit 0 ± 288 ml eine höhere Standardabweichung auf. Da es mit SVR gleich mehrere Modelle gibt, die einen signifikant geringeren Fehler haben, zeigt sich, dass die Modelle evident aus den Merkmalen gelernt haben und diese Modelle zu einer genaueren Regression des TV führen. Die persönlichen Merkmale scheinen dies nicht alleine zu klären, da der Probanden-Regressor als weiteres Vergleichsmodell mit -13 ± 291 ml einen höheren Messfehler aufweist.

Für die SVC sind die Ergebnisse in Tabelle 4.2b dargestellt. Im Gegensatz zum TV ist die Standardabweichung um mehr als 300 ml größer, was sich teilweise durch die größere Spannweite der SVC Verteilung erklären lässt. Eine weitere Ursache ist die Datengrundlage: Liegen für die Bestimmung des TV mehrere Atemzüge und rein manuell 52 ausgewählte Merkmale vor, so ist die Zeitspanne der forcierten Atmung deutlich reduziert und es gibt ausschließlich eine einzige Messwiederholung und 26 rein manuell vorgesehene, extrahierte Merkmale. Mit der Körperbewegung wird eine weitere Ursache ebenfalls in der Literatur durch [60–64, 71, 85, 86] beschrieben. Dies ist gleichzusetzen mit einer Oberkörperbewegung bei der forcierten Atmung. Diese Oberkörperbewegung superpositioniert die zu messende Thoraxbewegung und agiert als Störquelle. Bei starker Inspiration bewegt sich der gesamte Oberkörper nach vorne. Ebenso dehnt sich der Thorax in frontale Richtung aus. Die Überlagerung beider Vorgänge reduziert die Signalqualität.

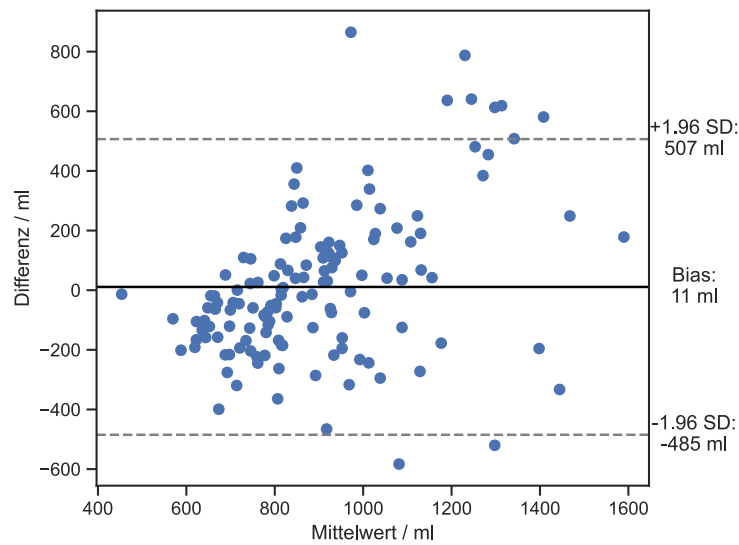
Im Gegensatz zum TV resultieren die Wrapper-basierten Merkmalsselektionsverfahren für die VC in besseren Ergebnissen als die Filter-Methoden. Die Ursache hierfür könnte im komplexeren Zusammenhang stecken, der zwischen den Regressoren und Regressanden gefunden werden muss und die sich bereits in der geringeren Genauigkeit widerspiegelt. Zur Detektion der Zusammenhänge der Merkmale untereinander und im Zusammenspiel mit dem Modell sind die Wrapper-Verfahren den Filter-basierten Methoden überlegen. Das bestes Verfahren stellt sich dennoch die SVR mit vorausgehender Merkmalsselektion über MI heraus, mit einer Messabweichung von 11 ± 624 ml.

Es ist herauszuheben, dass alle Verfahren sich signifikant vom Mittelwert-Regressor unterscheiden. Dieser weist eine Messabweichung von 2 ± 889 ml auf. Dies deutet auf die große Varianz im Datensatz hin. Die mittlere Messabweichung

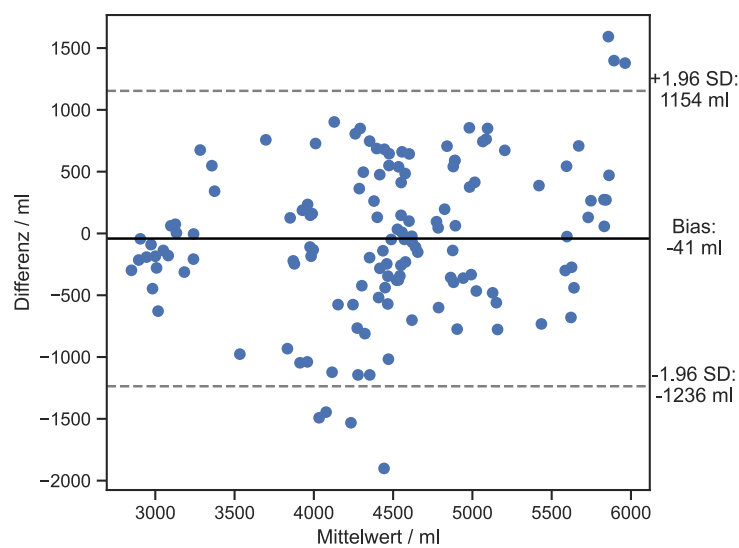
von 2 ml ist durch Rundungsungenauigkeiten zu erklären. Des Weiteren zeigt sich, dass der Probanden-Regressor mit -9 ± 611 ml ein besseres Ergebnis erzielt als die zu testenden Modelle. Dies unterstreicht, dass Merkmale nicht von ausreichender Qualität sind, um auf die SVC zu schließen. Die Begründungen hierfür sind in der Signalqualität zu suchen, welche aus den oben genannten Gründen schlechter als bei der Ruheatmung ist. Im Vergleich zu den einzelnen Modellen kann das EL für die SVC ein besseres Ergebnis erzielen, mit einem Messfehler von -41 ± 610 ml. Die Standardabweichung ist geringer als beim Probanden-Regressor.

Im Bland-Altman-Diagramm [106] werden zwei Messmethoden miteinander verglichen und graphisch dargestellt. Für diese Arbeit sind die besten Modelle entsprechend abgebildet. In Abbildung 4.6a ist das, mit FR und SVR, beste Modell für die Bestimmung der TV mit den vorliegenden Grundwahrheitswerten eines Spirometers gegenübergestellt. Abbildung 4.6b stellt das EL für die SVC den Spirometerdaten gegenüber. Es werden auf der Abszisse die Mittelwerte beider Messungen abgetragen und auf der Ordinate die Differenzen. Zur besseren Interpretation sind Mittelwert der Differenzen, sowie die Obergrenze der Zustimmung (ULoA, Upper Limit of Agreement) und die Untergrenze der Zustimmung (ULoA, Lower Limit of Agreement) eingetragen.

Beide Darstellungen zeigen eine konsistente Schätzung der Messwerte durch die entwickelten Modelle. Für TV befinden sich 92,6% der Messungen innerhalb der Zustimmungsgrenzen und für VC 94,8%. Vereinzelt sind Abweichungen von mehr als 800 ml für das TV zu erkennen, beziehungsweise 1500 ml für die VC respektive. Die Ausdehnung von ULoA und ULoA weisen darauf hin, dass die entwickelten Verfahren nicht die erforderliche klinische Genauigkeit eines Spirometers erreichen. Diese erfordert eine Messabweichung für die VC von nicht mehr als 150 ml [10].



(a) Tidalvolumen: F-Regression und Support Vector Regression



(b) Vitalkapazität: Mutual Information und Ensemble Learning

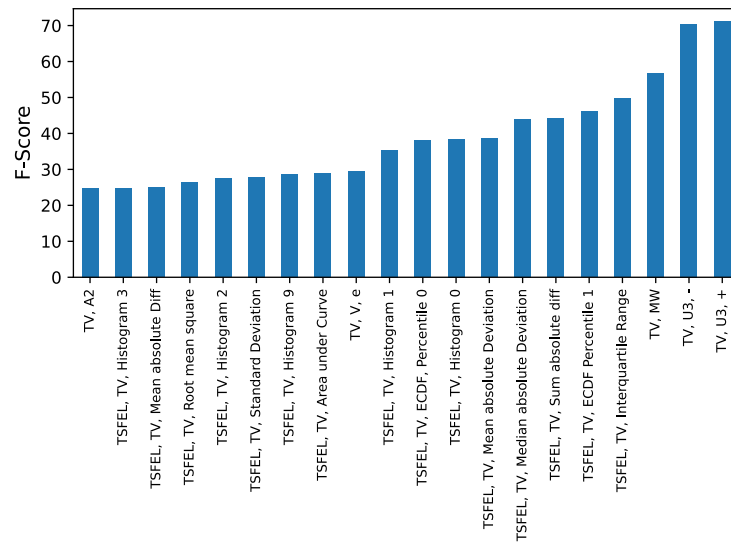
Abbildung 4.6: Bland-Altman-Darstellung zur Abweichung der entwickelten Modelle zum Referenzspirometer. Die besten Modelle sind für das Tidalvolumen (a) und die Vitalkapazität (b) dargestellt. Die mittlere Messabweichung, sowie Ober- und Untergrenze der Zustimmung sind eingetragen.

4.3.2 Relevante Merkmale für die Regression

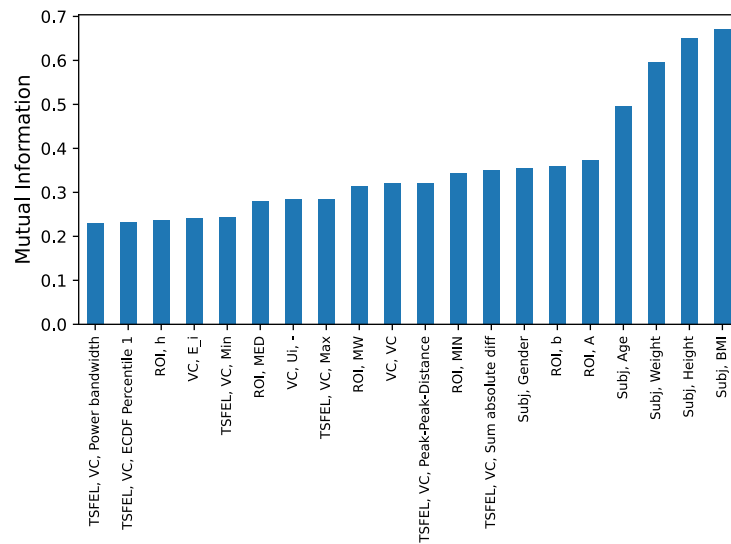
Die Merkmale haben einen wesentlichen Einfluss auf das Regressionsergebnis. Daher ist es von Interesse, die ausgewählten Merkmale zu diskutieren. Die 20 relevantesten Merkmale, sortiert nach dem F-Score mit dem TV als Regressand, sind in Abbildung 4.7a dargestellt. Dies entspricht dem besten Modell aus Abschnitt 4.3.1.

Für das TV weisen drei manuell ausgewählte Merkmale den höchsten F-Score auf. Dabei sind die Werte um das erste Extrema von besonderer Bedeutung sowie die mittlere Distanz zwischen den Peaks. Letzteres Merkmal lässt sich erklären, da dies der direkten Entsprechung des TV aus einem Volumen-Zeit-Signal entspricht. Die Umgebung der Signalpeaks kann von Bedeutung sein, da diese Rückschlüsse auf die Atemtechnik zulässt. Eine spezifische Wertigkeit des ersten Atempeaks im Signal kann nicht ausgemacht werden. Zu weiteren ausgewählten Merkmalen gehören 15 Merkmale aus der TSFEL. Es sind im Besonderen deskriptive Merkmale von Bedeutung, wie Interquartilsabstand oder die empirische kumulative Verteilungsfunktion. Mit geringerem F-Score gibt es unter den 60 besten Merkmalen insgesamt 16 Merkmale die auf spektrale Komponenten des Signals abzielen. Unter den 60 besten Merkmalen befinden sich mit BMI, Geschlecht und Gewicht drei Merkmale, welche direkten Bezug zum Patienten haben. Diese Merkmale sind ebenfalls für die VC relevant.

Die ausgewählten Merkmale anhand der MI für die VC sind in Abbildung 4.7b dargestellt. Es ist zu erkennen, dass drei Merkmale eine deutlich höhere MI aufweisen als die restlichen 17. Diese Merkmale sind: Größe, BMI und Gewicht, gefolgt vom Alter. Anschließend folgen Merkmale der ROI und mit dem Geschlecht das letzte persönliche Merkmal. Somit sind alle persönlichen Merkmale vertreten, was die Korrelation bestätigt, die in der Literatur zwischen diesen Größen und den VC aufgezeigt wurde, vergleiche Abschnitt 2.2.2. Die ROI wird anhand von Körperpositionen, wie den Schultern und der Hüfte definiert. Somit verändert sich die ROI-Ausdehnung bei größeren Patienten, was die Auswahl dieser ROI-Merkmale erklärt.



(a) F-Scores für die Bestimmung des TV



(b) Mutual Information für die Bestimmung der VC

Abbildung 4.7: 20 relevanteste Merkmale für Tidalvolumen (TV) (a) und Vi-talkapazität (VC) (b) anhand der entsprechenden Merkmalsselektion. Die Zuordnung der Merkmale ist mithilfe von Tabelle A.1 möglich.

Die Relevanz der persönlichen Merkmale wirft die Frage auf, welche Genauigkeit erzielt werden kann, wenn diese nicht zur Verfügung stehen. Daher sind die folgenden Ergebnisse mit den Merkmalen dargestellt, welche ausschließlich mit der Tiefenkamera aufgenommen werden. Zur Berechnung wird ansonsten die gleiche Vorgehensweise verwendet, wie für Abschnitt 4.3.1 dargestellt.

Für das TV in Tabelle 4.3a unterscheiden sich die Fehler im geringen Maße von den Modellen mit persönlichen Merkmalen. Als beste Vorgehensweise geht erneut die FR mit einem SVR Modell hervor. Der Fehler ist mit 18 ± 267 ml größer als bei Hinzunahme der Probanden-Merkmale. Es ist auffallend, dass die KNN Modelle bedeutend schlechter abschneiden und höhere Standardabweichungen aufweisen. Das EL weist eine Messabweichung von 27 ± 286 ml auf und ist somit erneut nicht performanter als die besten Einzelmodelle.

Deutliche Unterschiede sind erneut bei der VC festzustellen. Die Fehler für die Modelle ohne persönliche Merkmale weisen eine um mehr als 100 ml höhere Standardabweichung, als die korrespondierenden Modelle mit diesen Merkmalen. Die einzige Ausnahme ist SVR mit RFE. Als bestes Modell weist es einen Fehler von -9 ± 737 ml auf. Das EL liegt in einem vergleichbaren Fehlerbereich mit höherer mittlerer Abweichung von 78 ± 737 ml. Beide Modelle sind signifikant verschieden vom Mittelwert-Regressor.

Der Einfluss der persönlichen Merkmale wurde bereits in Abschnitt 2.2.2 dargestellt. Alter, Größe, Gewicht, Geschlecht und der Trainingszustand beeinflussen die Respirationsparameter. Der Einfluss dieser Merkmale ist besonders für VC deutlich. Die Überlegenheit des Probanden-Regressors (VC: -9 ± 611 ml), welcher ausschließlich mit den persönlichen Merkmalen lernt, verdeutlicht, dass die extrahierten Merkmale nicht geeignet sind, die VC zu bestimmen. Mögliche Ursachen wurden bereits in Abschnitt 4.3.1 genannt und sind in der Bewegung des Oberkörpers bei forcierter Atmung zu suchen.

Tabelle 4.3: Genauigkeiten der Verfahren ohne persönliche Merkmale für Tidalvolumen (a) und Vitalkapazität (b). Als Maschine Learning Verfahren werden Künstliche Neuronale Netze (KNN), Support Vector Regression (SVR) und Random Forest (RF) eingesetzt. Zur Merkmalsselektion kommen F-Regression (FR), Mutual Information (MI), modellbasierte Merkmale (MB) und rekursive Merkmalselimination (RFE) zum Einsatz. Die Angabe der Genauigkeit erfolgt als mittlerer Fehler \pm Standardabweichung.

(a) Tidalvolumen

	FR	MI	MB	RFE
RF	-23 ± 264 ml	-32 ± 282 ml	-19 ± 281 ml	-13 ± 285 ml
SVR	18 ± 267 ml *	12 ± 274 ml *	30 ± 288 ml	27 ± 271 ml **
KNN	6 ± 283 ml	21 ± 311 ml	24 ± 342 ml	-35 ± 343 ml

(b) Vitalkapazität

	FR	MI	MB	RFE
RF	-84 ± 842 ml	-65 ± 801 ml *	-35 ± 799 ml *	-44 ± 813 ml *
SVR	28 ± 907 ml	-18 ± 809 ml	-24 ± 909 ml	-9 ± 737 ml ***
KNN	-44 ± 881 ml	-55 ± 799 ml	-42 ± 848 ml	33 ± 765 ml *

* Signifikant geringer ggü. Mittelwert-Regressor (TV: 0 ± 288 ml, VC: 2 ± 889 ml) $p < 0.05$.

** Signifikant geringer ggü. Mittelwert-Regressor (TV: 0 ± 288 ml, VC: 2 ± 889 ml) $p < 0.01$.

*** Signifikant geringer ggü. Mittelwert-Regressor (TV: 0 ± 288 ml, VC: 2 ± 889 ml) $p < 0.001$.

4.3.3 Messgenauigkeit im Vergleich zur Literatur

Abschließend werden die für diese Arbeit entwickelten Modelle mit den bestehenden Modellen der Literatur verglichen. Da ein frei zugänglicher Datensatz verwendet wurde, werden die darauf angewendeten Algorithmen verglichen. Eine Zusammenfassung dieser Ergebnisse ist in Tabelle 4.4 dargestellt. Da in dieser Arbeit die mittleren 95 % der Datenverteilung verwendet wurden, wird dies ebenso für die Vergleichsdaten aus der Literatur angewendet. Die Entfernung der Werte mit starker Abweichung vom Mittelwert verringert die Varianz der Messgenauigkeit. Im Folgenden sind die Ergebnisse der Modelle von Soleimani zur besseren Vergleichbarkeit daher ebenso auf die Auswahl von 95 % der Daten angegeben. Die Genauigkeit mit höherer Varianz für alle Daten ist in Tabelle 3.1 dargestellt. Die Modelle aus der Literatur verwenden dennoch alle Daten zum Kalibrieren.

Die Gruppe um Soleimani hat zwei verschiedene Modelle für die Bestimmung der Respirationsparameter entwickelt. Für die Verwendung einer Kamera resultieren Abweichungen von 70 ± 226 ml für das TV und -143 ± 847 ml für die VC [71]. Dieses Modell verwendet jedoch Daten desselben Probanden zur Kalibration.

Ein anderes Modell der Gruppe [20] benutzt neben der frontal ausgerichteten Kamera, die Kamera, welche auf die Rückseite der Patienten ausgerichtet ist. Beide Kameras sind zueinander registriert. Auch dieses Modell benötigt eine patientenindividuelle Kalibration. Es ermöglicht eine Bestimmung des TV mit -23 ± 205 ml und der VC mit -318 ± 524 ml.

Tabelle 4.4: Vergleich der entwickelten Modelle mit der Literatur. Die Ergebnisse [20, 71] sind entgegen den Veröffentlichungen für 95 % der Daten angegeben.

Modell	Tidalvolumen	Vitalkapazität
Soleimani - 1 Kamera [71]	70 ± 226 ml	-143 ± 847 ml
Soleimani - 2 Kameras [20]	-23 ± 205 ml	-318 ± 524 ml
Mean-Regressor	0 ± 288 ml	2 ± 889 ml
Probanden-Regressor	-13 ± 291 ml	-9 ± 611 ml
Diese Arbeit mit persönlichen Merkmalen	11 ± 253 ml	11 ± 624 ml
Diese Arbeit ohne persönliche Merkmale	18 ± 267 ml	-9 ± 737 ml

Im Vergleich zu diesen Modellen auf dem gleichen Datensatz können die ML-Modelle dieser Arbeit für die VC deutlich besser abschneiden. Die Standardabweichung ist für die VC um mehr als 200 ml geringer, im Vergleich zum Modell, welches ebenfalls eine Kamera verwendet. Einzig das TV kann mit dem Einsatz von zwei Kameras durch Soleimani u. a. [20] genauer bestimmt werden. Für eine Kamera weist das Modell aus der Literatur einen um mehr als 50 ml größerem mittleren Fehler auf, aber eine um 25 ml geringere Standardabweichung. Es ist zusätzlich hervorzuheben, dass die ML-Modelle keine Kalibration auf die jeweiligen Patienten benötigen und mit den Daten von ausschließlich einer Kamera auskommen..

Vergleiche mit anderen Modellen, welche ohne Kalibration auf den entsprechenden Probanden auskommen, müssen hinsichtlich der Unterschiede im Messkonzept diskutiert werden. Ostadabbas u. a. [61] berichten von Messfehler von 70 ± 60 ml für das TV. Die Messungen der Probanden erfolge im Sitzen mit Rückenlehne, wobei die Patienten angehalten waren, sich nicht zu bewegen. Dies ist somit ein Unterschied, zu dem in dieser Arbeit verwendeten Datensatz, in dem es zum Teil zu starker Bewegung kommt und keine Fixierung des Rückens vorgenommen wird. Ebenfalls unterschiedlich ist die Messzeit von 8 min für das Modell von Ostadabbas. Eine längere Messzeit kann über die Mittelung vieler Einzelmessung zu besseren Ergebnissen führen. Kritisch zu sehen ist ebenfalls, dass die Referenzdaten nicht simultan aufgenommen wurden, sondern zum Teil aus Lungenfunktionstest von über einem Jahr stammen. Das Modell ist daher nicht als Referenz für diese Arbeit geeignet.

Ebenfalls ohne Patienten-Kalibration kommt das Modell von Oh u. a. [60] aus. Abweichend wurden Bewegungen der Probanden ausgeschlossen. Zur Evaluation wurde keine natürliche Atmung eingesetzt, sondern das eingestellte Volumen eines Beatmungsgeräts. Trotz der Abweichung von $8,41 \pm 2,16\%$ für das TV ist das Modell nicht alltagstauglich. Die Bearbeitung für 4s Bildaufnahmen benötigt circa 10 min.

Mit einer Abweichung von 57 ± 716 ml für die VC kommt das Modell von Takamoto u. a. [43] ohne Kalibration aus. Im verwendeten Testdatensatz sind insgesamt 30s Abschnitte von Patienten im Sitzen mit einer Evaluierung durch LOOCV sowie phänotypische Merkmale der Patienten. Der Messfehler ist mit den in dieser Arbeit beschriebenen Messfehlern vergleichbar. Jedoch konnte durch die

Hinzunahme von persönlichen Merkmalen in dieser Arbeit der Fehler reduziert werden auf 11 ± 624 ml.

Es lässt sich somit resultieren, dass die Modelle des ML aus dieser Arbeit geringere Fehler aufweisen, als vergleichbare Modelle in der Literatur. Insbesondere auf dem gleichen Datensatz sind deutlich geringere Messabweichungen zum Referenz-Spirometer durch nicht-lineare KI-Modelle möglich. Die schwerer zu bestimmende VC weist einen deutlich geringen Messfehler auf. Andere kalibrationsfreie Modelle sind nicht alltagstauglich oder weisen höhere Messabweichungen auf. Die in dieser Arbeit entwickelten Modelle bieten somit einen Mehrwert im Vergleich zum bisherigen Stand der Forschung.

4.4 Zusammenfassung

In diesem Kapitel wurde ein Ansatz der Literatur aufgegriffen und weiterentwickelt, Merkmale zu extrahieren, um daraus auf Respirationparameter zu schließen. Dafür wurden insgesamt 379 Merkmale berechnet. 78 Merkmale davon wurden entwickelt mit einem direkten physiologischen Bezug für TV und VC. Acht Merkmale entstammen der definierten ROI. Zusätzlich werden fünf persönliche Merkmale der Probanden erhoben sowie 288 weitere Merkmale automatisch extrahiert. Zur Selektion der besten Merkmale wurden Filter, Wrapper und Embedded Methoden eingesetzt. Als nicht-lineare ML Modelle wurden KNN, RF und SVR sowie ein EL entwickelt und miteinander verglichen. Modelle und Merkmalsselektionsverfahren wurden in allen Varianten kombiniert. Die Evaluation erfolgte auf einem frei verfügbaren Datensatz aus der Literatur mit insgesamt 148 Aufnahmen zur SVC und TV. Dabei wird eine LOPCV verwendet, die den Modellbildungsprozess evaluiert und ein Overfitting durch die Hyperparameter und die Modelle vermeidet. Somit wird jeder Proband iterativ als Testdatensatz verwendet. Eine Kalibration der Modelle auf die Probanden ist nicht notwendig.

Das Tidalvolumen kann mit einer Genauigkeit von 11 ± 253 ml bestimmt werden und die Vitalkapazität mit -41 ± 610 ml respektive. SVR und RF sind den KNN überlegen, was auf eine schwierigere Abstimmung der Hyperparameter zurückzuführen ist. Eine genauere Untersuchung der ausgewählten Merkmale zeigt eine starke Abhängigkeit von den subjektiven Patientenmerkmalen. Ohne diese sinkt die Genauigkeit zur TV Bestimmung auf 18 ± 267 ml und -9 ± 737 ml

für VC respektive. Dies lässt sich auf die physiologischen Zusammenhänge dieser Größen zurückführen und die damit verbundene starke Korrelation. Für das TV erzielt das ML Modell bessere Ergebnisse als die Referenz-Modelle, bestehend aus Mittelwert-Regressor und Probanden-Regressor. Die VC Bestimmung des Modells ist dem Probanden-Regressor unterlegen. Dies lässt den Schluss zu, dass die extrahierten Merkmale zur Bestimmung der VC ungenügend sind. Eine Ursache hierfür ist die Bewegung des Oberkörpers bei der forcierten Atmung. Diese Bewegung überlagert die zu messende Bewegung des Brustkorbs und führt zu starken Abweichungen. Im Vergleich zu bestehenden Modellen auf demselben Datensatz weisen die ML Modelle dieser Arbeit geringere Messabweichungen auf. Zusätzlich benötigen sie keine Kalibration auf die Probanden. Andere kalibrationsfreie Modelle aus der Literatur verwendeten unzureichende Referenzwerte zur Evaluierung, benötigen eine im Alltag nicht handhabbare Berechnungszeit oder weisen schlechtere Genauigkeiten auf.

Es lässt sich resultieren, dass die in dieser Arbeit entstandenen Methoden zur Entwicklung eines ML-Modells zur kontaktlosen Bestimmung der Atemparameter besser geeignet ist als bestehende klassische Verfahren. Die entwickelten Methoden zur Extraktion von Merkmalen aus dem Tiefen-Zeit-Signal sind von Bedeutung für die Regression. Die Modelle sind kalibrationsfrei auf neue Probanden anwendbar und erweitern den Stand der Technik. Als Problem bleibt die Bestimmung der Vitalkapazität bestehen, die durch Oberkörperbewegungen weiterhin erschwert messbar ist und eine klinische Nutzung nicht erlaubt.

Kapitel 5

Verbesserung der Signalextraktion

Die Ergebnisse aus Kapitel 4 zeigen die Überlegenheit von ML-Verfahren gegenüber herkömmlichen Modellen aus der Literatur. Dafür erhielten die Modelle durch gleiche Vorverarbeitungsschritte die gleichen Eingangssignale. Die ML-Modelle weisen geringe Messfehler auf und sind somit den bestehenden Modellen überlegen. Ziel dieses Kapitels ist es nun, die Vorverarbeitung der Daten zu verbessern. Insbesondere die Extraktion des Messsignals steht dabei im Fokus der Untersuchungen.

Bisherige Signalextraktionsmethoden weisen ein hohes Maß für Verbesserung hinsichtlich physiologischer Aspekte der Atmung auf. Wie in Abschnitt 3.1 bereits dargestellt, werden bisher meist simple, rechteckige Regionen auf der Brust der Patienten ausgewählt, um ein Messsignal zu extrahieren. Komplexere Verfahren verwenden Amplitudeninformationen zur Auswahl geeigneter Regionen sowie spatiale und temporale Informationen. Dies ist mit einem erhöhten Berechnungsaufwand verbunden und basiert nicht auf weitergehenden physiologischen Betrachtungen. Für eine gezielte Differenzierung von Subregionen werden derzeit meist markerbasierte Verfahren eingesetzt, die jedoch eine manuelle Anbringung benötigen. Die Untersuchungen in diesem Kapitel sollen die Vorteile bisheriger Verfahren kombinieren. Es soll eine Region identifiziert werden, welche für die automatische Anwendung geeignet ist und welche Rückschlüsse auf physiologische Atemprozesse zulässt. Zusätzlich soll die Qualität des somit extrahierten Signals gesteigert werden.

Ausgehend von dieser Signalextraktion und somit potentiell verbesserten Eingangsdaten für die Modelle aus Kapitel 4, soll untersucht werden, inwieweit die bestehenden ML-Modelle zu präziseren Bestimmungen der Respirationsparameter möglich sind. Während die bisherigen Modelle aus extrahierten Merkmalen direkt auf die Respirationsparameter schließen, ist zu untersuchen, inwiefern das vollständige Volumen-Zeit-Signal v aus dem Messsignal s approximiert werden kann. Dafür ist eine Transformation des Messsignals notwendig. In der Literatur [20] wird hierfür ein Skalierungsfaktor ($[\alpha] = 1 \text{ mm}^{-1}$) eingesetzt, siehe Gleichung (5.1). Es zeigt sich jedoch, dass α schon bei verschiedenen Messungen eines Probanden variiert. Soleimani u. a. [71] verwendete ebenfalls eine Unterteilung des Signals innerhalb einer Messung in Tidalatmung (Gleichung (5.1a)) und forcierte Atmung (Gleichung (5.1b)) mit entsprechenden Skalierungsfaktoren. Diese verschiedenen Skalierungsfaktor innerhalb einer Messung sind auf Bewegungsartefakte zurückzuführen.

$$v_{TV}[t] = \alpha_{TV} \cdot s_{TV}[t] \quad (5.1a)$$

$$v_{VC}[t] = \alpha_{VC} \cdot s_{VC}[t] \quad (5.1b)$$

Innerhalb dieses Kapitels wird der Einsatz von rekurrenten neuronalen Netzen zur Transformation des Messsignals analysiert. Diese können das gesamte Messsignal als Eingang nehmen, um das gesuchte Volumen-Zeit-Signal vollständig zu approximieren und werden bereits erfolgreich in der Biosignalverarbeitung eingesetzt [107]. Die Signal-Approximation erlaubt die Bestimmung weiterer Respirationsparameter anhand des Signals.

Die Untersuchungen dieses Kapitels zielen somit zusammenfassend auf die Beantwortung folgender Fragen ab:

- Ist eine automatisch angelegte und an die Physiologie angepasste Region zur Extraktion eines Tiefen-Zeit-Messsignals einer konventionellen rechteckigen Form überlegen?
- Welche physiologischen Zusammenhänge lassen sich aus der ausgewählten Region und der Atemmechanik ableiten?
- Wie genau können Respirationsparameter durch eine Volumen-Zeit-Signalapproximation mittels rekurrenter neuronale Netze bestimmt werden?

5.1 Methodik

Der Ansatzpunkt zur Verbesserung der Signalextraktion ist die Bestimmung einer neuen, physiologischen ROI. Dafür werden die statistischen Zusammenhänge der Signale eines jeden Bildpunkts mit dem Spirometersignal untersucht. Zusammenhängende Regionen mit einem hohen statistischen Zusammenhang mit dem Respirationssignal werden segmentiert. Aus der Mittelung aller Patientenmasken entsteht so eine generalisierte Maske zur Extraktion der Messsignale. Diese verbesserten Eingangssignale dienen als Eingang in die Modelle aus Kapitel 4 zur Bestimmung von Tidalvolumen und Vitalkapazität. Ebenfalls wird der Einsatz rekurrenter neuronaler Netze analysiert, welche das gesamte Volumen-Zeit-Signal approximieren. Eine Validierung erfolgt über LOOCV. Im Folgenden werden die Schritte im Detail dargestellt. Eine Übersicht des Vorgehens in Form eines Blockdiagramms ist in Abbildung 5.1 gegeben.

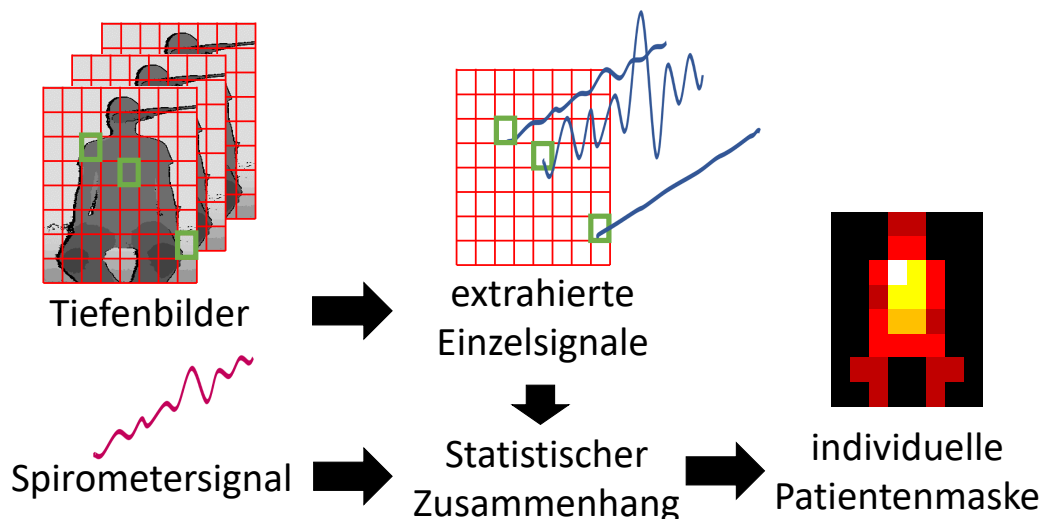


Abbildung 5.1: Allgemeines Vorgehen bei der Erstellung von individuellen Patientenmasken. Aus den Rohdaten (Tiefenbilder), welche über eine Tiefenkamera aufgezeichnet werden, werden pixelweise Zeitsignale extrahiert. Der statistische Zusammenhang dieser Signale mit dem Spirometersignal wird als individuelle Maske gespeichert. Das Maß des Zusammenhangs ist hier beispielhaft farbcodiert dargestellt mit hohem Zusammenhang in Weiß und keinem Zusammenhang in Schwarz.

5.1.1 Untersuchung der Region of Interest

Die Untersuchung einer optimierten ROI beruht auf der Annahme, dass verschiedene Regionen des Thorax am physiologischen Atemprozess beteiligt sind, wie zum Beispiel bei thorakaler und abdominaler Atmung. Zur Analyse der ROI werden pixelweise Messsignale generiert und mittels MI der statistische Zusammenhang mit dem Spirometersignal gemessen. Im Anschluss daran erfolgt die Auswahl derjenigen Regionen, die sich durch eine signifikante Beteiligung am Atemprozess auszeichnen.

Das allgemeine Vorgehen sieht zunächst eine Vorverarbeitung der Spirometersignale und der Synchronisation dieser Signale mit den Messdaten vor. Hierfür werden für jede Messung separate Masken generiert. Diese werden pro Proband zusammengefasst und schließlich für alle Probanden generalisiert. Die einzelnen Schritte sind im Folgenden detailliert beschrieben.

Vorverarbeitung

Zur Erreichung einer zeitlichen Synchronisierung von Spirometer- und Kamerasignal erfolgt zunächst ein Downsampling und anschließend ein Ausgleich der zeitlichen Verzögerung. Die Aufnahme des Volumen-Zeit-Referenz-Signals über das Spirometer erfolgte mit einer Abtastrate von 50 Hz, während die Kinect eine Abtastrate von 30 Hz verwendet. Zur Anpassung erfolgt ein Downsampling des Spirometersignals auf die Abtastrate der Kamera. Eine Zeitverschiebung zwischen beiden Signalen wird über Kreuzkorrelation bestimmt und ausgeglichen. Dies gewährleistet eine Vergleichbarkeit beider Signale im zeitlichen Kontext.

Individualisierte Maske

Für die Auswahl einer physiologischen ROI erfolgt eine Reduktion des Rechenbedarfs durch eine vorhergehende Hintergrunddetektion. MI wird anschließend zur Bestimmung von individuellen Masken eingesetzt.

Zur Detektion des Hintergrunds erfolgt zunächst eine Bestimmung des Messsignals s nach Abschnitt 4.1.1. Zusammen aus dem Mittelwert \bar{s} , der Standardabweichung $\text{std}(s)$ und einem Offset-Parameter $\beta = 0.2$ ergeben sich die Grenzen für die weiteren Betrachtungen wie folgt aus Gleichung (5.2). Alle Pixel mit dem Wert 0 gehören somit dem Hintergrund an.

$$I[x, y, t] = \begin{cases} 0, & \text{für } I[x, y, t] > (\bar{s} + \text{std}(s)) \cdot (1 + \beta) \\ 0, & \text{für } I[x, y, t] < (\bar{s} - \text{std}(s)) \cdot (1 - \beta) \\ I[x, y, t], & \text{sonst} \end{cases} \quad (5.2)$$

Anschließend wird pixelweise das Messsignal $s_{x,y}$ gebildet. Sollte zu einem beliebigen Zeitpunkt das Messsignal Hintergrund enthalten, wird der Maske M_n , dieser Messung n , für diesen Pixel der Wert 0 zugeordnet. Diese Segmentierung ermöglicht es direkt, Regionen mit geringem statistischen Zusammenhang zu kennzeichnen und folgende Berechnungen zu vereinfachen. Andernfalls erhält die Maske den Wert der MI mit dem Spirometersignal y , siehe Gleichung (5.3).

$$M_n[x, y] = \begin{cases} 0, & \text{wenn } s_{x,y}[t] = 0, \text{ für } t \in T \\ \text{MI}(s_{x,y}, y), & \text{sonst} \end{cases} \quad (5.3)$$

Abschließend erfolgt eine Normierung der MI anhand der maximalen MI für das Spirometersignal mit sich selbst, siehe Gleichung (5.4). Die MI berechnet sich dabei, wie in Abschnitt 4.1.3 angegeben, aus der Entropieabschätzung der k -nächsten Nachbarn [99].

$$M_n = M_n / \text{MI}(y, y) \quad (5.4)$$

Generalisierte Maske

Zur Vereinfachung der Berechnung wird mit der Annahme gearbeitet, dass keine Bewegungen der Probanden und insbesondere keine Positionswechsel zwischen aufeinanderfolgenden Messungen stattfinden. Diese Annahme erlaubt die Betrachtung von statischen Messsignalen, welche nicht in jedem Bild getrackt werden müssen. Somit lassen sich die Masken M_n der einzelnen Messungen N_p jedes Probanden p zu einer Probandenmaske M_p mitteln, siehe Gleichung (5.5).

$$M_p = \frac{1}{N_p} \sum_{n=0}^{N_p} M_n \quad (5.5)$$

Analog werden die Körpergelenke K_i einer Messung über eine Mittelwertbildung, der durch die Kinect bestimmten Körpergelenke, in allen Bilder gemittelt.

Diese erlauben die Transformation der Masken zwischen verschiedenen Probanden. Eine affine Transformationsmatrix T , welche es erlaubt, die Körpergelenke der Probanden aufeinander abzubilden, wird über das least-median-squares-Verfahren bestimmt, siehe Gleichung (5.6). K_i sind die Körpergelenke für einen Probanden, K'_i die entsprechenden Körpergelenke im zu matchenden Probanden. Es werden $n = 7$ Körpergelenke erkannt:

- Linke Schulter
- Rechte Schulter
- Wirbelsäule auf Schulterhöhe
- Wirbelsäulenbasis
- Linke Hüfte,
- Rechte Hüfte und
- Mittlere Wirbelsäule.

$$T = \underset{T}{\operatorname{argmin}} \sum_{i=1}^n \operatorname{med} (\|TK_i - K'_i\|_2) \quad (5.6)$$

Nach der Anwendung dieser Transformationsmatrix erfolgt die Mittelung aller Patientenmasken M_p zu einer Gesamtmaske M_{Gesamt} , siehe Gleichung (5.7) und beispielhaft Abbildung 5.2. N_P ist hier die Anzahl aller Probanden.

$$M_{\text{Gesamt}} = \frac{1}{N_P} \sum_{n=0}^{N_P} M_p \quad (5.7)$$

Es liegt nun eine Maske vor, welche den Beitrag zur Atmung gemessen anhand der MI anzeigt. Das Ziel einer ROI ist, ausschließlich Bereiche mit starker Signalqualität auszuwählen. Aus diesem Grund erfolgt eine Segmentierung der Maske anhand eines Histogramm-Schwellwerts. Die zugrunde liegende Annahme ist, dass verschiedene Bereiche mit unterschiedlich starker Mitwirkung am physiologischen Atemprozess beteiligt sind. Über die Peak-Prominence Methode [96] wird aus dem Histogramm das lokale Minimum θ mit höchster MI detektiert und anschließend als Schwellwert für die Segmentierung verwendet, siehe Gleichung (5.8) und Abbildung 5.2. Somit entsteht die generalisierte Maske $M_{\text{General, Tresh}}$.

$$M_{\text{General, Tresh}}[x, y] = \begin{cases} 0, & \text{für } M_{\text{Gesamt}}[x, y] < \theta \\ M_{\text{Gesamt}}[x, y], & \text{sonst} \end{cases} \quad (5.8)$$

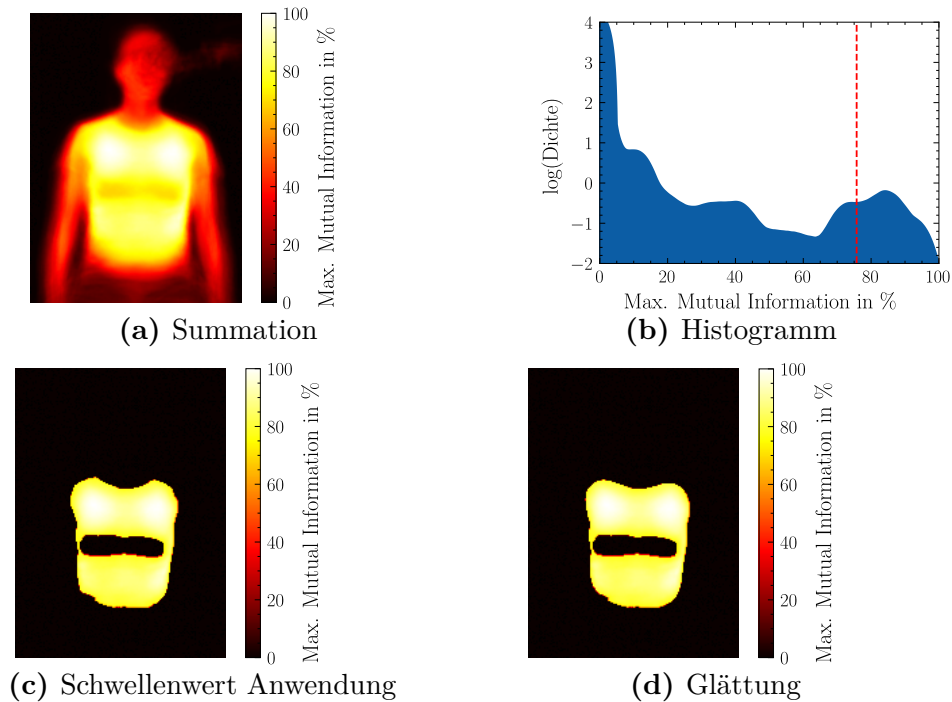


Abbildung 5.2: Vorgehen bei der Entwicklung einer generalisierten Maske. (a) Summation individueller Patientenmasken zu einer Gesamtmaske. (b) Schwellenwertbestimmung im Histogramm zur Bestimmung einer generalisierten Maske. Der Schwellenwert (rote Linie) wird als lokales Minimum mit dem höchsten Wert gewählt. Die Pixelintensitäten entsprechen dem statistischen Zusammenhang mit dem Spirometersignal, gemessen über Mutual Information und normiert auf 100. (c) Gesamtmaske nach Anwendung des Schwellenwerts. (d) Geglättete Gesamtmaske durch Snake-Algorithmus.

Ein solches globales Schwellenwertverfahren berücksichtigt keinerlei Nachbarschaftsverhältnisse. Diese können jedoch im physiologischen Prozess von Bedeutung sein. Um zu vermeiden, dass innerhalb einer homogenen Region einzelne Pixel aufgrund von statistischen Schwankungen ausgeschlossen werden und um die Kanten der Maske abzuglätten, wird die Kontur der Maske gebildet. Dies geschieht anhand des Snake-Algorithmus [108]. Die Kontur passt sich dabei an die Kanten der Segmentierung an, um eine minimale Energie zu erreichen. Die Energiefunktion setzt sich dabei aus innerer und äußerer Energie zusammen, wobei die innere Energie die Elastizität der Kontur und die äußere Energie die Anziehungskraft auf die Kanten des Objekts beschreiben. Es entsteht eine geglättete Maske. Die so abschließend erhaltene, generalisierte Maske wird als M_{General} bezeichnet.

5.1.2 Long-Short-Term-Memory-Modelle

Die Anwendung, der in Abschnitt 5.1.1 entwickelten Maske M_{General} auf die Eingangsbilder $I[x, y, t]$, ermöglicht es, ein Messsignal s zu erhalten, siehe Gleichung (5.9). X und Y sind die Zeilen, respektive Spalten im Bild.

$$s[t] = \frac{1}{X \cdot Y} \sum_{x=0}^X \sum_{y=0}^Y I[x, y, t] \cdot M_{\text{General}}[x, y] \quad (5.9)$$

Aus diesem Signal wurden in Kapitel 4 Merkmale extrahiert, welche auf die Respirationsparameter Tidalvolumen und Vitalkapazität schließen lassen. LSTM-Modelle sind in der Lage, längere zeitliche Abhängigkeiten in den Daten zu modellieren und ebenfalls mit nicht-stationären Daten umzugehen. Zum einen zeigte das vorherige Kapitel 4 bereits die Zusammenhänge des Signals mit den Messgrößen. Zum anderen wurde auch die zeitliche Abhängigkeit sowie die Nicht-Stationarität bei der forcierten Atmung verdeutlicht. Die Genauigkeiten gegenüber der Ruheatmung unterscheiden sich signifikant. Durch diese Eigenschaften sind LSTM prädestiniert für den Einsatz zur Approximation des Volumen-Zeit-Signals anhand der aufgenommenen Messsignale.

Erzeugung der Eingangsdaten

Zur Erzeugung der Eingangsdaten wird das vollständige Eingangssignal in Abschnitte mit einer Fenstergröße von W unterteilt. Diese Fenster werden um je ein Sample verschoben, sodass insgesamt $N - W + 1$ Abschnitte entstehen, wobei N die Anzahl der Samples des Messsignals ist. Aufgrund des abrupten Endes der Signale nach der forcierten Ausatmung, kann die regressive Bestimmung des mittleren Werts im Messfenster des synchronen Spirometersignals dazu führen, dass das Ende der Ausatemphase nicht bestimmt werden kann. Somit wäre eine Bestimmung der Vitalkapazität nicht möglich. Daher wird jeweils das letzte Sample des Fensters als Grundwahrheit für diesen Abschnitt verwendet. Dies führt dazu, dass es zu den ersten $W - 1$ Samples kein entsprechender Ausgang bestimmt werden kann. Dieser zeitliche Abstand entspricht dem Zeitraum, in dem das LSTM aus Merkmalen lernen kann. Verdeutlicht wird dieser Zusammenhang in Abbildung 5.3.

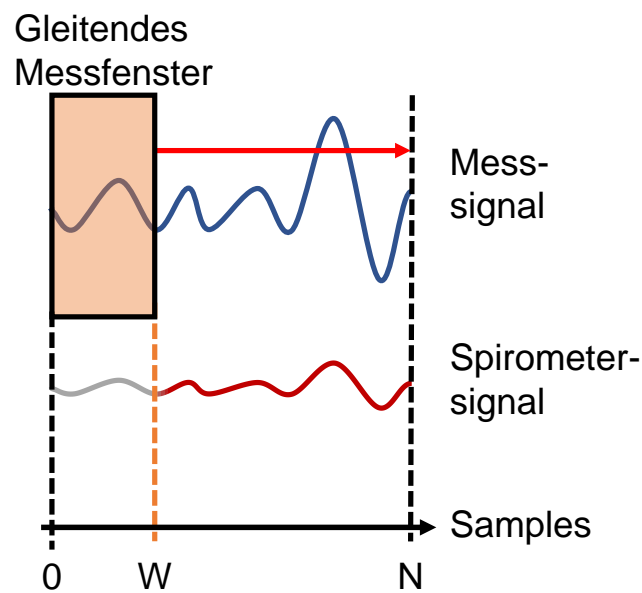


Abbildung 5.3: Signalunterteilung für rekurrente Neuronale Netze. Ein gleitendes Messfenster (Breite W) wird über das Messsignal geschoben. Die Werte innerhalb des Messfensters werden als Eingang in das Modell verwendet, während der aktuelle Wert am Ende des Messfensters im Spirometersignal als Grundwahrheit verwendet wird.

Hyperparameter des Modells

Wie bei allen Modellen des maschinellen Lernens, gibt es bei LSTM-Modellen verschiedene Hyperparameter, die die Leistung und das Verhalten des Modells beeinflussen. Einige der relevanten Hyperparameter für LSTM-Modelle sind die Anzahl der Neuronen im Netzwerk, die Anzahl der Schichten und die Lernrate [52].

Die Anzahl der Neuronen und Schichten beeinflussen die Kapazität des Modells und seine Fähigkeit, komplexe Sequenzmuster zu lernen. Eine höhere Anzahl von Neuronen und Schichten kann jedoch auch zu Overfitting führen, wenn das Modell zu komplex für die gegebene Datenmenge ist. Die Lernrate bestimmt, wie schnell das Modell während des Trainings lernt. Die Dropout-Rate ist ein Regularisierungsparameter, der das Overfitting des Modells reduziert, indem er zufällig Neuronen während des Trainings deaktiviert.

Der Suchraum der zu optimierenden Parameter des LSTM im Grid-Search-Verfahren beschränkt sich auf die in Tabelle 5.1 dargestellten Parameter. Diese

Parameter beziehen sich auf die Anwendung eines LSTM in einer eigenen Vorstudie [109]. Dabei wurden zehn bekleidete Probanden mit einer Orbecc Astra Pro Kamera aufgezeichnet und Atemsignale mit einer rechteckigen ROI extrahiert. Die LSTM-Modelle mit den Parametern aus Tabelle 5.1 schneiden bei einer Leave-One-Subject-Out-Kreuzvalidierung besser ab, als andere ML-Modelle. Mit der geringen Anzahl an Neuronen bei insgesamt drei Schichten soll ein Overfitting vermieden werden.

Tabelle 5.1: Verwendete Hyperparameter des Long Short-Term Memory Modells.

Parametername	Werte
Neuronen	[5, 10, 20]
Fenstergröße	[8, 30]
Aktivierung LSTM	[TanH]
Dropout LSTM	[0]
Aktivierung Dense	[ReLu]
Aktivierung Ausgang	[Linear]
Loss	[Mittlerer-Quadrierter-Fehler]
Lernrate	[0.005]
Epochen	[40]

5.1.3 Merkmalsbasierte Modell

Die Auswirkungen der Signalextraktion durch generalisierte, physiologische Masken ist hinsichtlich der bereits entwickelten Modelle des ML zu untersuchen. Die Ergebnisse in Kapitel 4 zeigen die Überlegenheit der ML-Modelle gegenüber linearen Modellen bei einer konventionellen Extraktion des Messsignals über eine rechteckige ROI auf dem Brustkorb der Probanden. Zur Bewertung der Auswirkungen der neu entwickelten Maske, werden die daraus extrahierten Signale als Eingang in die entwickelten Modelle gegeben und mit den bisherigen Ergebnissen aus Kapitel 4 verglichen.

Dafür werden wie zuvor 379 Merkmale extrahiert und eine Merkmalsselektion vorgenommen. Als Modell wird SVR verwendet, welches von den Modellen die beste Genauigkeit erreichte. Für eine genauere Beschreibung der Merkmalsauswahl (Abschnitt 4.1.2), Merkmalsselektion (Abschnitt 4.1.3) und Modellbildung (Abschnitt 4.1.4) sei auf die entsprechenden vorherigen Abschnitte verwiesen.

5.1.4 Evaluierungsstrategie

In einem ersten Schritt werden individuelle Masken für jeden Probanden erzeugt. Eine generalisierte Maske wird über Kreuzvalidierung anschließend für jeden Probanden aus den Masken der anderen Probanden erzeugt. Diese generalisierte Maske ist die Grundlage zur Extraktion der Messsignale. Diese Messsignale dienen als Eingang für die LSTM und SVR. Die Evaluierung der Modellarchitekturen erfolgt anhand geschachtelter Kreuzvalidierung, siehe bereits Abschnitt 4.1.5. Das Vorgehen ist zusammenfassend in Abbildung 5.4 verdeutlicht.

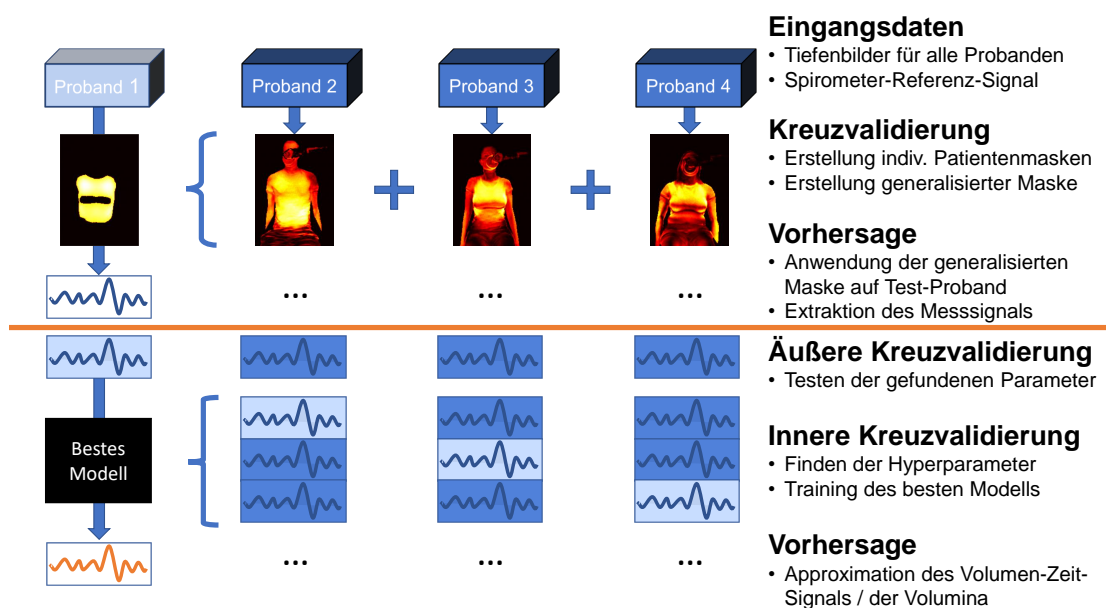


Abbildung 5.4: Validierung der Signalextraktion (oben) und Modelle des Maschinellen Lernens (unten) über Kreuzvalidierungstechniken. Über Leave-One-Out-Kreuzvalidierung werden die Messsignale für jeden Probanden erhoben. ML-Modelle verwenden diese Messsignale zur Approximation des Volumen-Zeit-Signals. Die Modellerstellung erfolgt über geschachtelte Kreuzvalidierung. Testdaten jeweils in hell- und Trainingsdaten in dunkelblau.

Als Datengrundlage wird erneut der Datensatz von Soleimani, Mirmehdi und Damen [94] verwendet. Für eine detaillierte Beschreibung sei auf die vorherige Ausführung in Abschnitt 4.2 verwiesen. An dieser Stelle werden jedoch die niedrigsten und höchsten 2,5 % der Daten für die LSTM-Modelle behalten. Dies begründet sich darin, dass das Modell an dieser Stelle nicht die reinen Messwerte lernen muss, sondern lediglich den Zusammenhang mit den Spirometerdaten. In der Literatur verwendete Skalierungsfaktoren wiesen bereits zwischen Messungen eines Probanden eine hohe Varianz auf und sind somit nicht alleine auf die Größe der Respirationsparameter zurückzuführen.

Die Bewertung des Erfolgs der Methode wird differenziert hinsichtlich der Signalqualität aus der physiologischen Maske und der resultierenden Extraktion der Respirationsparameter des LSTM-Modells. Das extrahierte Messsignal der entwickelten Maske wird mit dem Messsignal einer quadratischen ROI, wie sie in Abschnitt 4.1.1 beschrieben ist, verglichen. Als Referenz dient das Volumen-Zeit-Signal des Spirometers. Als Vergleichswerte werden die MI, der Pearson-Korrelationskoeffizient r und die Distanz d für die dynamische Zeitnormierung (DTW, dynamic time warping) zum Referenzsignal gebildet. Diese Metriken sollen dabei insgesamt ein breites Spektrum an Interpretationen abdecken. Es ist zu beachten, dass das Messsignal und das Volumen-Zeit-Signal zu diesem Zeitpunkt in anderen Maßstäben vorliegen. Die Einheit des Messsignals beträgt Millimeter, während das Volumen-Zeit-Signal in Litern gemessen wird. Aus diesem Grund kommen weitere Metriken wie zum Beispiel der mittlere quadratische Fehler nicht in Frage. Die MI hingegen gibt den statistischen Zusammenhang an und zeigte sich bereits als geeignet in der Merkmalsauswahl in Kapitel 4. Da die MI jedoch auch bereits für die Erstellung der Maske verwendet wurde - durch die Kreuzvalidierung nur für die anderen Probanden, nicht für den entsprechenden Probanden - kann sie als alleinige Bewertungsmetrik zu falschen Schlussfolgerungen führen. Daher wird des Weiteren der Pearson-Korrelationskoeffizient verwendet, welcher den linearen Zusammenhang untersucht. Darüber hinaus wird die DTW eingesetzt, um zusätzlich physiologische Unterschiede zwischen Thoraxbewegung und Atemvolumina zu berücksichtigen. Diese Metrik wird ebenfalls von Soleimani u. a. [85] eingesetzt. Die DTW berechnet eine optimale Pfadzuordnung zwischen den zwei Signalen, wodurch die Distanz zwischen ihnen minimiert wird. Sie wird oft bei der Analyse von Zeitreihendaten wie Spracherkennung, Musikerkennung und

Bewegungserkennung eingesetzt. Die DTW hat den Vorteil, dass sie in der Lage ist, Zeitreihen mit unterschiedlichen Skalierungen und Geschwindigkeiten zu vergleichen, indem sie die Signalform anpasst. Sie kann auch mit Daten umgehen, die zeitlich verzerrt oder unregelmäßig gesampelt wurden. Die DTW hat jedoch auch einige Nachteile hinsichtlich der hohen benötigten Rechenleistung. Es ist darüber hinaus schwierig, die DTW-Distanz zu interpretieren und zu visualisieren, da sie auf einem optimierten Pfad basiert, der nicht immer intuitiv ist. Die DTW-Distanz D wird anhand Gleichung (5.10) bestimmt, wobei X und Y die zwei Sequenzen mit der Länge n und m sind, x_i und y_j die i -te bzw. j -te Komponente der Sequenz sind und $w(i, j)$ die Gewichtungsfunktion ist, die den Beitrag der i -ten Komponente in der ersten Sequenz und der j -ten Komponente in der zweiten Sequenz zur Gesamtdistanz bestimmt. Die Optimierung der Gewichtsmatrix erfolgt nach den Ausführungen von Senin [110].

$$D_{DTW}(X, Y) = \sqrt{\sum_{i=1}^n \sum_{j=1}^m w(i, j) \cdot (x_i - y_j)^2} \quad (5.10)$$

Mit einem einseitigen verbundenen t-Test werden die Metriken anschließend auf Signifikanz geprüft. Das Signifikanzniveau wird auf 95 % festgelegt. Die Nullhypothese H_0 gibt an, dass die Metrik (MI und r) der physiologischen Maske zum Referenzsignal kleiner oder gleich des absoluten Fehlers der quadratischen ROI beträgt. Entgegengesetzt besagt die Alternativhypothese H_1 , dass die Metrik der physiologischen Maske größer als die der quadratischen ROI ist, siehe Gleichungen (5.11a) und (5.11b). Für die DTW verhalten sich H_0 und H_1 entgegengesetzt, da eine kleinere DTW als besser angesehen wird.

$$H_0 : r_{\text{quadratischeROI}} \geq r_{\text{physiologischeMaske}} \quad (5.11a)$$

$$H_1 : r_{\text{quadratischeROI}} < r_{\text{physiologischeMaske}} \quad (5.11b)$$

5.2 Ergebnisse und Diskussion

In diesem Abschnitt werden die Ergebnisse dargestellt und diskutiert. Dabei wird die Vorverarbeitung durch die Erstellung von individuellen und generalisierten Masken mit konventionellen quadratischen Regionen zur Signalextraktion verglichen. Dabei wird zum einen auf das Aussehen der Masken und mögliche physiolo-

gische Interpretationen eingegangen. Zum anderen wird die direkte Signalqualität anhand von ausgewählten Metriken gegenübergestellt.

Daraufhin werden die Auswirkungen der neuartigen Signalextraktion untersucht, indem die Ergebnisse einer direkten Approximation der Volumen-Zeit-Kurve durch LSTM präsentiert werden. Abschließend erfolgt eine Evaluierung der Resultate der Respirationsparameterbestimmung, wobei die extrahierten Signale als Eingangsgrößen in die zuvor konzipierten Modelle einfließen und mit den früheren Ergebnissen aus Abschnitt 4.3 in einen Vergleich gesetzt werden.

5.2.1 Patientenmasken

In diesem Abschnitt werden die individuellen Masken der einzelnen Messungen zusammenfassend dargestellt und vergleichend analysiert. Im Folgenden wird das Aussehen der generalisierten Maske über alle Probanden diskutiert. Die Signale, welche anhand dieser Masken generiert werden, werden den Signalen einer herkömmlichen ROI gegenüber gestellt.

Unterschiede in den individuellen Masken

Allgemein betrachtet, sind in allen individuellen Masken die anatomischen Strukturen der Probanden zu erkennen, siehe auch Abbildung 5.5. Dies umfasst Kopf, Arme, Rumpf und Beine. Des Weiteren ist das gestützte und am Mund fixierte Spirometer zu erkennen. Aus dem Hintergrund sind keine Strukturen zu erkennen. Dies deutet darauf hin, dass die Hintergrunddetektion funktioniert und erfolgreich einen Schwellenwert zwischen Messperson und Hintergrund festlegen kann. Da das Spirometer unmittelbar am Mund angebracht wird, kann es über die Distanz nicht ausgeschlossen werden. Es zeigt sich, dass die MI an der Position des Spirometers ($< 20\%$) deutlich niedriger ist, als die MI des Oberkörpers ($> 40\%$). Der geringe Informationsbeitrag des Spirometers lässt sich zum Teil auf stochastische Prozesse zurückführen, wie zum Beispiel die Messungenauigkeiten der Tiefenkamera. Eine weitere Erklärung kann in der geringen Bewegung des Kopfes bei der Atmung gefunden werden. Diese überträgt sich auf das Spirometer. Der aufgewiesene statistische Zusammenhang zwischen den Armbewegungen und dem Spirometersignal ist auf indirektem Weg zu erklären. So führt die Anhebung des Thorax und die Anhebung der Schultern zu einer Anhebung der Arme und

somit zu einer synchronen Bewegung, welche sich im statistischen Zusammenhang durch die MI zeigt. Mit statistischen Schwankungen und geringen Übertragungen von Bewegungen sind auch die geringen Zusammenhänge zu erklären, die im Bereich der Oberschenkel der Probanden zu erkennen sind. Diese Beiträge variieren zwischen den Probanden und sind nicht gleichermaßen konstant in allen Messungen. Dem entgegen stehen die Beiträge des Oberkörpers, welche in allen Messungen auftreten und wie erwartet durch die Atmung hervorgerufen werden. Auffällig ist, dass nur sehr kleine Regionen mit sehr starker MI ($> 90\%$ weiß) erkennbar sind. Regionen im Oberkörper weisen weitestgehend eine hohe MI ($> 70\%$ gelb) auf. Dies ist in Übereinstimmung mit den abgeleiteten Ergebnissen aus der Literatur. So wurden in vorherigen Studien rechteckige ROI auf dem Brustkorb verwendet und eine starke Korrelation mit dem Spirometersignal berichtet [62, 71].

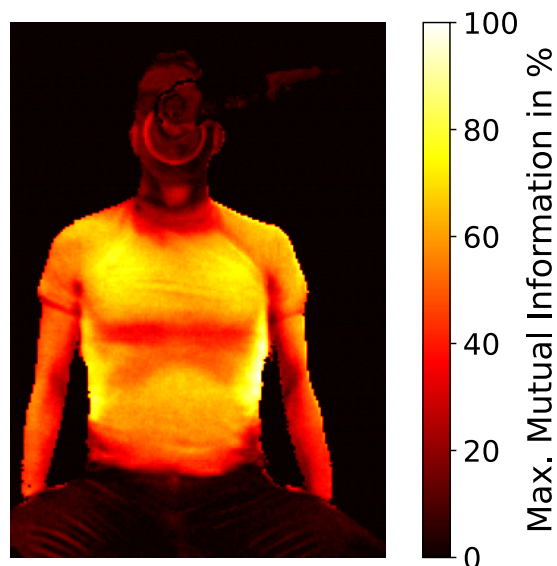


Abbildung 5.5: Beispielhafte Darstellung einer über Mutual Information erstellten Maske für Proband 5 Messung 39.

Eine explorative, visuelle Betrachtung der Oberkörper in den individuellen Masken lässt es zu, weitere Unterschiede in den hervorgehobenen Regionen der Heatmaps zu gruppieren. Aufgrund der großen Anzahl an Messungen wird an dieser Stelle nur auf einzelne Beispiele verwiesen, die diese Gruppen beispielhaft darstellen.

So sind bei einzelnen Messungen folgende Besonderheiten erkennbar:

- ein ausgeprägter Rippenbogen,
- eine stärkere Ausprägung des Bauchs im Gegensatz zum Brustkorb,
- eine Kantenbildung im Brustbereich und
- Faltenbildung durch Kleidung.

In Abbildung 5.6 werden beispielhaft die Masken dargestellt, die einen ausgeprägten Rippenbogen aufweisen. Bei diesen Messungen weisen die unteren Rippen einen starken Zusammenhang mit dem Spirometersignal auf. Bei den Messungen 150, 256, 265 und 291 weisen jeweils alle Rippen und somit der komplette Thorax einen großen statistischen Zusammenhang ($> 80\%$ der max. MI) mit dem Referenzsignal auf. Wenngleich keine anatomische Referenz verwendet werden kann, ist der Rippenbogen bei den Messungen 196 und 265 besonders stark ausgeprägt. Der resultierende Bogen aus den Messungen 150 und 291 entspricht nicht dem anatomischen Aufbau des Brustkorbs, vergleiche Abschnitt 2.1.1. Die schwächer ausgeprägte Form kann zum einen durch eine vermehrte Bauchatmung oder Oberkörperbewegungen innerhalb der Messung erklärt werden. Es lässt sich somit resultieren, dass anatomisch bedeutsame Strukturen allein durch die Beiträge zum Atemsignal deutlich gemacht werden können.

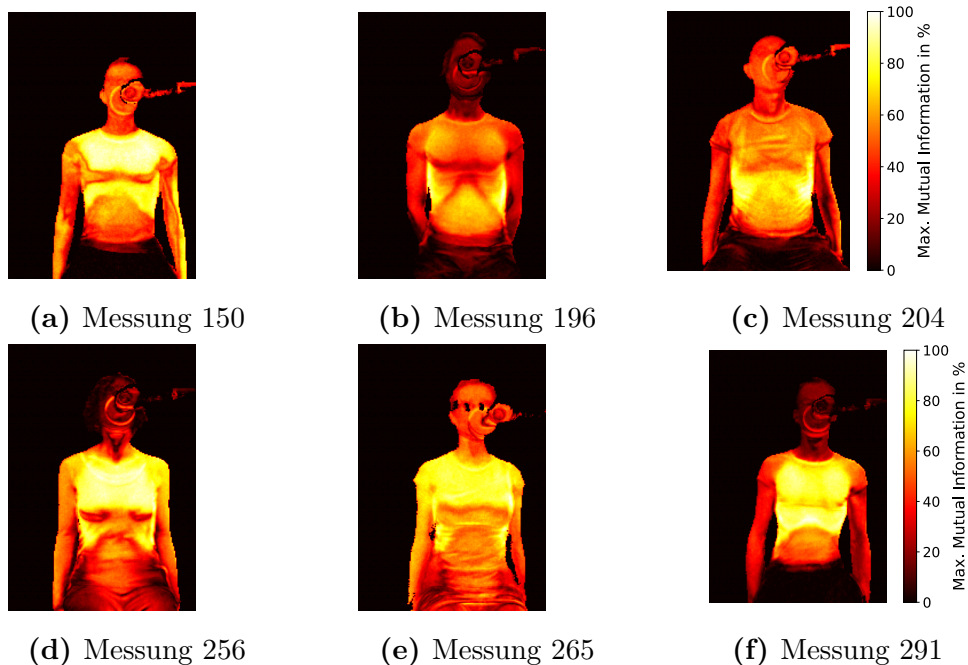


Abbildung 5.6: Individuelle Patientenmasken bei thorakaler Atmung.

Im Gegensatz zur thorakalen Atmung sind in Abbildung 5.7 stärkere Indizien auf abdominale Atmung erkennbar. Bei den Messungen 130, 178 und 230 weist das Abdomen einen stärkeren Zusammenhang mit dem Referenzsignal auf. Wie bei der thorakalen Atmung sind auch hier anatomische Regionen sichtbar.

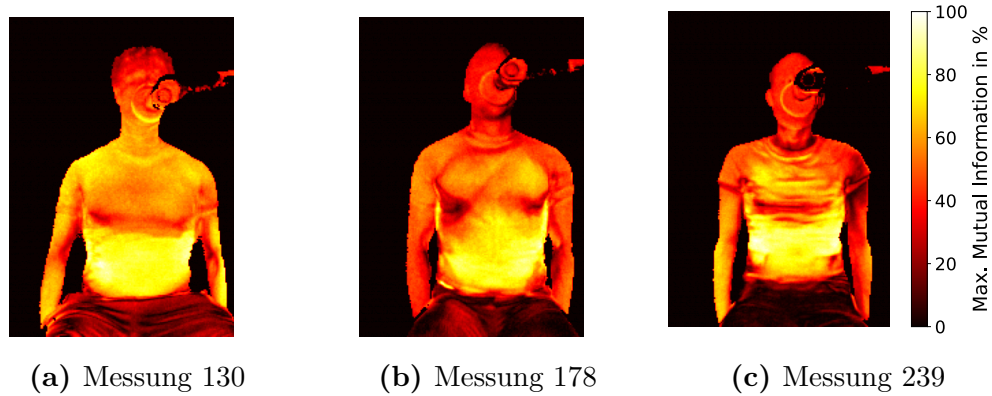


Abbildung 5.7: Individuelle Patientenmasken bei abdominaler Atmung.

Eine starke Kante im Brustbereich hingegen, weisen insbesondere die Messungen 82, 150 und 265 in Abbildung 5.8 auf. Diese zeichnet sich sowohl bei männlichen, als auch bei weiblichen Probanden durch einen horizontalen Streifen geringerer MI ($< 50\%$ der max. MI) im Vergleich zum suprathorakalen und abdominalen Bereich aus. Aufgrund der Physiologie der Atmung ist nicht anzunehmen, dass diese Region geringer an der Atmung beteiligt ist. Vielmehr können technische Ursachen der Versuchsdurchführung eine mögliche Erklärung liefern. Die Messsignale s beziehen sich auf die horizontale Abstandsänderung zwischen Kamera und Proband. Die Bewegung des Brustkorbs bei der Einatmung erfolgt in ebendiese ventrale Richtung, sowie nach kranial. Diese kraniale Bewegung ist jedoch bei der Maskenentstehung nicht abbildbar. So wurden die Signale bei festen Pixelwerten mit dem Spirometersignal verglichen. Die Thoraxmorphologie weist auf der Longitudinalachse zwei starke Gradienten auf. Dies ist die positive, sowie negative Abstandsänderung der Brust. Bewegt sich nun die Brust bei der Atmung nach oben, rückt die Brust raus aus den vorherigen Pixeln, welche somit eine Zunahme des Abstands messen. Dies ist konträr zum Spirometersignal und schlägt sich somit in einer verringerten MI wieder. Die horizontale Ausdehnung dieses Streifens kann auf die Spannung der Kleidung über dem Sternum erklärt werden. Die Kleidung liegt möglicherweise nicht auf der Haut auf. Dies führt dazu, dass diese Region einen geringen statistischen Zusammenhang mit dem Volumen-Zeit-

Signal des Spirometers aufweist und zur Extraktion des Messsignals potentiell ausgelassen werden könnte.

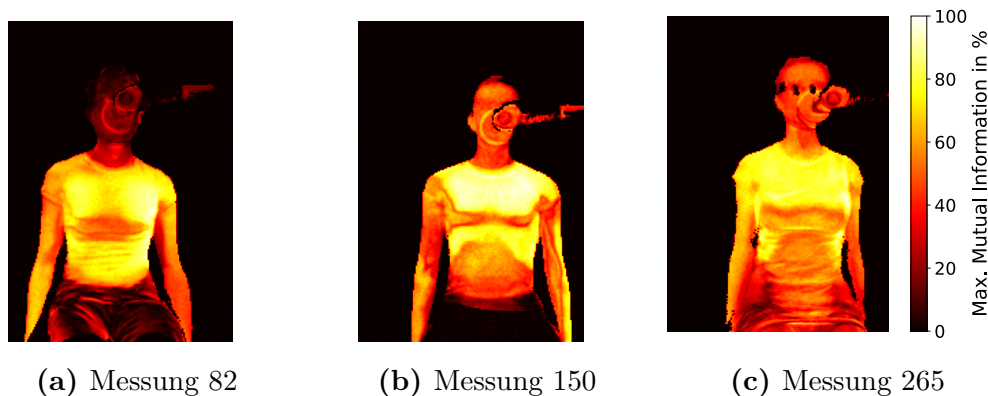


Abbildung 5.8: Individuelle Patientenmasken mit ausgeprägter Kante.

Ebendiese Faltenbildung lässt sich auch bei den Messungen 62, 180 und 222 in Abbildung 5.9 ausgeprägt feststellen. In der Heatmap sind klare Falten ($< 60\%$ der max. MI auf Hintergrund mit $> 80\%$ der max. MI) der Bekleidung zu erkennen. Dies verdeutlicht, warum in bisherigen Messungen in der Literatur die Probanden gebeten wurden, eng anliegende Kleidung zu tragen [61]. Diese Kleidungs falten führen zu einer Verschlechterung des Messsignals.

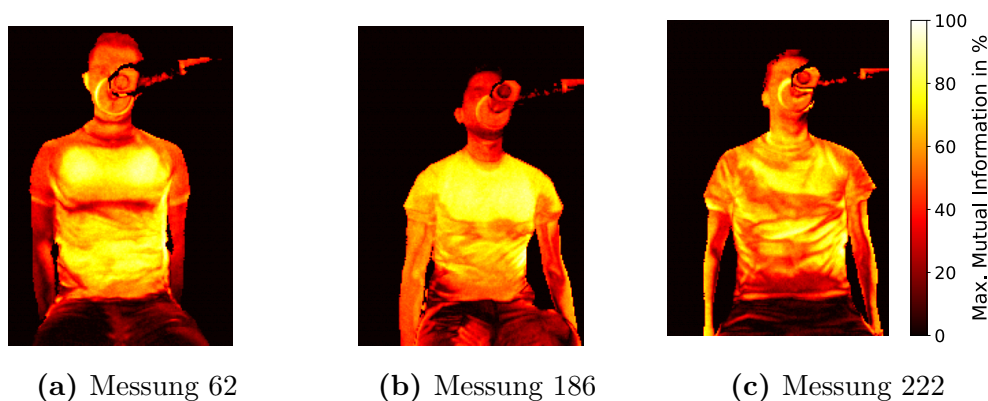


Abbildung 5.9: Individuelle Patientenmasken bei faltenschlagender Kleidung.

Zusammenfassend lässt sich sagen, dass individuelle Masken für die einzelnen Messungen die anatomische Struktur der Atemmechanik widerspiegeln. Sie weisen daher eine neuartige Möglichkeit auf, die Beiträge der Atmung in Form einer Heatmap zu visualisieren. Ein solches diagnostisches Mittel könnte den Diagnoseprozess von Krankheiten zusätzlich unterstützen. Insbesondere bei restriktiven

Erkrankungen können die Atemmuster optisch dargestellt und disfunktionale Regionen identifiziert werden. Im Vergleich zu anderen Methoden, wie der OEP [81] sind die zu untersuchenden Regionen nicht durch die vorherige Anbringung von Markern beschränkt. Es ist jedoch festzuhalten, dass die oben beschriebene Analysen des Atemstils (thorakale oder abdominale Atmung) sich für diesen Datensatz nicht quantifizieren lassen, da keine entsprechende Information gegeben ist. Die entwickelte Methode zur Identifikation geeigneter Oberkörperregionen, mit einem hohen statistischen Beitrag am Volumen-Zeit-Verlauf der Atmung durch den Einsatz von MI, setzt das Vorhandensein eines Referenzsignals eines Spirometers voraus.

Physiologische Prozesse bei der generalisierten Maske

Der vornehmliche Untersuchungsgegenstand dieser Arbeit stellt die Zusammenführung der Erkenntnisse zu den individuellen Regionen dar und ob sie sich zu einer neuen, allgemeinen ROI verallgemeinern lassen. Dafür werden die einzelnen, individuellen Masken summiert und ein optimaler Schwellenwert zur Segmentierung festgelegt. Die Kanten der finalen, generalisierten Maske werden anschließend geglättet. Die generalisierte Maske ist in Abbildung 5.10 dargestellt. Zum besseren Abgleich mit anatomischen Strukturen sind zusätzlich die relevanten Körpergelenke durch die Kinect-Kamera abgebildet.

Es ist zu erkennen, dass weitere Körperpartien, wie die Beine, Arme und der Kopf nicht mit in der Maske enthalten sind (0 % der max. MI) und genauso wie das Spirometer, aufgrund einer geringeren MI, aus der generalisierten Maske entfernt wurden. Somit verbleibt als Region, mit maßgeblichen Beitrag zur Atmung, die Thoraxregion erhalten (> 70 % der max. MI). Dabei sind Beobachtungen zutreffend, die bereits für die Einzelmasken getroffen wurden. Tendenziell weist der obere Bereich den stärksten statistischen Zusammenhang mit dem Spirometer-signal auf. Besonders die lateralen Brustbereiche weisen einen stärkeren Beitrag (bis zu 100 % der max. MI) auf. Einen geringeren (> 80 %) medialen Beitrag im Bereich des Brustbeins lässt sich ebenfalls mit der Auflagefläche durch Kleidung erklären. Gleiches ist auf die ausgesparte, untere Brustkante zutreffend. Durch einen zu geringen Beitrag, im Vergleich zu den anderen Regionen, wurde dieser Bereich im automatischen Schwellenwertverfahren ausgespart. Im oberen Thoraxbereich verläuft die Grenze der Maske bogenförmig und folgt somit dem Rippen-

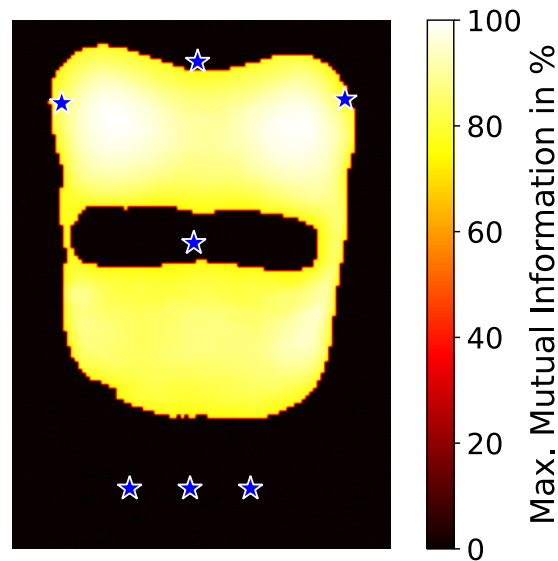


Abbildung 5.10: Verallgemeinerte Maske mit Körpergelenken (blaue Sterne) durch Kinect. Körpergelenke von links oben nach rechts unten: rechte Schulter, Wirbelsäule auf Schulterhöhe, linke Schulter, mittlere Wirbelsäule, rechte Hüfte, Wirbelsäulenansatz, linke Hüfte.

verlauf, mit Auslassung des Brustbeins. Ein Verlauf der Maske am Rippenbogen war bereits bei den Einzelmasken zu erkennen, zeigt sich in der verallgemeinerten Maske jedoch nicht. Die Mittelung über alle Probanden kann die einzelnen Effekte verwischen. An den Körpergelenken der linken und rechten Hüfte zeigt sich dennoch, dass ein subabdominaler Bereich in der Maske ausgespart wurde.

Signalqualitäten

Neben der visuellen Inspektion und explorativen Analyse der Masken steht die daraus erhobene Signalqualität im Vordergrund der Untersuchungen dieser Arbeit. Es wurde über LOOCV eine generalisierte Maske aus den einzelnen Messungen erstellt und für jeden Probanden angewendet. Das somit extrahierte Signal wird anhand verschiedener Metriken mit einem Signal aus einer quadratischen ROI auf dem Thorax verglichen.

Die statistische Übereinstimmung zwischen dem Signal der generalisierten Maske und dem Spirometersignal ist höher im Vergleich zum Signal der konventionellen ROI, was durch die MI-Metrik ($p < 0.0001$) belegt wird. Auch wenn durch die LOOCV die Spirometerdaten des getesteten Probanden nicht

in die Entwicklung der generalisierten Maske eingehen, ist die MI das Maß zur Entwicklung ebendieser generalisierten Maske. Dementsprechend ist es in Anbetracht der zugrundeliegenden Entwicklungslogik der generalisierten Maske nicht unerwartet, dass eine statistisch signifikante Verbesserung für die entsprechende Metrik zu verzeichnen ist. Ebenso ist der Pearsons-Korrelationskoeffizient für die generalisierte Maske höher ($p < 0.005$), was ebenfalls auf eine höhere Übereinstimmung hinweist. Die DTW-Distanz ist für die Signale einer generalisierten Maske geringer ($p < 0.0001$), wobei dies einem besseren Ergebnis entspricht.

Das Messsignal einer generalisierten Maske ist dem Messsignal einer konventionellen Maske hinsichtlich der MI, der Pearson-Korrelation und der DTW-Distanz überlegen. Im Detail ist somit der Informationsgehalt des extrahierten Signals (MI) höher, der lineare Zusammenhang mit dem Spirometersignal (Pearson-Korrelation) erhöht und die Signalform (DTW) exakter. Eine Maske, welche über MI spezielle Regionen auswählt, welche einen hohen statistischen Zusammenhang mit einem Referenzspirometer aufweisen, führt zu einem besseren Messsignal, im Vergleich zu einer konventionellen, quadratischen ROI auf dem Thorax. Die, über das in dieser Arbeit entwickelte Verfahren, erhaltene Masken sind auf andere Probanden übertragbar.

5.2.2 Auswirkungen auf die Modelle des Maschinellen Lernens

Ziel einer verbesserten Signalextraktion durch generalisierte Masken ist es, den ML-Modellen genauere Eingangsdaten zu liefern. Die daraus folgenden Auswirkungen werden einmal für LSTM-Modelle und für die bestehenden merkmalsbasierten Ansätze aus Kapitel 4 im Folgenden dargestellt.

LSTM-Modelle

LSTM-Modelle können das Tidalvolumen mit einer Genauigkeit von 191 ± 255 ml und die Vitalkapazität mit einer Genauigkeit von 488 ± 938 ml gegenüber dem Referenzspirometer bestimmen. Diese Ergebnisse sind bedeutend ungenauer als die Modelle im vorherigen Kapitel 4. Es fällt dabei die hohe mittlere Abweichung auf, welche im Vergleich zu den vorherigen ML-Modellen von nahezu 0 ml auf mehrere hundert ml ansteigt. Dies zeigt einen systematischen Fehler der LSTM-

Modelle. Die Ergebnisse sind vergleichend in Tabelle 5.3 dargestellt.

Eine Erklärung ist in der Struktur der LSTM-Modelle zu suchen. Zum einen kann der begrenzte Satz an Hyperparametern zu einem schlechteren Ergebnis führen. Zum anderen betrachten LSTM-Modelle nur einen geringen Teil des Signals gleichzeitig. Während für RF alle extrahierten Merkmale des kompletten Signals zur Verfügung stehen, sind für das LSTM nur die Samples innerhalb der Fenstergröße verfügbar. Eine größere Fenstergröße ist für die Modelle in diesem Datensatz nicht anwendbar, da nicht ausreichend viele Samples der Atemmänner ermittelt werden könnten. So wäre die VC nicht vollständig bestimmbar, vergleiche Abbildung 5.3. Aufgrund der auftretenden Ungenauigkeiten, bereits bei den statischen Parametern, erfolgt an dieser Stelle keine weitere Betrachtung von dynamischen Respirationparametern.

Im Vergleich zu den Ergebnissen einer Voruntersuchung [109] auf einem kleineren Datensatz mit LSTM-Modellen sind diese Ergebnisse auf diesem Datensatz [94] nicht reproduzierbar. Die Unterschiede sind dem diverseren Datensatz zu zuschreiben. So liegt in dem in dieser Arbeit verwendeten Datensatz eine höhere Varianz der Probanden (Alter, Größe, Gewicht) und in den Respirationparametern (TV und VC) zu Grunde. Dies kann dazu führen, dass die Ergebnisse nicht vergleichbar sind.

Merkmalsbasierte Modelle

Die Verbesserungen in der Signalextraktion, durch die Anwendung von generalisierten Masken, hat einen positiven Einfluss auf die Bestimmung von TV und VC durch merkmalsbasierte SVR-Modelle. Alle Modelle verbessern sich in ihrer Genauigkeit (mittlere Abweichung \pm Standardabweichung) sowohl für Tidalvolumen als auch für Vitalkapazität, vergleiche Tabelle 5.2 und Tabelle 4.3 der vorherigen Ergebnisse. Vereinzelt sind die Modelle exakter ohne persönliche Merkmale der Probanden, zum Beispiel zur Bestimmung des Tidalvolumens und Merkmalsselektion durch F-Regression und rekursive Merkmalselimination. Der Fehler des besten Modells für die TV-Bestimmung kann von 11 ± 253 ml auf 8 ± 206 ml reduziert werden. Dies entspricht einer Verringerung der Standardabweichung um 18,6%. Zusätzlich werden keine persönlichen Merkmale benötigt, was potentielles Overfitting zukünftiger Modelle weiter reduzieren kann. Für die VC kann die Standardabweichung durch generalisierte Masken ebenfalls reduziert werden.

Dies ist jedoch mit einem höheren mittleren Fehler verbunden. Die beste Genauigkeit für die VC beträgt -43 ± 599 ml, im Vergleich zur Genauigkeit beim vorherigen Einsatz einer konventionellen, rechteckigen ROI mit 11 ± 624 ml, siehe auch Tabelle 5.3.

Tabelle 5.2: Genauigkeit der Support Vector Regression mit Eingangssignalen aus generalisierten Masken für Tidalvolumen (a) und Vitalkapazität (b). Zur Merkmalsselektion werden F-Regression (FR), Mutual Information (MI), modellbasierte Merkmale (MB) und rekursive Merkmalselimination (RFE) eingesetzt. Angabe der Genauigkeit als mittlerer Fehler \pm Standardabweichung.

(a) Tidalvolumen (TV)

Persönliche Merkmale	FR	MI	MB	RFE
mit	-2 ± 222 ml ^{***}	12 ± 214 ml ^{***}	7 ± 212 ml ^{***}	3 ± 217 ml ^{***}
ohne	11 ± 209 ml ^{***}	17 ± 218 ml ^{***}	6 ± 230 ml ^{***}	8 ± 206 ml ^{***}

(b) Vitalkapazität (VC)

Persönliche Merkmale	FR	MI	MB	RFE
mit	-9 ± 651 ml ^{***}	-43 ± 599 ml ^{***}	63 ± 684 ml ^{***}	36 ± 611 ml ^{***}
ohne	40 ± 801 ml	-43 ± 705 ml ^{***}	-7 ± 820 ml	20 ± 689 ml ^{***}

* Signifikant geringer ggü. Mittelwert-Regressor (TV: 0 ± 288 ml, VC: 2 ± 889 ml) $p < 0.05$.

** Signifikant geringer ggü. Mittelwert-Regressor (TV: 0 ± 288 ml, VC: 2 ± 889 ml) $p < 0.01$.

*** Signifikant geringer ggü. Mittelwert-Regressor (TV: 0 ± 288 ml, VC: 2 ± 889 ml) $p < 0.001$.

Diese Ergebnisse belegen, dass die Signalextraktion durch eine generalisierte Maske zu genaueren Ergebnissen durch ML-Modelle führt. Die generalisierte Maske ist somit einer rechteckigen ROI vorzuziehen. Es zeigt sich, dass die Auswirkungen auf die Bestimmung der Tidalatmung größeren Einfluss haben, als auf die forcierte Atmung. Eine Ursache kann in der Statik der generalisierten Maske gefunden werden, die nur einmalig für die gesamte Messung angelegt wird. Bei einer forcierten Atmung können sich durch starke Bewegung im Thorax die atmungsrelevanten Bereiche aus der Maske verschieben. Dies kann dazu führen,

dass somit nicht mehr das optimale Signal extrahiert werden kann.

Tabelle 5.3: Auswirkungen auf die Modelle des Maschinellen Lernens durch generalisierte (gen.) Masken und den Einfluss von persönlichen (pers.) Merkmalen. Ergebnisse dargestellt für existierende Modelle aus der Literatur, Support Vector Regression (SVR) und Long Short-Term Memory (LSTM) Modelle im Vergleich mit quadratischen (quadr.) Regionen von Interesse (ROI).

Modell	Tidalvolumen	Vitalkapazität
Soleimani - 1 Kamera [71]	70 ± 226 ml	-143 ± 847 ml
Soleimani - 2 Kameras [20]	-23 ± 205 ml	-318 ± 524 ml
SVR mit pers. Merkmalen und quadr. ROI	11 ± 253 ml	11 ± 624 ml
SVR ohne pers. Merkmale und quadr. ROI	18 ± 267 ml	-9 ± 737 ml
SVR mit pers. Merkmalen mit gen. Maske	7 ± 212 ml	-43 ± 599 ml
SVR ohne pers. Merkmale mit gen. Maske	8 ± 206 ml	20 ± 689 ml
LSTM mit gen. Maske	191 ± 255 ml	488 ± 938 ml

5.3 Zusammenfassung

Der Untersuchungsgegenstand dieses Kapitels ist die Verbesserung der Signalextraktion durch generalisierte Patientenmasken. Es wurde eine Methode entwickelt, um mittels MI Regionen des Körpers zu identifizieren, die einen hohen Beitrag zum Signal des Referenzspirometers aufweisen. Dies ermöglicht neue Einblicke in die Betrachtung der Dynamik des Brustkorbs ohne im Vorfeld festgelegte Regionen und Vorbereitung durch Marker. Es ist potentiell möglich Atemmechaniken wie Brustatmung und Bauchatmung zu unterscheiden. Aufgrund fehlender Informationen, lassen sich diese jedoch für diesen Datensatz nicht verifizieren. Die Erkenntnisse der ausgewählten Regionen lassen sich zusammenfassen und auf neue Probanden anwenden. Es zeigt sich, dass insbesondere eine Kante in der Brustregion nicht förderlich für die Signalextraktion ist. Mit diesen generalisierten Ansätzen lässt sich die Signalqualität signifikant erhöhen, gemessen an den drei

verschiedenen Metriken. Eine solche generalisierte Maske lässt sich automatisch anlegen und ist einer konventionellen, rechteckigen ROI vorzuziehen.

Ein weiterer Untersuchungspunkt war die Approximation eines Volumen-Zeit-Signals durch LSTM-Modelle. LSTM-Modelle zeichnen sich dadurch aus, Zusammenhänge im zeitlichen Kontext verarbeiten zu können. Da im vorherigen Kapitel 4 bereits gezeigt wurde, dass ML eingesetzt werden kann, um aus extrahierten Signalmerkmalen auf die Respirationsparameter zu schließen, ist das Ziel dieser Untersuchung die Möglichkeit der kompletten Approximation eines Volumen-Zeit-Signals durch LSTM-Modelle. Die Anwendung von LSTM-Modellen in dieser Arbeit lässt die Messung von dynamischen Respirationsparametern auf diesem Datensatz nicht zu. Eine mögliche Ursache ist im begrenzten Eingang des Modells zu sehen. Das Modell betrachtet den zeitlichen Verlauf der Eingangsdaten. Somit können jedoch die Parameter der gesamten Messung nicht gleichzeitig zur Verfügung stehen. Die Ergebnisse der LSTM-Modelle sind bedeutend ungenauer als merkmalsbasierte ML-Modelle.

Die verbesserte Signalextraktion durch generalisierte Masken führt zu einer genaueren Bestimmung von Tidalvolumen und Vitalkapazität durch die in Kapitel 4 entwickelten Support-Vector-Regression-Modelle. Es kann eine Genauigkeit von 8 ± 206 ml für das Tidalvolumen und 43 ± 599 ml für die Vitalkapazität erreicht werden. Dies verdeutlicht, dass die generalisierten Masken einer rechteckigen ROI vorzuziehen sind.

Auch wenn die gewünschte Genauigkeit durch die LSTM-Modelle nicht ausreichend ist, so zeigt die explorative Analyse der individuellen Masken einen hohen potentiellen Mehrwert in der Detektion von disfunktionalen Respirationsprozessen. Solche Masken könnten in der Klassifikation von respiratorischen Krankheiten eingesetzt werden, insbesondere in der Bestimmung von restriktiven Erkrankungen, die sich durch ebenjene disfunktionale Atemmechanik auszeichnen. Der Einsatz von generalisierten Masken durch MI ist den konventionellen, rechteckigen ROI überlegen und führt zu einer Verbesserung in der Bestimmung der Respirationsparameter.

Kapitel 6

Klassifikation von respiratorischen Erkrankungen

Die vorangegangenen Kapitel haben die Durchführbarkeit von ML-basierten Verfahren zur Bestimmung des Respirationsvolumens aus extrahierten Merkmalen (Kapitel 4) sowie die Verbesserung der Signalextraktion durch die Anwendung einer physiologischen ROI (Kapitel 5) demonstriert. Die Untersuchung der ausgewählten ROI zeigt, dass verschiedene Thoraxregionen unterschiedliche Beiträge zur Extraktion des Messsignals liefern und Einblicke in die Mechanik der Atembewegung ermöglichen. Die entsprechenden Masken wurden in Abschnitt 5.2 veranschaulicht. Eine quantifizierte Zuordnung zu thorakaler und abdominaler Atemdynamik ist jedoch aufgrund fehlender Daten derzeit nicht möglich.

Eine asynchrone Atemdynamik kann eine Folge einer respiratorischen Erkrankung, wie zum Beispiel COPD sein [92]. Die Beiträge von Rippen und Abdomen zum Gesamtvolumen können sich dabei im Vergleich zu einer gesunden Kontrollgruppe signifikant unterscheiden [91]. Es wird in diesem Kapitel untersucht, ob DPG in Verbindung mit KI in der Lage ist, respiratorische Erkrankungen zu klassifizieren.

ML-Modelle sind in der Lage, relevante Bereiche und Merkmale autonom zu extrahieren und daraus Zielgrößen abzuleiten, wie bereits in Kapitel 4 gezeigt. In weiteren Beispielen werden CNN zur Extraktion der Respirationsrate aus Farbbildern ohne Tiefeninformation [111] oder zur Klassifikation von Lungenentzündungen auf Röntgenbildern [112] eingesetzt. Bisherige Untersuchungen in der Literatur zur Klassifikation von respiratorischen Erkrankungen über DPG verwen-

den zuvor extrahierte respiratorische Merkmale, wie zum Beispiel FEV_1 [43, 56, 61, 80]. Dabei werden die Möglichkeiten der DPG zur Extraktion von spatio-temporalen Daten nicht vollständig ausgenutzt. In dieser Arbeit soll daher untersucht werden, welche Modalitäten der DPG zur Klassifikation geeignet sind.

In diesem Kapitel stehen folgende Fragestellungen im Fokus:

- Wie kann DPG in Verbindung mit KI genutzt werden, um respiratorische Erkrankungen zu klassifizieren?
- Inwiefern liefern die, aus den extrahierten Masken gewonnenen, physiologischen Informationen eine solide Grundlagen für die Klassifikation von respiratorischen Erkrankungen?

Zur Beantwortung dieser Fragestellungen ist der bestehende Datensatz [94] aus der Literatur nicht ausreichend, da keine kranken Personen inkludiert sind. Auch, wenn anderweitig Aufnahmen mit Erkrankten existieren, so sind diese nicht öffentlich zugänglich [43, 56, 61, 80]. Aus diesem Grund wird für die Untersuchung in diesem Kapitel ein neuer Datensatz aufgenommen. Zur Klassifikation werden Methoden des ML eingesetzt und die Eignung verschiedener Datenmodalitäten der DPG analysiert.

6.1 Methodik

Die Hypothese, dass eine Klassifikation von respiratorischen Erkrankungen über DPG möglich ist, basiert auf den Erkenntnissen aus Kapitel 5. Die individuelle Patientenmaske ermöglicht eine differenzierte Betrachtung der Beiträge zur Atemmechanik und die Signalextraktion, aus einer generalisierten Maske, ermöglicht eine genauere Bestimmung der Respirationsparameter. Da sowohl die Atemmechanik als auch Atemparameter wichtige Indikatoren bei respiratorischen Erkrankungen sind, nehmen die Methoden zur Maskenerzeugung bei der Klassifikation dieser Erkrankungen eine zentrale Rolle ein.

Um ein systematisches Vorgehen sicherzustellen und die Ursachen der Klassifikation zu untersuchen, werden in diesem Kapitel insgesamt fünf verschiedene Methoden miteinander verglichen:

- das **Referenzmodell** (im Folgenden kurz M1) erstellt mit den gemessenen Größen des Bodyplethysmographen eine Klassifikation und stellt den Vergleichswert für die anderen Modelle dar,
- für die **merkmalsbasierte Klassifikation** (M2) werden über die generalisierte Maske aus Kapitel 5 Signale extrahiert, aus denen die Merkmale aus Kapitel 4 extrahiert werden,
- für die **signalbasierte Klassifikation** (M3) werden ausschließlich die Signale, welche anhand der generalisierten Maske erhoben werden, klassifiziert,
- die **Klassifikation anhand der physiologischen Maske** (M4) erfolgt, indem die individualisierten Patientenmasken klassifiziert werden,
- ein **End-to-End Maschine Learning Modell** (M5) wird angewendet, um über ein Faltungsnetzwerk automatisch Merkmale aus den Tiefeneingangsbildern zu extrahieren und dann mit einem anschließenden Neuronalen Netz zu klassifizieren.

Eine Übersicht dieser Methoden und des Vorgehens in diesem Kapitel ist in Abbildung 6.1 dargestellt und in den folgenden Abschnitten im Detail beschrieben.

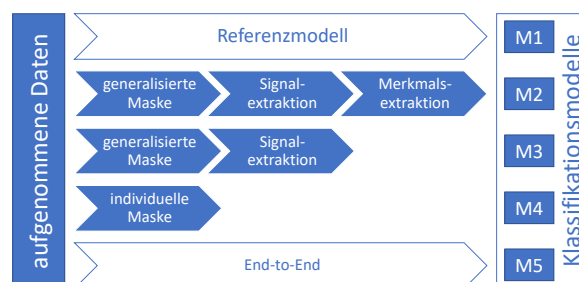


Abbildung 6.1: Vorgehensweise zur Klassifikation von respiratorischen Erkrankungen. Als Daten stehen RGB- und Tiefenbilder, sowie die Messparameter der Bodyplethysmographie zur Verfügung. Blaue Pfeile stellen eine Signalverarbeitung dar, während das Referenzmodell und der End-to-End Ansatz direkt mit den bereitgestellten Daten arbeiten.

Die Leistungsfähigkeit der verschiedenen Methoden wird durch die Evaluierung mittels Konfusionsmatrizen, Accuracy und Cohens Kappa κ untersucht. Die Berechnungsvorschrift für κ ist in Gleichung (6.1) gegeben. Dabei ist P_o die Übereinstimmung zwischen Klassifikator und Grundwahrheit und P_e die zufällig erwartete Übereinstimmung. Somit kann κ einen Wertebereich von $[-1, 1]$ annehmen, wobei ein Wert von 1 eine perfekte Übereinstimmung mit der Grundwahrheit darstellt, ein Wert 0 anzeigt, dass der Klassifikator nicht besser ist, als zufälliges Raten und bei einem Wert < 0 der Klassifikator schlechtere Ergebnisse aufweist, als zufälliges Raten. Ab einem Wert von $\kappa > 0,2$ gilt die Übereinstimmung als ausreichend [113].

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad (6.1)$$

6.1.1 Datensatzgenerierung

Zur Evaluierung der folgenden Methoden wird ein Datensatz benötigt, der über 3D-Tiefenaufnahmen verfügt und sowohl gesunde als auch kranke Testpersonen beinhaltet. Diese Studie wurde bei Dr. med Sohrab in der Neudorfer Lungenpraxis (Gustav-Adolf-Straße 5, 47057 Duisburg) durchgeführt. Ein positives Ethikvotum der Ärztekammer Nordrhein liegt unter der Nummer 2022142 vor.

Datenaufnahme

Die kontaktlose Messung der Respirationsparameter verläuft parallel zu einer herkömmlichen Bodyplethysmographie, wie in Abbildung 6.2 dargestellt. Die Messung mit Hilfe der Orbbec Astra Pro Kamera wird parallel zur Untersuchung im Bodplethysmographen (PowerCube Body+, GANSHORN Medizin Electronic, Niederlauer) gestartet und nach beendeter Bodyplethysmographie gestoppt. Die Kamera befindet sich dabei in einer Höhe von circa 75 cm und ist mit einem Abstand von circa 160 cm zentral auf den Brustkorb der Testpersonen ausgerichtet. Die Verwendung der Orbbec Astra Pro Kamera wurde aufgrund ihrer Untersuchung verschiedener Einflussgrößen, wie Störlicht, Winkel und Reflexionen, zur Detektion von Atembewegungen bei Säuglingen empfohlen [114]. Weitere Informationen der Messungen stellen Alter, Größe, Geschlecht, Gewicht und Diagnose dar, welche pseudonymisiert zur Verfügung gestellt werden. Als Grundwahrheit

stehen die folgenden Parameter der Bodyplethysmographie zur Verfügung:

- inspiratorische Vitalkapazität ($[VC_{in}] = l$),
- forcierte Vitalkapazität ($[FVC] = l$),
- Einsekundenkapazität ($[FEV_1] = l$),
- maximale Atemstromstärke bei forcierter Expiration (engl. peak expiratory flow, $[PEF] = l s^{-1}$),
- mittlere Atemstromstärke bei noch 25 % VC in der Lunge (engl. mean expiratory flow, $[MEF_{25}] = l s^{-1}$),
- Totalkapazität ($[TC] = l$),
- expiratorisches Reservevolumen ($[ERV] = l$),
- Residualvolumen ($[RV] = l$),
- funktionelle Residualkapazität ($[FRC] = l$),
- Thorakales Gasvolumen ($[TGV] = l$),
- Atemwegswiderstand bei normaler Atmung (engl. airway resistance, $[R_{Aw,eff}] = kPa s l^{-1}$),
- Atemwegswiderstand bei normaler Atmung pro Liter Lungenvolumen (engl. airway resistance, $[R_{Aw,eff}] = kPa s l^{-1}$),
- Atemwegswiderstand bei maximaler Aus- oder Einatmung pro Liter Lungenvolumen (engl. airway resistance, $[R_{Aw,tot}] = kPa s l^{-1}$) und
- Atemwegsleitfähigkeit bei normaler Atmung pro Liter Lungenvolumen (engl. airway conductance, $[sG_{Aw,eff}] = kPa^{-1} s$).

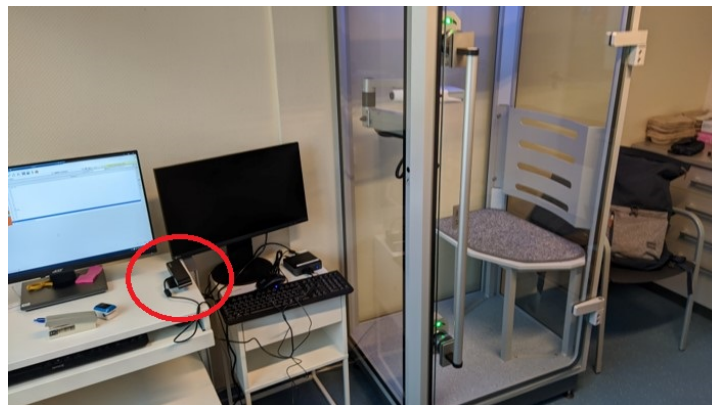


Abbildung 6.2: Messaufbau in der Lungenpraxis. Die Testperson sitzt im Bodyplethysmographen (rechts) und wird von der Orbbec Astra Pro Kamera (roter Kreis) gefilmt.

Ein- und Ausschlusskriterien

Die Einschlusskriterien für diese Studie sind wie folgt definiert:

- vorliegendes schriftliches Einverständnis,
- fließende Deutschkenntnisse und
- Zugehörigkeit zu einer der Gruppen: Gesund, COPD, Asthma, restriktive Erkrankung.

Die Ausschlusskriterien nach Jörres [115] für diese Studie enthalten im Wesentlichen Kontraindikationen für Bodyplethysmographie:

- Durchführung Bodyplethysmographie aufgrund fehlender Mitarbeit oder Klaustrophobie nicht möglich,
- vorliegende absolute Kontraindikationen für forcierte Atemmanöver: akute lebensbedrohliche Krankheitsbilder,
- vorliegende relative Kontraindikation für forcierte Atemmanöver: ausgedehnter Pneumothorax (innerhalb der ersten Wochen), Abdomen-/Thoraxoperation (je nach Befund 1-4 Wochen postoperativ), Augen-/Hirn-/Ohrenoperation (variabel, Rücksprache Operateur), Hämoptysen unklarer Genese,
- Tragen stark faltenschlagender Kleidung (oversized), mehrerer Kleidungsschichten (Unterhemden, Jacken, Schals) oder atmungseinschränkender Kleidung (Korsetts),
- bestehende Schwangerschaft.

6.1.2 Referenzmodell

Um eine Vergleichsbasis für die nachfolgenden Modelle zu schaffen, wird ein Referenzmodell (M1) verwendet. Das Referenzmodell dient als Grundlage zur Bewertung der Leistungsfähigkeit und Genauigkeit anderer Modelle im Vergleich zu einer automatisierten Methode des maschinellen Lernens. Das Referenzmodell verwendet die gemessenen Daten des Bodyplethysmographen als Eingabe und erzeugt eine Klassifikation der Krankheit als Ausgang.

Für die Implementierung des Referenzmodells wird die Auto-Maschine-Learning-Library `auto-sklearn` [116, 117] verwendet. Diese Methode nutzt Algorithmen des maschinellen Lernens, um automatisch Muster und Zusammenhänge in den Daten zu erkennen und eine Klassifikation der Krankheit vorzunehmen. Insgesamt werden 15 Klassifikatoren, 14 Methoden zur Merkmalsvorverarbeitung und

4 Methoden zur Datenvorverarbeitung eingesetzt. Auto-sklearn berücksichtigt Leistungen auf ähnlichen Datensätzen und erzeugt Ensembles aus den einzelnen optimierten Modellen. Das automatische ML-System ist dabei auf mehr als 30 Datensätzen validiert. Die Parameter für auto-sklearn sind in Tabelle 6.1 dargestellt. Auto-sklearn bietet den Vorteil, dass sie den Prozess der Modellentwicklung automatisiert und die Notwendigkeit manueller Eingriffe minimiert. Dadurch wird eine effiziente und objektive Analyse der Daten ermöglicht.

Tabelle 6.1: Hyperparameter und technische Komponenten der Klassifikation für das Referenzmodell durch automatisches Maschinelles Lernen.

Parameter	Größe
Memorylimit	10000 MB
Zeitlimit	15 min
CPU	Intel(R) Xeon(R) Silver 4214 at 2.20 GHz
Memory	96 GiB

Die Wahl dieses Referenzmodells basiert auf mehreren Faktoren. Erstens ist die Auto-Maschine-Learning-Methode weit verbreitet und hat sich in verschiedenen Anwendungsgebieten [118] als erfolgreich erwiesen. Zweitens ermöglicht die Verwendung eines Referenzmodells eine standardisierte Vergleichsbasis für die anderen Modelle, da es als Maßstab für die Leistungsfähigkeit dient. Dadurch können die Stärken und Schwächen der anderen Modelle besser identifiziert und bewertet werden.

6.1.3 Merkmalsbasierte Klassifikation

Die zu untersuchende Fragestellung mit der merkmalsbasierten Klassifikation (M2) lautet: Sind die Merkmale, welche bereits für eine Bestimmung der Respirationsvolumina geeignet sind, auch für die Klassifikation von respiratorischen Erkrankungen geeignet? Dafür werden die in Kapitel 4 definierten Merkmale verwendet.

Signalverarbeitung

Zur Extraktion der Signale aus den Bildern wird die in Kapitel 5 entwickelte generalisierte Maske verwendet und an die detektierten Körpergelenke angepasst. Die Detektion der Körpergelenke erfolgt mithilfe von MediaPipe [119] auf den RGB-Bildern. Die Entscheidung für MediaPipe basiert auf der Echtzeitleistung, der Hardwareunterstützung und der Portabilität, was eine einfache Integration in verschiedene Plattformen und Anwendungen in einem späteren Anwendungsfall ermöglicht, wie vergleichsweise in einem Hand-Tracking-Szenario auf einem modernen Smartphone [120]. Es sei darauf hingewiesen, dass die detektierten Körpergelenke von MediaPipe in diesem Datensatz und der Kinect-Kamera, die zur Entwicklung der generalisierten Maske verwendet wurde, nicht exakt übereinstimmen. Daher werden Korrekturfaktoren k eingeführt, siehe Gleichungen (6.2a) und (6.2b). Diese Korrekturfaktoren beziehen sich auf die Breite b und Höhe h , indem sie auf die x- und y-Koordinaten der Körpergelenke angewendet werden, vergleiche Gleichungen (6.2c) bis (6.2f). Die Schultern S und Hüften H werden sowohl links l als auch rechts r detektiert und entsprechend transformiert.

$$k_b = \frac{x_{lS} - x_{rS}}{12} \quad (6.2a)$$

$$k_h = \frac{\max(x_{lH}, x_{rH}) - \min(x_{lS}, x_{rS})}{7} \quad (6.2b)$$

$$x_{rS} = x_{rS, \text{MediaPipe}} + k_b \quad (6.2c)$$

$$x_{lS} = x_{lS, \text{MediaPipe}} - k_b \quad (6.2d)$$

$$y_{pos} = y_{pos, \text{MediaPipe}} + k_h \text{ für } pos \in [rS, rH] \quad (6.2e)$$

$$y_{pos} = y_{pos, \text{MediaPipe}} - k_h \text{ für } pos \in [lS, lH] \quad (6.2f)$$

Anhand dieser Korrekturen kann über eine affine Transformationsmatrix die generalisierte Maske an die Probanden angepasst werden, siehe bereits die ausführliche Beschreibungen in Abschnitt 5.1.1.

In Übereinstimmung mit Kapitel 4 erfolgt die Signalvorverarbeitung. Das generierte Signal wird mit einem Butterworth-Tiefpass-Filter dritter Ordnung mit einer Grenzfrequenz bei 1.5 Hz geglättet und anschließend um den linearen Trend des Signals bereinigt, in Anlehnung an bisherige Arbeiten [75]. Die Extraktion der 374 Merkmale erfolgt identisch zu den Ausführungen in Kapitel 4. Persönliche

Merkmale wie Größe, Alter und Geschlecht werden nicht einbezogen, um sicherzustellen, dass das Klassifikationsergebnis nicht von der Multikollinearität dieser Merkmale abhängt, sondern ausschließlich von den respiratorischen Merkmalen.

Modellerstellung

Ziel der Untersuchung, ist die Feststellung, ob die extrahierten Merkmale zur Klassifikation geeignet sind. Aus diesem Grund werden Modelltypen eingesetzt, die für eine Klassifikation geeignet sind. Zum einen werden SVMs eingesetzt und zum anderen ein Auto-ML Ansatz.

In vorherigen Abschnitten (Kapitel 4) zeigte sich, dass eine Regression über Support Vektoren (SVR) die besten Ergebnisse mit den zuvor extrahierten Merkmalen erzielte. Daher werden in diesem Abschnitt zur Klassifikation ebenfalls Support Vektoren (SVM) eingesetzt. Es werden die gleichen Hyperparameter verwendet, wie bereits zuvor beschrieben, siehe Abschnitt 4.1.4. Dies geschieht ebenso in Kombination mit Merkmalsselektionsverfahren, siehe Abschnitt 4.1.3.

Da das zuvor entwickelte SVR-Modell zwar beste Ergebnisse für die Regression erzielte, jedoch nicht zwangsläufig die besten Ergebnisse für die Klassifikation anhand der extrahierten Merkmale liefert, wird ein Vergleichswert durch einen Einsatz des Auto-ML-Ansatzes über `auto-sklearn` [116, 117] erstellt. Die Parameter bleiben dabei im Vergleich zum Referenzmodell (M1) in Tabelle 6.1 unverändert, um die Vergleichbarkeit zwischen M1 und M2 zu gewährleisten. Die Klassifikation anhand der Merkmale in M2 erfolgt somit über SVM und einen Auto-ML-Ansatz.

6.1.4 Signalbasierte Klassifikation

In der merkmalsbasierten Klassifikation (M2) wird die Wirksamkeit der zuvor extrahierten Merkmale auf ihre Eignung zur Klassifikation getestet. Trotz der hohen Anzahl von 374 Merkmalen besteht die Möglichkeit, dass relevante Aspekte nicht berücksichtigt werden. Um sicherzustellen, dass potenziell relevante Merkmale nicht unbeachtet bleiben, wird in diesem Abschnitt ein Modell (M3) verwendet, welches eigenständig relevante Merkmale extrahiert.

CNNs sind in der Lage, eigenständig Merkmale durch Faltungsoperationen zu extrahieren, siehe Abschnitt 2.4. Sie haben sich in der Biosignalanalyse als äußerst

leistungsfähig erwiesen [111, 121]. Aus diesem Grund wird ein CNN ausgewählt, um die Signale zu klassifizieren. Die Extraktion der Signale erfolgt dabei, wie zuvor in Abschnitt 6.1.3 beschrieben, anhand der generalisierten Maske.

Signalvorverarbeitung

Die so extrahierten Signale werden über ein Moving-Average Filter mit der Größenordnung 15 geglättet und anschließend um den linearen Trend des Signals bereinigt. Die Schätzung des linearen Trends basiert auf der linearen Regression, bei der die Trendgerade durch die Parameter a (Steigung) und b (Achsenabschnitt) dargestellt wird, siehe Gleichung (6.3a). Die Parameter a und b werden so gewählt, dass der quadratische Abstand zwischen den geschätzten Trendwerten \hat{s}_i und den ursprünglichen Datenpunkten s_i minimiert wird (Least-Mean-Squares). Das bereinigte Signal \tilde{s} erhält man durch Subtraktion des geschätzten Trends von den ursprünglichen Datenpunkten, siehe Gleichung (6.3b).

$$\hat{s}_i = a \cdot i + b \quad (6.3a)$$

$$\tilde{s}_i = x_i - \hat{s}_i \quad (6.3b)$$

Eine Signalskalierung zwischen 0 und 1 erfolgt anhand der globalen Extrema im entsprechenden Probandensignal. Da diese Extrema mit der maximalen forcierten Ein- und Ausatmung übereinstimmen sollten, sind alle weiteren Atemvolumina direkt als relationale Größen anzusehen, wenn sie von dem CNN extrahiert werden.

Als Signalabschnitt werden insgesamt 300 Samples ausgewählt. Dies entspricht bei einer Framerate von 30 s^{-1} 10 s des Atemsignals. Als Signal wird dabei der Einstieg in die forcierte Atmung gewählt. So werden 5 s Tidalatmung und anschließend 5 s forcierte Atmung ausgewählt, siehe Abbildung 6.3.

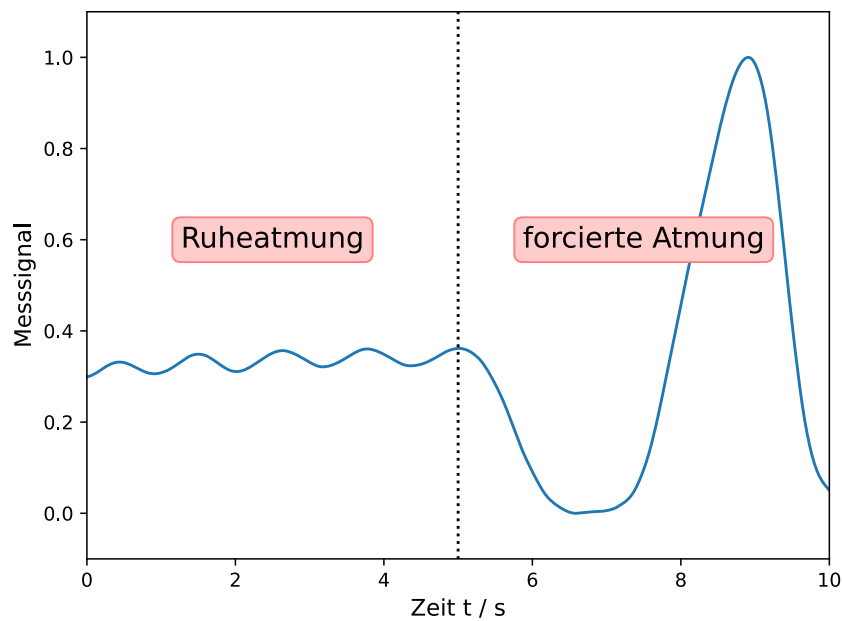


Abbildung 6.3: Ausgewählter Messabschnitt des skalierten Messsignals mit den Bereichen der Ruheatmung und forcierter Atmung.

Modellerstellung

Die Klassifikation erfolgt mithilfe eines CNN. Als relevante Hyperparameter eines 1D-CNN [122] werden die folgenden Hyperparameter optimiert:

- Anzahl der versteckten Faltungsschichten,
- Größe des Faltungskernels,
- Subsampling-Faktor in den Faltungsschichten,
- die Wahl der Pooling und Aktivierungsoperatoren.

Der Bereich der Parameter ist in Tabelle 6.2 dargestellt.

Tabelle 6.2: Hyperparameter für die Signal-Klassifikation durch ein CNN.

Parameter	Größe
Faltungsschichten	3
Filter	[3, 5, 12, 24, 36]
Kernel	[3, 5, 7, 11]
MaxPooling	[3, 5, 7, 11]
Neuronen in Dense-Schicht	100
Hidden-Aktivierungsfunktion	ReLU
Output-Aktivierungsfunktion	Softmax
Loss	Categorical-Crossentropy
Lernrate	[0.1, 0.01, 0.001]

6.1.5 Klassifikation anhand der physiologischen Maske

In Kapitel 5 werden die individualisierten Patientenmasken als Repräsentation der Atemmechanik verwendet. Als Heatmap dargestellt sind die Bereiche hervorgehoben, die eine hohe Mutual Information mit dem Spirometersignal aufweisen. Diese Darstellung ermöglicht es, diejenigen Bereiche auszuwählen, die besonders zum Messsignal beitragen. Da sich COPD durch Asynchronität in der Atemmechanik bemerkbar machen kann [92], besteht der Ansatz der in diesem Abschnitt beschriebenen Methode darin, anhand der individuellen Patientenmaske respiratorische Erkrankungen zu klassifizieren (M4). Es wird die Hypothese aufgestellt, dass die Asynchronität sich in der individuellen Patientenmaske zeigt und somit ein wichtiger Faktor für die Klassifikation ist. Beispielsweise könnten so thorakale und abdominale Areale in der Maske durch die Asynchronität der Respirationsmechanik verschieden gewichtet werden. Die Ergebnisse von Motamedi-Fakhr, Wilson und Iles [92] lassen eine solche Hypothese zu.

Untersuchung individueller, korrelationsbasierter Masken

Anders als in Kapitel 5 stehen für die Messungen in diesem Kapitel keine Spirometersignale zur Verfügung, siehe Abschnitt 6.1.1. Es ist somit nicht möglich, die MI für die Erzeugung der Maske zu verwenden. Aus diesem Grund wird im Folgenden die Erzeugung von individuellen Masken über den Pearson-Korrelationskoeffizienten untersucht.

Zunächst wird, wie bereits in Abschnitt 5.1.1 beschrieben, eine Hintergrundelimination durchgeführt und die Signale $s_{x,y}$ für jedes Pixel (x, y) extrahiert. Aus den verbleibenden Signalen wird nun jeweils der Pearson Produkt-Moment-Korrelationskoeffizient $r \in [-1, 1]$ berechnet und eine Korrelationsmatrix aufgestellt. Gemessen wird somit der lineare Zusammenhang der Signale. Da bereits der Hintergrund entfernt ist, bleiben ausschließlich Signale übrig, die potentiell zum Atemsignal beitragen. Es ist somit anzunehmen, dass der Anteil an nutzbringenden Signalen den Rauschanteil übersteigt. Zur Gewichtung und Identifikation von ähnlichen Bereichen, werden anschließend die Korrelationskoeffizienten r für jedes Signal summiert (Gleichung (6.4a)). Mit einem hohen Wert dieser Summe $m_{x,y}$ zeigt sich eine hohe lineare Übereinstimmung mit vielen anderen Signalen der segmentierten Probanden. Eine hohe negative Gewichtung weist auf eine negative Korrelation hin und ein Wert um die 0 auf wenig Korrelation zu den anderen Signalen, potentiell verursacht durch Rauschanteile. Nach Gleichung (6.4b) entsteht die finale Korrelationsmaske M , mit den oben genannten Bedingungen für die Hintergrundelimination.

$$m_{x,y} = \sum_i \sum_j r_{s_{x,y}, s_{i,j}} \quad (6.4a)$$

$$M[x, y] = \begin{cases} 0, & \text{für } s_{x,y}[t] = 0, \text{ für } t \in T \\ m_{x,y} & \text{sonst} \end{cases} \quad (6.4b)$$

In dieser Untersuchung wird ausschließlich die Tidalatmung als Signal verwendet. In früheren Studien wurden signifikante Unterschiede in der Atemmechanik bei der Tidalatmung festgestellt [91, 92]. Zudem können die nicht-linearen Effekte, die bei der Bewegung in der forcierten Atmung beobachtet wurden, wie bereits in Kapitel 4 und Kapitel 5 beschrieben, zu Komplikationen bei der Interpretation der linearen Pearson-Korrelationskoeffizienten führen. Daher wird

die Tidalatmung als geeignetes Signal für die Klassifikation von respiratorischen Erkrankungen betrachtet.

Es ist beabsichtigt, dass die Patientenmaske Rückschlüsse auf die Synchronität der Atemmechanik zulässt. Jedoch sind durch die Kontur und Größe der Maske eventuell weitere Rückschlüsse auf die einzelnen Patienten möglich. So wäre vorstellbar, dass die Größe und Gewicht durch die Maskenform abgeschätzt werden kann. Da eine Klassifikation ausschließlich anhand der Atemdynamik am Oberkörper stattfinden soll, wird im Folgenden eine ROI ausgewählt auf dem Oberkörper ausgewählt und entsprechend auf 100 x 60 Pixel skaliert, vergleiche Abbildung 6.4. Anhand dieses Bereichs bleiben die ursprünglichen Größenverhältnisse weitestgehend erhalten. Die Auswahl der ROI erfolgt anhand der ermittelten Körpergelenke durch MediaPipe [119], dargestellt durch Gleichung (6.5). Es sind b die Breite der ROI und h die Höhe sowie x und y die entsprechenden Koordinaten. Die Körpergelenke sind entsprechend gekennzeichnet: l - links, r - rechts, S - Schulter, H - Hüfte. Die Unterschiede zu Gleichung (4.1) entstehen durch eine andere Abschätzung der Körperpositionen.

$$b = x_{lS} - x_{rS} \quad (6.5a)$$

$$h = \max(x_{lH}, x_{rH}) - \min(x_{lS}, x_{rS}) \quad (6.5b)$$

$$x_{ROI} = x_{rS} \quad (6.5c)$$

$$y_{ROI} = \max(x_{lS}, x_{rS}) + 0,1 \cdot h \quad (6.5d)$$

Zusammenfassend bietet diese Methode eine neue Möglichkeit zur Generierung einer individuellen Patientenmaske. Ohne das Vorhandensein von Spirometersignalen wird eine gewichtete Maske anhand von Korrelationskoeffizienten gebildet. Somit werden Regionen mit ähnlichen Verhalten in der Atemmechanik betont und Rauschen reduziert. Über den Pearson-Korrelationskoeffizienten werden die statistischen Zusammenhänge bis zur zweiten Ordnung berücksichtigt. Zusammenhänge höherer Ordnung können nicht berücksichtigt werden, dies war insbesondere bei der Merkmalsauswahl anhand von MI jedoch von entscheidender Bedeutung.

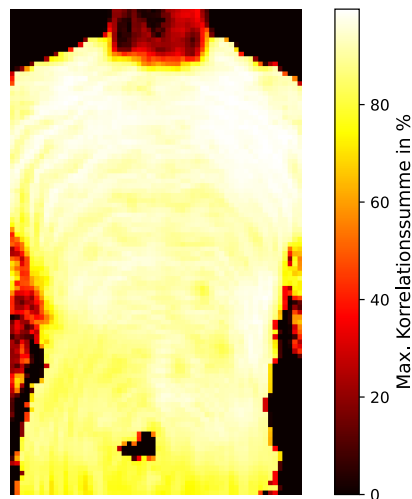


Abbildung 6.4: Ausgewählte Region of Interest der individuellen Maske.

Modellerstellung

Ziel ist es, die individuelle Patientenmaske, also ein 2-dimensionales Bild, zu klassifizieren. Für den Bereich der Bildklassifikation wird ein CNN verwendet. Die zu optimierenden Parameter befinden sich in Tabelle 6.3.

Tabelle 6.3: Hyperparameter für die Masken-Klassifikation durch ein CNN.

Parameter	Größe
Faltungsschichten	3
Filter	[8, 16, 32]
Kernel	[3, 5, 7, 11]
MaxPooling	[2, 3, 5]
Neuronen in Dense-Schicht	64
Hidden-Aktivierungsfunktion	ReLU
Output-Aktivierungsfunktion	Softmax
Loss	Categorical-Crossentropy
Lernrate	[0.1, 0.01, 0.001]

6.1.6 End-to-End Maschine Learning Modell Ansatz

Im Gegensatz zu den vorherigen Abschnitten, in denen eine Signalverarbeitung und Datenreduktion stattfindet, bevor ein Klassifikationsmodell eingesetzt wird, soll in diesem Abschnitt ein anderes Vorgehen verwendet werden. Ein End-to-End Ansatz verarbeitet die Eingangsdaten direkt und ist in der Lage wichtige Informationen zu extrahieren, die anschließend für eine automatische Klassifikation genutzt werden (M5). Zu diesem Zweck wird im Folgenden der Aufbau einer hybriden CNN-LSTM-Architektur, respektive CNN-CNN, dargestellt. Diese Architekturen ermöglichen die effektive Verarbeitung von Daten mit räumlichen und zeitlichen Abhängigkeiten.

Das 2D-CNN eignet sich zur Extraktion von räumlichen Merkmalen aus Tiefendaten. Durch die Verwendung von Faltungsoperationen lernt das CNN automatisch relevante Merkmale auf verschiedenen Skalenebenen. Diese Merkmale werden anschließend an das LSTM weitergeleitet, das die zeitliche Abhängigkeit zwischen den extrahierten Merkmalen modelliert. Dadurch wird eine ganzheitlichere Erfassung der Datenstruktur ermöglicht, was zu einer verbesserten Leistung bei der Klassifikation führen kann. Das LSTM ist für die Modellierung von zeitlichen Abhängigkeiten besonders gut geeignet. Im Gegensatz zu herkömmlichen rekurrenten Neuronalen Netzen kann das LSTM lange Abhängigkeiten über einen längeren Zeitraum hinweg erfassen, ohne dass das Problem des Verschwindens oder Explodierens des Gradientens auftritt. Diese Fähigkeit ist entscheidend, um Zeitreihendaten angemessen zu verarbeiten. Die Integration eines 2D-CNN mit nachgeschaltetem LSTM kann auch dazu beitragen, die Robustheit gegenüber Variabilität und Rauschen in den Daten zu verbessern. Das CNN kann invariant gegenüber bestimmten Transformationen sein und dadurch unempfindlich gegenüber geringfügigen Variationen in den Daten werden. Das LSTM-Modell kann wiederum die temporale Konsistenz in den Daten nutzen, um Rauschen oder unerwünschte Variationen zu glätten und die Klassifikationsleistung zu verbessern.

Datenvorverarbeitung und Modellerstellung

Die Datenvorverarbeitung erfolgt identisch zu Abschnitt 6.1.4 und Abschnitt 6.1.5. Die Bilder werden skaliert $[0, 1]$ und ausschließlich die entsprechende ROI ausgewählt. Als Signalabschnitt werden insgesamt 300 Samples selektiert. So werden

5 s Tidalatmung und anschließend 5 s forcierte Atmung verwendet, siehe bereits Abbildung 6.3.

Die grundlegende Struktur des zweistufigen Modells ist empirisch bestimmt. Die zu optimierenden Parameter des CNN der ersten Stufe finden sich in Tabelle 6.4, genauso wie das LSTM, respektive CNN, der zweiten Stufe.

Tabelle 6.4: Hyperparameter für die Klassifikation durch End-to-End Maschinelles Lernen. Es wird eine zweistufige Architektur eingesetzt mit einem 2D-CNN Modell (erster Abschnitt) und anschließend einer Klassifikation durch ein LSTM (zweiter Abschnitt) oder ein 1D-CNN (dritter Abschnitt).

Parameter	Größe
Faltungsschichten	[1, 2]
Filter	[8, 16]
Kernel	[3, 7]
MaxPooling	[2, 5]
LSTM-Neuronen	128
LSTM-Aktivierung	TanH
LSTM-Aktivierung rekurrent	Sigmoid
LSTM-Dropout	0
CNN-Filter-Schichten	[3, 7, 9]
CNN-Kernel-Schichten	[11, 5, 3]
CNN-MaxPooling	3
Neuronen in Dense-Schicht	64
Hidden-Aktivierungsfunktion	ReLU
Output-Aktivierungsfunktion	Softmax
Loss	Categorical-Crossentropy
Lernrate	0.01

6.2 Ergebnisse und Diskussion

6.2.1 Aufbau des Datensatzes

Die vorliegende Vorabuntersuchung der geplanten Studie für 120 Testpersonen umfasst 92 Aufnahmen. Aus diesen Aufnahmen werden 53 für die Weiterverarbeitung verwendet. Ausschlusskriterien waren ein Widerruf der Einwilligungserklärung ($N = 2$), technische Schwierigkeiten¹ ($N = 17$), eine fehlerhafte Kamerapositionierung ($N = 5$) sowie das Tragen von Jacken oder Schals ($N = 15$). Restriktive Patienten befanden sich noch nicht unter den gemessenen Patienten.

Die Mittelwerte und Standardabweichungen der beschreibenden Datensatzparameter sind in Tabelle 6.5 dargestellt. Der Datensatz weist eine Ungleichverteilung auf, wobei die Asthma-Patienten die Majoritätsklasse darstellen. Am seltensten vertreten sind gesunde Probanden, deren Messungen am häufigsten aufgrund von technischen Problemen ausgeschlossen wurden. Die verschiedenen Klassen unterscheiden sich hinsichtlich des mittleren Alters, der Größe und des Geschlechts der Testpersonen, wobei die COPD-Patienten im Durchschnitt die älteste, kleinste und leichteste Klasse bilden. Ebenfalls weist diese Klasse die geringste durchschnittliche FVC auf.

Tabelle 6.5: Deskriptive Beschreibung des aufgenommenen Datensatzes und der inkludierten Messungen für gesunde Probanden, COPD-Patienten und Asthmatiker. Dargestellt sind die Mittelwerte \pm Standardabweichung.

Parameter	Gesamt	Gesund	COPD	Asthma
Personen (davon ♀)	53 (26)	9 (2)	18 (10)	26 (14)
Alter in Jahren	55.2 \pm 17.9	40.3 \pm 14.4	69.5 \pm 10.7	50.5 \pm 16.4
Größe in cm	171 \pm 9	179 \pm 7	165 \pm 8	173 \pm 8
Gewicht in kg	78.0 \pm 14.9	82.0 \pm 15.5	72.1 \pm 15.4	80.6 \pm 13.2
FVC in l	3.54 \pm 1.46	4.87 \pm 1.04	2.19 \pm 0.62	3.99 \pm 1.29

¹In diesen Fällen wurden aus ungeklärten Gründen in der parallelen Aufnahme entweder nur RGB, oder nur Tiefenbilder aufgezeichnet. Das Vorhandensein beider Modalitäten ist jedoch für die Auswertung von Bedeutung.

6.2.2 Klassifikationsgenauigkeiten

Die Klassifikationsergebnisse sind in Tabelle 6.6 zusammenfassend dargestellt. Dabei ist nach Landis und Koch [113] die Stärke der Übereinstimmung mit Cohens Kappa κ wie folgt zu interpretieren:

- $\kappa < 0$ schlecht (engl. poor),
- $0 \leq \kappa \leq 0,20$ geringfügig (engl. slight),
- $0,20 < \kappa \leq 0,40$ ausreichend (engl. fair),
- $0,40 < \kappa \leq 0,60$ mäßig (engl. moderate),
- $0,60 < \kappa \leq 0,80$ erheblich (engl. substantial),
- $0,80 < \kappa \leq 1,00$ fast perfekt (engl. almost perfect).

Es ist erkenntlich, dass das Referenzmodell (M1) die höchste Genauigkeit und das höchste κ aufweist. Die Modellübereinstimmung ist dabei als angemessen zu interpretieren. Eine größere Abweichung von $\kappa = 1$ ist durch das Entfernen von persönlichen Patientendaten, wie Alter, Größe, Geschlecht und Gewicht zu begründen. In der klinischen Praxis werden anhand dieser Daten die Referenzgrößen gebildet und Abweichungen für eine Diagnostik herangezogen. Des Weiteren ist die Bodyplethysmographie kein eigenständiges Diagnosekriterium in der klinischen Praxis. Dem Referenzmodell entgegen stehen die Modelle basierend auf den Merkmalen (M2) mit einer Genauigkeit von 49,1%. Ohne eine Verarbeitung der Merkmale, mit einer Klassifikation anhand des Signals durch ein 1D-CNN (M3) ist eine bessere Klassifikation möglich (Genauigkeit 55,8% und κ 0,24). Das 2D-CNN zur Klassifikation der individuellen Patientenmasken (M4) ist hingegen besser zur Klassifikation geeignet (57,7%, κ 0,28). Die End-to-End Modelle (M5) zeigen eine geringfügige Übereinstimmung für CNN (52,0%, κ 0,015) und LSTM respektive (52,0%, κ 0,015). Es ist somit festzustellen, dass die entwickelten Modelle keine bessere Klassifikationsgenauigkeit erzielen können, als die Daten des Bodyplethysmographen (Referenzmodell, M1). Die Interpretation von Cohens Kappa lässt jedoch die Schlussfolgerung zu, dass Informationen aus der DPG-Messung extrahiert werden können, die eine Klassifikation grundsätzlich (ausreichend, $\kappa > 0,20$) ermöglichen.

Im Vergleich zur Literatur werden geringere Klassifikationswerte erreicht. Dies ist jedoch entsprechend einzuordnen. Feitosa u. a. [56] erreichen eine Sensitivität von 93,75% und eine Spezifität von 83,33% bei der Erkennung von Belastungsasthma bei Kindern. Der Einsatz von markerbasierten OEP und 8 Kameras ist

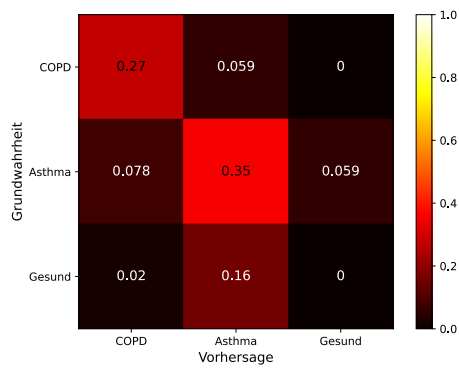
Tabelle 6.6: Übersicht der Modelle hinsichtlich der Klassifikationsgütern Genauigkeit und Cohens Kappa.

Modell	Genauigkeit in Prozent	Cohens Kappa
M1 Referenz	62,8	0,37
M2 Merkmale SVM	49,1	0,11
M2 Merkmale Auto-ML	49,1	0,13
M3 Signal	55,8	0,24
M4 Maske	57,7	0,28
M5 End-to-End CNN	51,9	0,15
M5 End-to-End LSTM	50,9	0,05

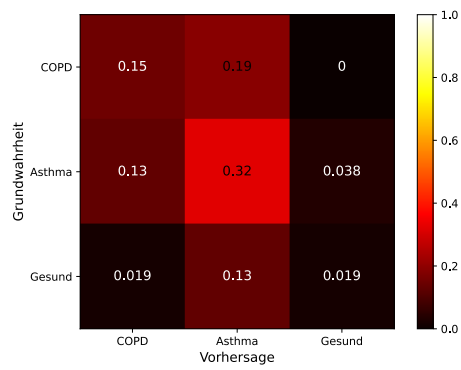
jedoch nicht mit dem Messaufbau in dieser Arbeit vergleichbar. Die Klassifikation von unterschiedlichen Obstruktionsschweren in Ostadabbas u. a. [61] erfolgt anhand einer Kalibration des Systems auf Messungen des selben Probanden. Die Accuracy von 76,2% ist somit nicht mit dem kalibrationsfreien Ansatz in dieser Arbeit vergleichbar. Sharp u. a. [80] kalibrieren ebenfalls ihre Messsysteme auf andere Messungen der selben Probanden. Der berichtete positive Vorhersagewert für COPD Patienten von 83,3% kann daher nicht auf die Ergebnisse dieser Arbeit übertragen werden. Einen vergleichbaren, kalibrationsfreien Ansatz entwickelte Takamoto u. a. [43]. Über multiple-lineare Regression unter Verwendung von LOOCV werden Atemparameter aus dem Messsignal extrahiert. FEV_1/FVC eignet sich anschließend zur Klassifikation von COPD Patienten und Gesunden und erreicht 81% Sensitivität, 90% Spezifität und eine Accuracy von 85%. Diese Modellgüte wird in dieser Studie auch nicht vom Referenzmodell (M1) erreicht, welches zur Klassifikation die Grundwahrheiten des Bodyplethysmographen als Eingang verwendet. Unter diesen Eingangsdaten befinden sich ebenfalls sowohl FEV_1 als auch FVC, sodass das Verhältnis FEV_1/FVC für diese Studie keine ausreichenden Klassifikationsgenauigkeiten erreicht. Wenn dies bereits aus den Grundwahrheiten nicht möglich ist, können die weiteren Modelle (M2-M5) ebenfalls nur eine geringere Genauigkeit aufweisen. Weitere Ursachen für die, in dieser Arbeit, geringe Klassifikationsgenauigkeit werden in Abschnitt 6.2.3 diskutiert.

Die detailliertere Betrachtung der Konfusionsmatrizen in Abbildung 6.5 lässt eine genauere Betrachtung der Modellprädiktionen zu. Die Konfusionsmatrizen geben Aufschluss über die Vorhersageergebnisse eines Klassifikationsmodells, das zwischen den drei Klassen Gesund, COPD und Asthma unterscheidet. Die Matrix zeigt die Anzahl der korrekt und inkorrekt klassifizierten Instanzen in Prozent für jede Klasse.

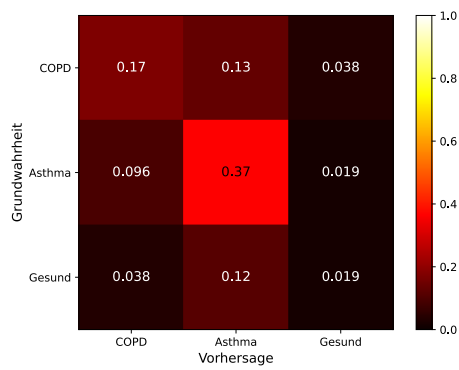
Für das Referenzmodell (M1) ist in Abbildung 6.5a zu erkennen, dass die meisten COPD- (35 %) und Asthma-Patienten (27 %) richtig klassifiziert werden. Gleichzeitig wird kein Gesunder als solcher erkannt und die Mehrheit (56,9 %) der Prädiktionen der Asthma-Gruppe zugeordnet. Dies deutet auf ein Problem des Klassenbalancings im Modell hin. Verstärkt ist dies in Abbildung 6.5b für das SVM-Modell der Merkmalsklassifikation (M2) zu erkennen. Die Mehrheit der Prädiktionen aller Klassen (gesamt 60 %) existiert für die Asthma-Klasse. Die Klassifikation der Messsignale (M3) in Abbildung 6.5c erlaubt es, COPD-Patienten (17 %) und Asthma-Patienten (37 %) mehrheitlich richtig zu klassifizieren. Jedoch werden die wenigsten Gesunden als gesund erkannt (1,9 %). Die Gruppe an COPD-Patienten, welcher fälschlicherweise Asthma vorhergesagt wird (13 %), ist größer als die Gruppe an Asthmatikern, die fälschlicherweise zu COPD zugeordnet werden (9,6 %). Die Konfusionsmatrix des Modells zur Klassifikation der individuellen Masken (M4) ist in Abbildung 6.5d dargestellt. Gesunde Personen werden zu gleichen Teilen den Klassen zugeordnet. Asthmatiker werden überwiegend richtig klassifiziert (37 %). COPD Patienten werden größtenteils ebenfalls den Asthma-Patienten zugeordnet. Jedoch wird kein COPD Patient als gesund klassifiziert. Das Modell ist am Besten in der Lage, Asthma-Patienten richtig zu klassifizieren. Trotz der schlechteren Genauigkeit bei der Klassifikation von COPD-Patienten ist die Modellgenauigkeit insgesamt den anderen Modellen überlegen. Dies wird durch die ungleiche Klassenverteilung verstärkt. Die hybriden End-to-End Modelle (M5) unterscheiden sich in ihrem Klassifikationsverhalten nicht von den anderen Klassifikatoren, dargestellt in Abbildungen 6.5e und 6.5f. Der geringste κ -Koeffizient wird ersichtlich für das CNN-LSTM Modell. Bei diesem werden 94 % aller Eingangsdaten als Asthma-Patienten klassifiziert, der Majoritätsklasse folgend.



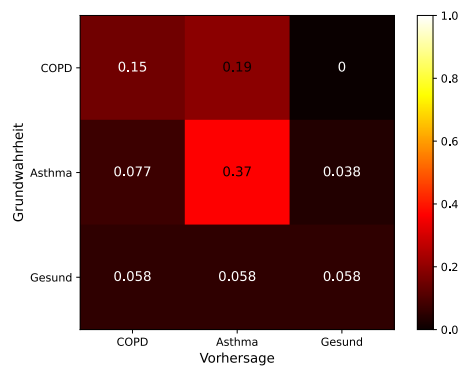
(a) Referenzmodell



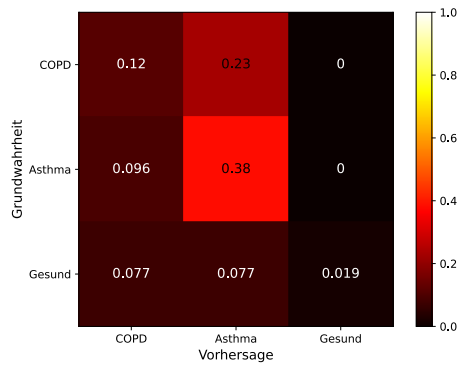
(b) Merkmalsbasiert-SVM



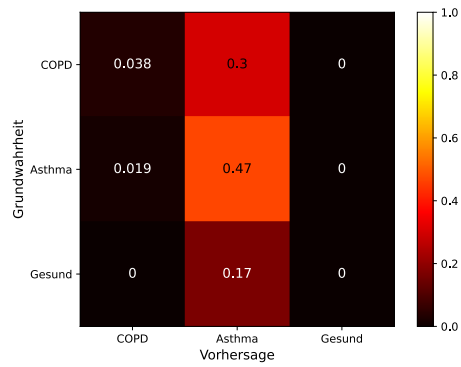
(c) Signalbasiert



(d) Masken



(e) CNN-CNN



(f) CNN-LSTM

Abbildung 6.5: Konfusionsmatrizen der einzelnen Klassifikationsmodelle mit der Prädiktion in Abhängigkeit vom wahren Wert für die Gruppen Asthma, COPD und Gesund. Die Häufigkeit ist relational zur Gesamtprobandenanzahl angegeben.

6.2.3 Diskussion der Klassifikationsgenauigkeiten

Mögliche Ursachen für die niedrige Klassifikationsgenauigkeit lassen sich in den verschiedenen Schritten der DPG-Messung diskutieren: Datenerhebung, Vorverarbeitung und Modellbildung.

Datenerhebung

Die Daten der Studie wurden simultan zur Bodyplethysmographie in der Neudorfer Lungenpraxis erhoben. Dabei befand sich die Kamera außerhalb der Glaskammer im Abstand von 160 cm und einer Höhe von 75 cm. Das projizierte Muster der SL-Kamera wird zum Teil bereits an der Glaswand des Bodyplethysmographen reflektiert und nur ein Teil erreicht den Brustkorb der Patienten. Dies kann unter Umständen dazu führen, dass der Abstand zum Brustkorb nicht korrekt gemessen werden kann. Eine visuelle Inspektion zeigt jedoch keine Störungen im Messsignal. Die Spiegelung des Glases des Bodyplethysmographen kann dennoch zu einem erhöhten Rauschanteil im Messsignal führen [46].

In diesem Kapitel wurde die Signalvorverarbeitung angepasst, um der Verwendung der Orbbec Astra Pro Kamera (640×480 Pixel bei 30 Hz und 45° Field of View [123]) gerecht zu werden. Diese Kamera unterscheidet sich von der zuvor verwendeten Microsoft Kinect Kamera (512×424 Pixel bei 30 Hz und 60° Field of View [124]) in den vorliegenden Datensätzen. Die Anpassung war erforderlich, um die unterschiedliche Signalqualität effektiv zu verarbeiten und zu berücksichtigen. Die Orbbec Kamera ist grundsätzlich für den Einsatz in dieser Studie geeignet und wird auch hinsichtlich einiger Parameter wie Reflexion empfohlen [114].

Ein hoher Anteil an getätigten Aufnahmen ($N = 15$) musste nachträglich verworfen werden, da Patienten Kleidungsstücke trugen, die den Oberkörper zusätzlich verdeckten, wie zum Beispiel Jacken oder Schals. Dies ließ sich aus den aufgenommenen RGB-Bildern rekonstruieren. Es lässt sich jedoch nicht posteriori feststellen, ob die Patienten ein Unterhemd trugen. Dies ist ebenfalls nach Studienprotokoll ausgeschlossen, da es die messbare Oberkörperbewegung an der Oberfläche einschränkt und Faltenbildung anregt [61]. Eine mögliche Anzahl an Probanden mit Unterhemd und potentielle Auswirkungen lassen sich nicht genauer quantifizieren, sind aber eine potentieller Einflussfaktor der Störgrößen.

Die Rekrutierung von erkrankten Patienten erfolgte über die Neudorfer Lun-

genpraxis. Dies führt dazu, dass ein großer Anteil an gemessenen Patienten nach der Erstdiagnose auch bereits in Behandlung ist. Dies kann die Symptome, insbesondere von Asthma-Patienten, signifikant abschwächen [125]. Der für die gesamtheitliche Studie zu untersuchende Teil von restriktiven Patienten wurde nicht für diese hier vorgestellte Teiluntersuchung inkludiert. Insbesondere bei dieser Patientengruppe sind atemmechanische Veränderungen zu erwarten, mit der Hypothese, dies durch DPG messen zu können. So ist durch die zu untersuchenden Gruppen an Patienten noch keine repräsentative Stichprobe für alle Krankheiten ausgewählt und vorhandene Erkrankungen durch eine Behandlung eventuell abgeschwächt.

Die zu untersuchende Hypothese dieses Kapitels basiert unter anderem auf der Annahme, dass Patienten mit respiratorische Erkrankungen eine paradoxen Atemmechanik aufweisen [91]. Dies ist jedoch nicht bei allen Erkrankten der Fall [93]. Aus diesem Grund wurde die zu untersuchende Patientengruppe in der Studie hinreichend groß gewählt (jeweils $N = 30$). Aufgrund der Ausschlüsse kann es nun in der verminderten, inkludierten Patientenzahl dieser Voruntersuchung der Fall sein, dass auch nur in wenigen Fällen die paradoxe Atemmechanik tatsächlich sichtbar ist. Auch die ungleiche Klassenverteilung selbst kann dazu führen, dass eine Klassifikation weniger erfolgreich ist, da ML-Modelle dazu tendieren können, die Majoritätsklasse besser zu erkennen und die selteneren Klassen zu vernachlässigen.

Als weiterer Punkt ist der Einfluss der Bewegung der Patienten bei der DPG Messung zu diskutieren. Insbesondere bei der forcierten Atmung ist Bewegung des Oberkörpers zu beobachten, der die Brustkorbbewegung superpositioniert. Gleiches Verhalten ist bei anderen Studien untersagt [60, 61]. Diese Arbeit untersucht jedoch eine freie und natürliche Atmung, wie bereits in Kapitel 4 und 5 dargestellt, sodass die Bewegungen in der Studie nicht eingeschränkt wurden.

Vorverarbeitung

Neben der Datenaufnahme kann die Datenvorverarbeitung einen erheblichen Einfluss auf das schlussendliche Klassifikationsergebnis aufweisen. Das Vorgehen unterscheidet sich hier zwischen den verschiedenen Modellen. Im merkmalsbasierten Ansatz (M2) ist die Qualität der Merkmale maßgeblich entscheidend für das Klassifikationsergebnis. Die verwendeten Merkmale zur Klassifikation waren zuvor für

die Regression der Atemvolumina erstellt und evaluiert, siehe Kapitel 4. Einige klinische Parameter, wie zum Beispiel FEV_1 , werden nicht als zu extrahierendes Merkmal erfasst. Da FEV_1 jedoch entscheidende Veränderungen bei Asthma und COPD aufweist [28], ist zu erwarten, dass dieses Merkmal einen Beitrag zur Klassifikation leisten könnte. Die Abwesenheit dieser klinischen Parameter in den extrahierten Merkmalen kann somit die Klassifikationsleistung der merkmalsbasierten Modelle verschlechtern.

Das extrahierte Messsignal für M2 und M3 entstammt der gematchten Maske, welche aus dem Datensatz von Soleimani, Mirmehdi und Damen [94] in Kapitel 5 entstanden ist. Diese Maske ist auf den Daten von Gesunden generiert und somit eventuell nicht gänzlich geeignet die beitragenden Regionen bei Erkrankten abzubilden. Dies kann zu einem höheren Rauschanteil und verringerten Signalanteil im Messsignal führen. Des Weiteren komprimiert die Maske das Messsignal vom ganzflächigen Oberkörper. Eine Unterscheidung in einzelne Regionen, wie in der Literatur [82, 90, 91] ist somit nicht möglich. Dies kann dazu führen, dass räumliche Informationen unterdrückt werden und sich somit die Klassifikationsleistung von merkmals- und signalbasierten Modellen verschlechtert.

Darüber hinaus können beim Anlegen der generalisierten Maske aus Kapitel 5, auf die in diesem Kapitel erhobenen Daten, Ungenauigkeiten zu einer schlechteren Signalqualität führen, mit Auswirkungen auf M2 und M3. Dies ist auf die unterschiedliche Detektion von Körpergelenken zurückzuführen: In Kapitel 5 über die Microsoft Kinect v2 Kamera und in diesem Kapitel über MediaPipe [119]. Ein Umrechnungsfaktor wurde eingeführt, um dies auszugleichen. Insbesondere bei Bewegung kann es dennoch zu Verschiebungen kommen, die dazu führen, dass die ursprünglich ausgewählten Regionen in der generalisierten Maske nicht exakt auf den Aufnahmen der Probanden anliegen und somit spatiale Informationen nicht extrahiert werden können.

Bei der maskenbasierten 2D-Klassifikation (M4) ist zu beachten, dass die Methode zur Entwicklung der individuellen Masken sich von der vorherigen Methode in Kapitel 5 unterscheidet. Dies ist darin begründet, dass das Referenzsignal des Spirometers nicht zur Verfügung steht und somit keine MI ermittelt werden kann. Die Auswirkungen auf die Darstellung der Atemmechanik in beiden Verfahren ist nicht quantifiziert. Dennoch weist die maskenbasierte Klassifikation eine höhere Genauigkeit und bessere Übereinstimmung mit den wahren Werten auf, als merk-

malsbasierte Verfahren. Dies deutet darauf hin, dass auch der in diesem Kapitel entwickelte, korrelationsbasierte Ansatz der Maskenentwicklung in der Lage ist, die spatialen Informationen entsprechend zu verarbeiten und für eine Klassifikation gewinnbringend zu verarbeiten.

Weiteren Einfluss auf die Klassifikation (M2-M5) kann die Signalauswahl haben. Zu einer besseren Vergleichbarkeit untereinander wurden in dieser Arbeit stets 10s Abschnitte verwendet, die zur Hälfte die Tidalatmung und zur anderen Hälfte die forcierte Atmung beinhalten. Es ist jedoch zu beachten, dass die tatsächliche Messung im Bodyplethysmographen für alle Patienten unterschiedlich lang dauert. Durch die fixierte Zeitauswahl sind somit einige Atemzüge der Patienten nicht in der Messung enthalten. Diese können jedoch ebenso wertvolle Informationen zur Klassifikation enthalten. Eine zeitliche Skalierung wurde dennoch nicht eingesetzt, da dies dynamische Atemkennwerte wie FEV_1 verändern würde.

In dieser Arbeit erfolgte die Skalierung jeweils auf die Messung eines Probanden und nicht auf den gesamten Trainingsdatensatz, für die Signal- (M3) und End-to-End-basierte Klassifikation (M5). Dies berücksichtigt insbesondere den Erhalt von wichtigen Parametern, wie das Verhältnis von FEV_1 zu FVC, welches ein wichtiges Kriterium zur Diagnose darstellt [28]. Gleichzeitig kann durch diese Art der Skalierung der relative Bezug zu den Patienten untereinander verloren gehen.

Weitere wichtige Parameter in der Diagnose von respiratorischen Erkrankungen lassen sich im Bodyplethysmographen bestimmen, jedoch gibt es keinen Anhaltspunkt, dass eine optische Messung über Verfahren wie DPG möglich sind. Dazu zählen das Residualvolumen [6], welches immer in der Lunge verbleibt und nach derzeitigem Kenntnisstand nicht von außen bestimmt werden kann. Somit ist eine Klassifikation über DPG der Bodyplethysmographie in diesem Punkt unterlegen.

Modellentwicklung

Maßgebliche Einflussfaktoren auf die Klassifikationsgenauigkeit finden sich auch bei der Entwicklung der Modelle. Der begrenzte Erfolg der maschinellen Lernmethoden lässt sich auch damit erklären, dass eine visuelle Inspektion der Signale (M2 - M4) und der End-to-End-Daten (M5) keine deutlichen Unterschiede

zwischen den Klassen aufzeigt. Daher wird deutlich, dass die Modelle nicht auf simple Merkmale zur Klassifikation zurückgreifen können und eine komplexere Merkmalsberechnung erforderlich ist.

Die korrekte Konfiguration der Hyperparameter eines Modells ist von entscheidender Bedeutung für dessen Leistungsfähigkeit, da eine unangemessene Einstellung zu suboptimalen Ergebnissen führen kann. In dieser Studie wurde eine begrenzte Anzahl von Parametern für die Hyperparametersuche verwendet, was potenziell die Leistung der Modelle beeinflussen kann. Da diese Parameter auch direkten Einfluss auf die Anzahl der Zwischenschichten und Neuronen haben, besteht die Möglichkeit von Overfitting aufgrund zu komplexer Modelle oder Underfitting durch unzureichendes Training. Das Finden eines solchen Gleichgewichts kann durch eine gezielte Optimierung der Hyperparameter erreicht werden. Eine höhere Anzahl an Parametern hätte durch die in dieser Arbeit verwendete Gittersuche zu einem exponentiellen Anstieg der Berechnungszeit geführt. Zukünftig kann der Einsatz von fortgeschrittenen Varianten, wie zum Beispiel evolutionäre Algorithmen [126], zu einer Reduktion der Optimierungszeit führen.

In dieser Arbeit wurde bewusst auf den Einsatz vortrainierter Modelle verzichtet. Dadurch wurden die verwendeten Modelle von Grund auf trainiert, ohne auf bereits vorhandenes Vorwissen zurückzugreifen. Diese Entscheidung wurde getroffen, um die Leistungsfähigkeit und Flexibilität der Modelle in Bezug auf das eigenständige Erlernen von Merkmalen zu evaluieren und zu untersuchen, ohne von externem Vorwissen abhängig zu sein. Transfer-Lernen kann jedoch auch dazu führen, dass antrainiertes Vorwissen zu einer verbesserten Merkmalsextraktion und somit zu einer höheren Klassifikationsgenauigkeit führen kann. Durch zum Beispiel ResNet-Architekturen [127] wäre es denkbar, dass insbesondere eine Oberkörperdetektion bereits implizit enthalten ist und somit nicht erneut trainiert werden muss.

Zusammenfassung der Diskussion

Eine Klassifikation von respiratorischen Erkrankungen ist mittels DPG und KI über verschiedene Modalitäten möglich. Eine geringfügige Klassifikationsgüte ($0 < \kappa \leq 0,20$) ist über merkmalsbasierte (M2) und End-to-End Modelle (M5) möglich. Eine angemessene Klassifikationsgüte ($0,20 < \kappa \leq 0,40$) wird über signal- (M3) oder maskenbasierte (M4) Modelle erreicht. Diese Klassifikationsergebnisse bei der DPG-Messung können durch verschiedene Faktoren in den Datenerfassungs-, Vorverarbeitungs- und Modellbildungsschritten erklärt werden.

Die Identifizierung und Berücksichtigung dieser Faktoren sind entscheidend, um die Leistung der Klassifikationsmodelle zu verbessern. Eine mögliche Verbesserung könnte durch die Einbeziehung weiterer klinischer Parameter und eine optimierte Signalvorverarbeitung erreicht werden. Zudem könnte der Einsatz von Transfer-Learning mit vortrainierten Modellen die Klassifikationsgenauigkeit steigern. Ein weiterer Forschungsbedarf besteht in der Erweiterung der Datenerhebung und Untersuchung zusätzlicher Atemparameter für eine umfassende Diagnose von respiratorischen Erkrankungen.

6.2.4 Korrelationsbasierte Patientenmasken

In diesem Kapitel ist ein neuartiger Ansatz entstanden, Patientenmasken zu generieren, ohne ein Referenzspirometersignal zu verwenden. Stattdessen werden die jeweiligen Korrelationskoeffizienten der einzelnen Pixelsignale summiert und entsprechend dargestellt. Dieser Ansatz (M4) weist von den untersuchten Modalitäten die besten Klassifikationsergebnisse auf, sodass die einzelnen Ergebnisse der individuellen und generalisierten Patientenmasken im Folgenden genauer betrachtet werden.

Generalisierte Maske

Die Summation aller Patientenmasken, vor dem Thresholding und Smoothing, ist in Abbildung 6.6 dargestellt. Es ist zu erkennen, dass die höchste Summe aller Korrelationskoeffizienten nur in einem kleinen Bereich erreicht wird. Es ist grob die Kontur einer Person zu erkennen, mit den schattenhaften Beinen ($< 20\%$) im unteren Bildbereich. Radial verringert sich diese Summe beinahe stufenweise auf 80% und 40% . Dies kann durch eine ungenaue Überlagerung der individuellen Patientenmasken erklärt werden. Insbesondere die rechte Schulter scheint in diesem Fall nicht exakt überlagert zu sein. Dies führt dazu, dass durch die Transformation der Körpergelenkspunkte die anatomischen Strukturen der Masken nicht exakt übereinander gelegt und anschließend gemittelt werden können. Dies kann im Weiteren eine verschwommene und stufenhafte Darstellung erklären. Eine Ursache dafür kann die ungenaue Detektion der Körpergelenke selbst sein, in dieser Arbeit durch MediaPipe [119], oder eine starke Bewegung der Patienten. Die Betrachtung der Region $> 70\%$ weist dennoch Ähnlichkeiten zur generalisierten Maske über MI in Kapitel 5 auf. So ist die Form im Wesentlichen auf den oberen Teil des Thorax begrenzt, mit einer Auslassung im subabdominalen Bereich. Von oben nach unten verläuft die Maske dabei schmaler. Ein wesentlicher Unterschied ist das Nichtvorhandensein der Brustkante, was durch ein fehlerhaftes Überlagern jedoch erklärt werden kann. Die generalisierte Maske weist somit Übereinstimmungen mit den bisherigen Untersuchungen auf. Weitere Ursachen für die nicht korrekte Überlagerung können verdeckte Gelenke sein, wie zum Beispiel im Beckenbereich durch davor verschränkte Arme.

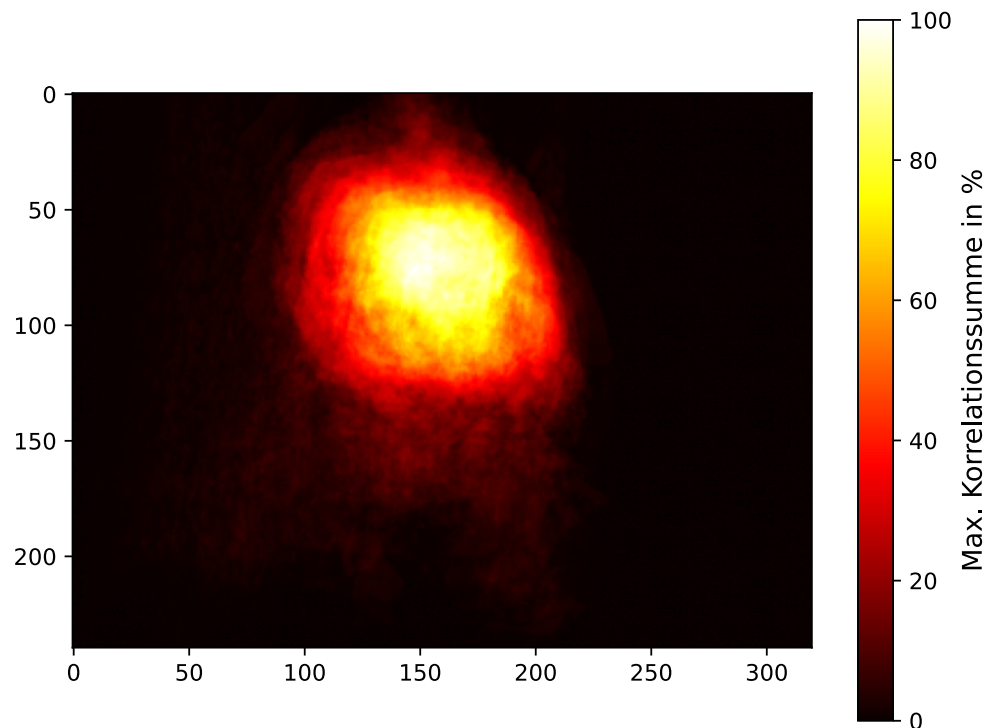


Abbildung 6.6: Summierte korrelationsbasierte Maske aller individuellen Einzelmasken. Darstellung als Relation zum Maximalwert.

Individuelle Masken

Bei Betrachtung der individuellen Masken sind verschiedene Effekte sichtbar. Im Folgenden werden diese an ausgewählten Beispielen verdeutlicht.

Wie bereits in Kapitel 5 sind bei einigen Testpersonen starke Gradienten im Unterbrust-Bereich feststellbar, siehe Abbildung 6.7. Die Entstehung dieses Effekts wird der kranialen Thoraxbewegung, zusammen mit den starken Gradienten in der Longitudinalachse der Thoraxmorphologie, zugeschrieben. Auch die Darstellung von Kleidungsfalten (Abbildung 6.8) ist erneut festzustellen.

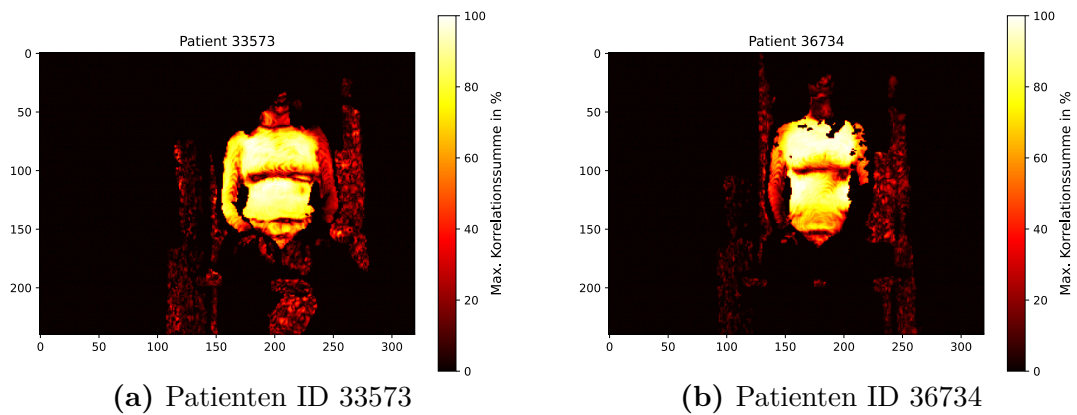


Abbildung 6.7: Korrelationsbasierte Masken mit ausgeprägten Gradienten an der Brustunterkante.

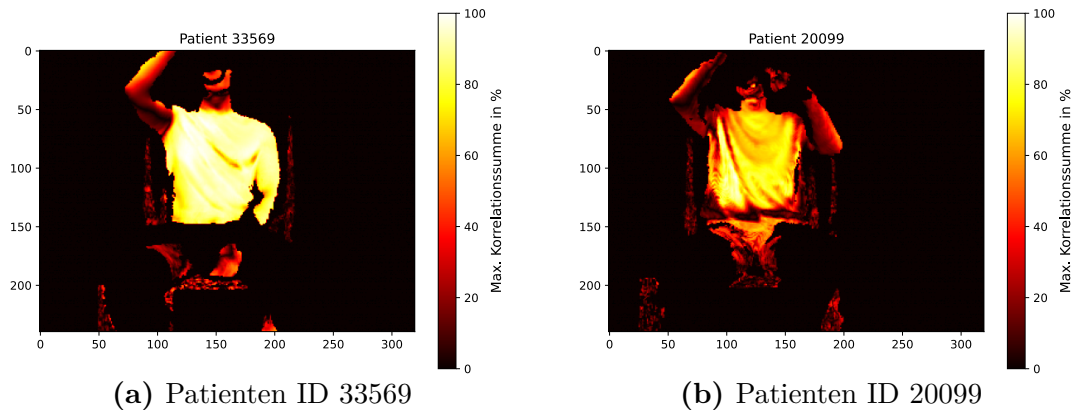


Abbildung 6.8: Korrelationsbasierte Masken mit erkennbaren Kleidungsfalten.

Bei weiteren Testpersonen (32885, 33573, 34189, 11329 und 14457) sind deutliche Diskontinuitäten in der individuellen Segmentierungsmaske feststellbar, beispielhaft in Abbildung 6.9 dargestellt. Die generierten Regionen weisen keine Kontinuität auf und entsprechen nicht den erwarteten anatomischen Regionen. Die größeren Lücken auf der linken Seite der Patienten resultieren aus Bewegungen des Arms während der Messung. Zunächst wird durch Anwendung eines Thresholding-Verfahrens eine Separation des Probanden vom Hintergrund und Vordergrund durchgeführt. Eine ungenaue Einstellung des automatischen Thresholds kann dazu führen, dass ein nach vorne gehaltener Arm, möglicherweise zum Halten des Mundstücks, fälschlicherweise dem Hintergrund zugeordnet wird. Eine Bewegung dieses Arms während der Messung vor dem Brustkorb kann dazu

führen, dass dieser Bereich nicht korrekt berechnet wird und somit keine aussagekräftigen Korrelationskoeffizienten aufweist. Kleinere Löcher innerhalb der segmentierten Thoraxregion könnten möglicherweise auf einen erhöhten Rauschanteil zurückzuführen sein.

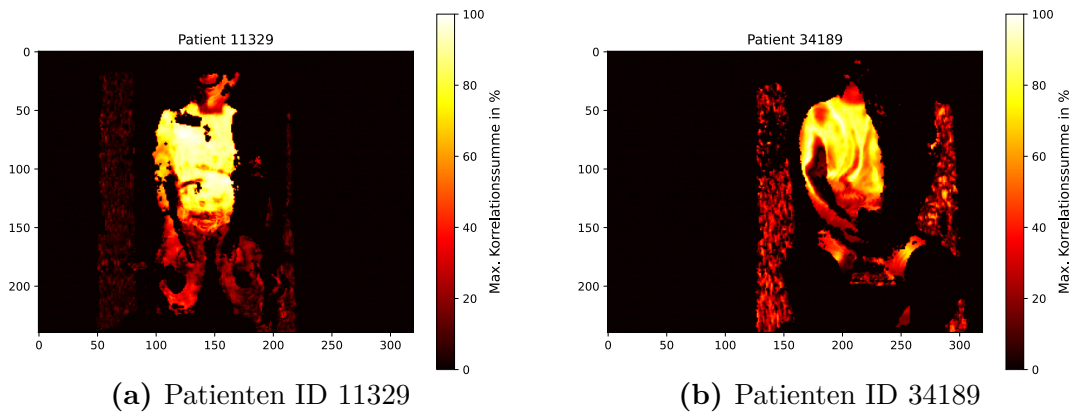


Abbildung 6.9: Korrelationsbasierte Masken mit Lücken in der Segmentierung des Oberkörpers.

Isolinien, die ein topologisches Profil andeuten, sind in den individuellen Masken der Patienten 36702, 5265, 16482, 17261, 18494, 25068, 26525 und beispielhaft in Abbildung 6.10 zu erkennen. Physiologisch sind diese Muster nicht zu erklären. Vielmehr könnte es sich um ein Alias-Artefakt beim Down-Sampling der Eingangsbilder handeln.

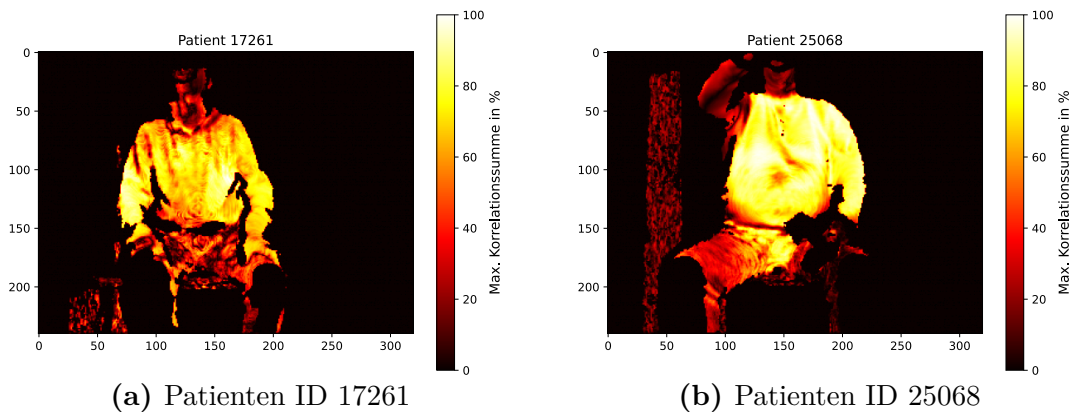


Abbildung 6.10: Korrelationsbasierte Masken mit kreisförmigen Isolinien auf dem Oberkörper.

Insgesamt sind in den individuellen Patientenmasken ähnliche Effekte zu beobachten, wie in der Methode über MI. Dies deutet daraufhin, dass die kalibrationsfreie Methode zur Entwicklung von Patientenmasken prinzipiell geeignet ist. Gleichauf sind Störungen durch Lückenbildung stärker ausgeprägt. Der fehlende Abgleich mit der Grundwahrheit bringt höhere Ungenauigkeiten mit sich. Durch die Erkenntnisse aus Abschnitt 6.2.2 ist bekannt, dass die so ermittelten Masken Informationen enthalten, die zur Klassifikation von respiratorischen Störungen geeignet sind. Diese präzisen Informationen sind jedoch durch eine visuelle Inspektion allein nicht zu erkennen. Somit lässt sich der Zusammenhang der durch die Maske extrahierten physiologischen Informationen mit der Klassifikationsgüte nicht final bewerten.

6.3 Zusammenfassung

Ziel der Untersuchungen in diesem Kapitel war die Beantwortung der folgenden Fragestellungen:

1. Wie kann DPG in Verbindung mit KI genutzt werden, um respiratorische Erkrankungen zu klassifizieren?
2. Inwiefern liefern die aus den extrahierten Masken gewonnenen physiologischen Informationen eine solide Grundlagen für die Klassifikation von respiratorischen Erkrankungen?

Zu diesem Zweck wurde eine Vorabuntersuchung an einem eigens dafür aufgenommenen Datensatz durchgeführt. Dieser Datensatz umfasst 53 verwertbare Aufnahmen von COPD-Patienten (N=18), Asthmatikern (N=26) und Gesunden (N=9), bei denen ein Bodyplethysmograph als Referenzmessgerät verwendet wurde. Mit den so gemessenen Grundwahrheitswerten wurde ein auto-ML Modell trainiert und als Referenz verwendet. Durch die Anwendung der in Kapitel 5 entwickelten generalisierten Masken wurden Signale der Testpersonen extrahiert. Die in Kapitel 4 entwickelte Merkmalsextraktion diente als Eingang in das zuvor ebenfalls erfolgreich getestete SVM-Modell und den auto-ML-Ansatz. Zusätzlich wurde ein 1D-CNN entworfen, um die so extrahierten Signale der einzelnen Testpersonen direkt zu klassifizieren. Da im vorherigen Kapitel die Patientenmasken Hinweise auf die Atemmechanik darstellten, wurde in diesem Kapitel eine korrelationsbasierte Methode entwickelt, welche ohne Verwendung eines Referenz-Spirometersignals

individuellen Patientenmasken erzeugen kann. Diese Masken dienen des Weiteren als Eingang in ein 2D-CNN zur Klassifikation. Über eine skalierte Auswahl einer Region of Interest wird ausgeschlossen, dass Patienteninformationen, wie die Körpergröße, erfasst werden können. Abschließend wurde als End-to-End Ansatz eine hybride CNN-LSTM Architektur, respektive CNN-CNN, entwickelt, welche im ersten Schritt Merkmale aus den Tiefenbildern extrahiert und anschließend über die zweite Stufe auch temporale Informationen zur Klassifikation einsetzt. Die Modelle wurden mittels einer Kreuzvalidierung über die Metriken Accuracy und Cohens Kappa κ validiert.

Die Klassifikationsgenauigkeit des Referenzmodells (Accuracy 62,8% und κ 0,37) kann von den entwickelten Modellen nicht erreicht werden. Merkmalsbasierte (SVM: 49,1%, κ 0,11; Auto-ML: 49,1%, κ 0,13) und End-to-End-Modelle (CNN: 51,9%, κ 0,15; LSTM: 50,9%, κ 0,05) erreichen lediglich eine geringfügige Klassifikationsübereinstimmung. Eine ausreichende Klassifikationsübereinstimmung wird durch die signalbasierte (55,8%, κ 0,24) und maskenbasierte Klassifikation (57,7%, κ 0,28) erreicht. Die Methode zur Entwicklung der korrelationsbasierten Masken erzielt dabei vergleichbare Ergebnisse zu den zuvor entwickelten individuellen MI-basierten Masken. Der Vorteil ist jedoch, dass kein Referenzsignal benötigt wird. Eine fehlerhafte Überlagerung dieser individuellen Masken führt jedoch nicht zu vergleichbaren Ergebnissen mit der generalisierten MI-Maske. Dies ist möglicherweise auf eine ungenaue Detektion der Körpergelenke durch das verwendete Skelettierungsverfahren zurückzuführen. Auf welcher Informationsgrundlage die Klassifikation der maskenbasierten Modelle basiert, lässt sich nicht final beantworten. Weitere Ursachen für die, dem Referenzmodell unterlegene, Klassifikationsgenauigkeit sind in der Datenerhebung, der Signalvorverarbeitung und Modellentwicklung zu suchen. Die Bewegung der Testpersonen und Reflexionen am Glas beeinflussen die Signalqualität bei der Aufnahme. Ungenauigkeiten bei der Körpergelenkdetektion führen dazu, dass durch das Anlegen der Maske zur Signalextraktion nicht alle relevanten Bereiche korrekt ausgewählt werden. Die entwickelten Modelle können aufgrund einer hohen Anzahl an Gewichten, aber einer geringen Anzahl an Hyperparametern möglicherweise nicht optimal auf das Problem angepasst werden.

Insgesamt ist in diesem Kapitel ein neuer Datensatz für die Klassifikation von respiratorischen Erkrankungen mittels DPG entstanden. Die Anwendungen von verschiedenen Modellen des ML auf verschiedenen Modalitäten lässt eine Klassifikation mit geringer Genauigkeit dieser Erkrankungen zu. Insbesondere die korrelationsbasierten, individuellen Patientenmasken scheinen dafür notwendige Informationen zu enthalten, welche möglicherweise auf Asynchronitäten in der Atemmechanik zurückzuführen sind. Um eine gleiche oder bessere Klassifikationsgenauigkeit als die Referenz Bodyplethysmographie zu erreichen, sind weitere Untersuchungen notwendig.

Kapitel 7

Zusammenfassung und Ausblick

7.1 Zusammenfassung der Forschungsergebnisse

DPG ist eine neuartige Methode zur Bestimmung von respiratorischen Parametern. Um aus dem gemessenen Signal auf das Atemvolumen schließen zu können, ist eine Transformation des gemessenen Signals erforderlich. Dies geschieht in der Regel mit Hilfe linearer Skalierungsfunktionen. In bisherigen Ansätzen der Literatur hat sich gezeigt, dass auch somatotypische Faktoren einen Einfluss haben.

In Kapitel 4 wurde ein Ansatz der Literatur aufgegriffen und weiterentwickelt, um aus extrahierten Merkmalen auf Respirationsparameter zu schließen. Insgesamt wurden 379 Merkmale extrahiert, darunter 78 mit physiologischen Bezug, acht bezogen auf die ROI, fünf phänotypische Merkmale der Probanden und 288 automatisch extrahierte Merkmale. Zur Auswahl geeigneter Merkmale wurden Filter, Wrapper und Embedded Methoden eingesetzt. Systematisch wurde der Einsatz von nicht-linearen ML-Modelle (KNN, RF und SVR) analysiert. Für SVC und TV erfolgte die Evaluation auf einem frei verfügbaren Datensatz aus der Literatur mit 148 Aufnahmen. Für SVR und RF kann das Tidalvolumen mit einer Genauigkeit von 11 ± 253 ml bestimmt werden und die Vitalkapazität mit -41 ± 610 ml respektive. Auf Grund der physiologischen Zusammenhänge sinkt die Genauigkeit ohne phänotypische Patientenmerkmale zur TV-Bestimmung auf 18 ± 267 ml und auf -9 ± 737 ml für VC. Die willkürliche Bewegung des gesamten Oberkörpers bei der forcierten Atmung erklärt die größeren Abweichungen bei der

VC. Diese Oberkörperbewegung superpositioniert die zu messende Distanzveränderung des Brustkorbs und führt somit zu einem Störanteil. Die in dieser Arbeit entwickelten ML-Modelle zeigen im Vergleich zu existierenden Modellen auf demselben Datensatz geringere Messabweichungen. Ein weiterer Vorteil besteht darin, dass sie keine Kalibration auf die Probanden erfordern. Andere kalibrationsfreie Modelle in der Literatur verwenden unzureichende Referenzwerte zur Evaluierung, benötigen eine Berechnungszeit, die im Alltag nicht praktikabel ist, oder weisen eine geringere Genauigkeit auf.

Der Untersuchungsgegenstand von Kapitel 5 ist die Verbesserung der Signalextraktion. Dafür wurde der Einsatz von Mutual Information zur Auswahl geeigneter Regionen des Oberkörpers untersucht. Regionen werden identifiziert, welche einen hohen statistischen Zusammenhang zum Signal des Referenzspirometers aufweisen. Die so entstehenden Patientenmasken ermöglichen neue Einblicke in die Betrachtung der Dynamik des Brustkorbs. Die individuellen Masken lassen sich zu einer generalisierten Maske zusammenfassen und auf neue Probanden anwenden. Somit lässt sich die Signalqualität signifikant erhöhen. Eine darauf aufbauende Anwendung von LSTM-Modellen lässt die Messung von dynamischen Respirationsparametern auf diesem Datensatz nicht zu. Eine mögliche Ursache ist im begrenzten Messfenster des Modells zu sehen und in der Abwesenheit von persönlichen Merkmalen. Die Ergebnisse der LSTM-Modelle sind bedeutend ungenauer als merkmalsbasierte ML-Modelle. Durch die verbesserte Signalextraktion durch generalisierte Masken, können jedoch Tidalvolumen (8 ± 206 ml) und Vitalkapazität (43 ± 599 ml) durch die zuvor entwickelten merkmalsbasierten Support-Vector-Regression-Modelle genauer bestimmt werden.

In Kapitel 6 wurden verschiedene Modalitäten der DPG zur Klassifikation von respiratorischen Erkrankungen analysiert. Die Untersuchung erfolgte auf einem Datensatz mit insgesamt 53 verwertbare Aufnahmen von COPD-Patienten (N=18), Asthmatikern (N=26) und Gesunden (N=9), bei denen ein Bodyplethysmograph als Referenzmessgerät verwendet wurde. Klassifikationsmodelle wurden anhand von Merkmalen, Signalen, Masken und den Tiefen-Rohdaten trainiert und gegeneinander evaluiert. Eine ausreichende Klassifikationsübereinstimmung wird dabei ausschließlich durch die signalbasierte (Accuracy 55,8%, Cohens Kappa κ 0,24) und maskenbasierte Klassifikation (57,7%, κ 0,28) erreicht. Die untersuchte Methode zur Entwicklung der korrelationsbasierten Masken erzielt dabei

vergleichbare Ergebnisse zu den zuvor entwickelten individuellen MI-basierten Masken. Es werden dabei jedoch keine Referenzsignale benötigt. Die geringe Klassifikationsgenauigkeit kann auf verschiedene Faktoren zurückgeführt werden, darunter die Datenerhebung, die Signalvorverarbeitung und die Modellentwicklung. Während der Aufnahme können die Bewegung der Testpersonen und Reflexionen am Glas die Qualität des Signals beeinflussen. Darüber hinaus führen Ungenauigkeiten bei der Körpergelenksdetektion dazu, dass nicht alle relevanten Bereiche korrekt ausgewählt werden, wenn die Maske zur Signalextraktion angelegt wird. Es ist möglich, dass die entwickelten Modelle aufgrund einer hohen Anzahl von Gewichten, aber einer geringen Anzahl von Hyperparametern nicht optimal an das vorliegende Problem angepasst werden können.

In dieser Arbeit wurde strukturiert der Einsatz von KI für die DPG zur Bestimmung von respiratorischen Parametern untersucht. Es zeigt sich, dass eine merkmalsbasierte Regression über ML eine kalibrationsfreie DPG ermöglicht – mit einer besseren Genauigkeit, als vergleichbare konventionelle Verfahren. Über die Mutual Information zum Spirometersignal können individuelle Probandenmasken erzeugt werden, welche Hinweise auf atemmechanische Prozesse geben können. Eine Signalextraktion über eine verallgemeinerte Maske ist einer konventionellen ROI überlegen und führt zu einer exakteren Bestimmung der Respirationsparameter. Die Informationen von individuellen Patientenmasken über korrelationsbasierte Ansätze ermöglichen eine Klassifikation von respiratorischen Erkrankungen.

7.2 Wissenschaftliche Schlussfolgerungen und Contributions

Aus dem Stand der Forschung ergaben sich für diese Arbeit wesentliche Arbeitshypothesen. Diese waren die Grundlagen in den Untersuchungen der einzelnen Kapitel. Die Beantwortung der dazugehörigen Forschungsfragen wird im Folgenden zusammenfassend dargestellt.

Der Einsatz von KI mit der Berücksichtigung von phänotypischen Merkmalen eignet sich dazu, eine kalibrationsfreie DPG zu etablieren.

- *Wie genau können Respirationsparameter aus Merkmalen des Tiefen-Zeit-Signals bestimmt werden?*

Bei einer gleichbleibenden Signalaufnahme, über eine konventionelle rechteckige ROI, können die in dieser Arbeit entwickelten Signalmerkmale als Eingang von Modellen des Maschinellen Lernens das Tidalvolumen mit einer Genauigkeit von 11 ± 253 ml bestimmen, respektive die Vitalkapazität mit 11 ± 624 ml. SVR ist zur Bestimmung dieser Größen geeignet. Die MI eignet sich für die Selektion der Merkmale. Zur Ermittlung der Respirationsparameter ist keine Kalibration auf die jeweiligen Probanden notwendig.

- *Kann ein Ensemble-Learning-Ansatz, bestehend aus den einzelnen Modellen, die Genauigkeit der Respirationsparameterbestimmung erhöhen?*

EL erreicht eine Genauigkeit von 18 ± 283 ml für die Bestimmung des Tidalvolumen und -41 ± 610 ml für die Vitalkapazität. EL bietet somit keine nennenswerten Vorteile gegenüber Support Vector Regression.

- *Welchen Einfluss haben phänotypische Merkmale der Probanden auf die Genauigkeit der Respirationsparameterbestimmung?*

Ohne persönliche Merkmale der Probanden sinkt die Genauigkeit für die Bestimmung des Tidalvolumens auf 18 ± 267 ml und für die Vitalkapazität auf -9 ± 737 ml. Insbesondere für die Vitalkapazität sind Alter, Größe, Gewicht, BMI und Geschlecht relevante Merkmale zur Bestimmung der Respirationsparameter.

Statistische Ansätze führen zu einer überlegenen Maske für die Messsignalextraktion in der DPG im Vergleich zu einer konventionellen, rechteckigen ROI.

- *Ist eine automatisch angelegte und an die Physiologie angepasste Region, zur Extraktion eines Tiefen-Zeit-Messsignals, einer konventionellen rechteckigen Form überlegen?*

Die Berechnung von individuellen Patientenmasken über Mutual Information und die Anwendung einer generalisierten Maske auf neue Probanden führen zu einer besseren Signalqualität. Sowohl der Pearson-Korrelationskoeffizient, als auch die Mutual Information und die Dynamic Time Warping Distance zu dem Spirometersignal sind signifikant besser als bei der konventionellen Variante. Tidalvolumen (8 ± 206 ml) und Vitalkapazität (-43 ± 599 ml) können somit genauer bestimmt werden.

- *Welche physiologischen Zusammenhänge lassen sich aus der ausgewählten Region und der Atemmechanik ableiten?*

Durch die individuellen Masken werden anatomische Strukturen, wie zum Beispiel der Rippenbogen sichtbar. Die relationalen Gewichte der Masken deuten auf abdominale oder thorakale Atmung der Probanden hin. Dies lässt sich auf Grund fehlender Daten nicht validieren. Die verallgemeinerte Maske lässt den Rückschluss zu, dass die meisten Informationen aus dem thorakalen Bereich entnommen werden können. Erstmals wird eine Auslassung im unteren Brustbereich vorgeschlagen und eine konkrete Form einer ROI bestimmt. Die reduzierten Informationen in dieser Region werden einer topologischen Kante des Oberkörpers und daraus resultierender Faltenbildung der Kleidung zugeschrieben.

- *Wie genau können Respirationparameter durch eine Volumen-Zeit-Signalapproximation mit Hilfe von rekurrenten neuronalen Netze bestimmt werden?*

Eine Approximation des Volumen-Zeit-Signals mithilfe von LSTM-Modellen führt zu einer ungenaueren Bestimmung der Respirationparameter. Das Tidalvolumen kann mit einer Genauigkeit von 191 ± 255 ml und die Vitalkapazität mit 488 ± 938 ml bestimmt werden.

Die durch die DPG verfügbaren spatio-temporalen Daten sind ausreichend, um eine Klassifikation von respiratorischen Erkrankungen vorzunehmen.

- *Wie kann DPG in Verbindung mit KI genutzt werden, um respiratorische Erkrankungen zu klassifizieren?*

Cohens Kappa $\kappa > 0.20$ deutet auf die prinzipielle Machbarkeit der Klassifikation von Gesunden, Asthmatikern und COPD-Patienten mittels DPG hin. Dabei ist die Klassifikation von individuellen Masken (Genauigkeit 57,7% und Cohens Kappa $\kappa 0,28$) besser als eine Klassifikation des Messsignals (55,8%, $\kappa 0,24$), der daraus extrahierten Merkmale (49,1%, $\kappa 0,13$) oder eines End-to-End ML-Ansatz (51,9%, $\kappa 0,15$). Die Klassifikationsgenauigkeit eines Referenzmodells (62,8%, $\kappa 0,37$), trainiert mit den Daten des Bodyplethysmographen, wird jedoch nicht erreicht.

- *Inwiefern liefern die aus den extrahierten Masken gewonnenen physiologischen Informationen eine solide Grundlagen für die Klassifikation von respiratorischen Erkrankungen?*

Die Klassifikationsgenauigkeit und Cohens Kappa der relevanten Modelle lassen den Rückschluss zu, dass eine Klassifikation anhand der 2D-Patientenmasken möglich ist. Diese Masken werden über korrelationsbasierte Verfahren, ohne Anwendung eines Referenzsignals, erstellt. Eine physiologische Ursache, durch zum Beispiel asynchrone Atemmechanik, lässt sich jedoch nicht feststellen. Auch aufgrund der geringen Klassifikationsgenauigkeit kann diese Frage somit nicht abschließend beantwortet werden.

Insgesamt lassen sich die getroffenen Hypothesen bestätigen, wenngleich nicht alle Forschungsfragen vollumfänglich beantwortet werden können. Die wesentlichen Beiträge dieser Arbeit lassen sich wie folgt zusammenfassen:

1. Die Analyse von Merkmalen und einer kalibrationsfreien Pipeline zur Anwendung von ML-Modellen zur regressiven Bestimmung von Tidalvolumen und Vitalkapazität, inklusive des Nachweis zur Überlegenheit dieser Modelle, gegenüber konventioneller linearer Signalverarbeitung.
2. Die Untersuchung eines Verfahrens zur Darstellung von individuellen Patientenmasken, deren physiologische Interpretation sowie die Verallgemeinerung zur verbesserten Signalextraktion gegenüber einer konventionellen rechteckigen ROI.
3. Der Nachweis zur Machbarkeit einer Klassifikation von respiratorischen Erkrankungen durch die spatio-temporalen Merkmale einer korrelationsbasierten Patientenmaske.

7.3 Ausblick und zukünftige Forschungsfelder

Wichtige zukünftige Herausforderungen sind die Reduktion von Thoraxbewegungen, die Verarbeitung des Messsignals als Zeitsignal zur Extraktion dynamischer Parameter und die Weiterentwicklung der Klassifikationsmodelle. Im Folgenden werden mögliche Ansätze skizziert.

Insbesondere bei der forcierten Atmung wurde, wie bereits in der Literatur, auch in dieser Arbeit eine Superposition der willkürlichen Oberkörperbewegung mit der Brustkorbbewegung beobachtet. Während Signalverarbeitungsmethoden wie die EMD eine Signalverbesserung erzeugen, sollte dem Problem direkt bei der Signalextraktion begegnet werden. Mithilfe der Körpergelenksdetektion wäre es möglich, relevante Fixpunkte wie Hals, Schultern und Hüfte zu detektieren und darüber eine planare Oberfläche zu bestimmen. Die Brustkorbbewegung ist dann relativ zu dieser Ebene zu messen. Da sich die Bewegung des Brustkorbs mit der Bewegung des gesamten Oberkörpers überlagert und letztere als Referenzebene erfasst wird, kann somit die Signalqualität erhöht werden. Es gilt dabei jedoch zu analysieren, welche Thoraxausdehnungen auch in dorsaler und lateraler Richtung erfolgen.

Die Erfassung von dynamischen Messgrößen ist ein wesentlicher Vorteil bei der Anwendung eines linearen Skalierungsfaktors. Eine Möglichkeit, die Performance der LSTM-Modelle zur Zeitsignalapproximation dahingehend zu verbessern, besteht darin, sie mit den extrahierten Merkmalen zu kombinieren und ein hybrides Modell zu entwerfen. Durch die ausgewiesene Fähigkeit, Merkmale zu verarbeiten, mit der Möglichkeit, ein Zeitsignal zu generieren, können somit zukünftig auch dynamische Kenngrößen erfasst werden. Darüber hinaus ist zu untersuchen, inwieweit mehrere Inputsignale die Regression verbessern können. Auch hier können die weiteren Signale aus zum Beispiel die Schulter- und Hüftregionen extrahiert werden, mit dem Ziel, die generelle Oberkörperbewegung zu modellieren.

Für die Klassifikation von respiratorischen Erkrankungen ist insbesondere der Einsatz von vortrainierten Modellen wie zum Beispiel ResNet50 [127], EfficientNet [128], Inceptionv3 [129] oder weiteren Modellarchitekturen wie Transformer-Modellen [130] zu analysieren. Die Auswirkungen des Vorhandenseins von Vorwissen auf die Merkmals- und Signalextraktion ist zu untersuchen. Ziel zukünftiger Studien sollte auch eine parallele Verarbeitung von RGB und Tiefendaten sein, mit dem Ziel Informationen aus beiden Domänen zu kombinieren. Des Weiteren ist zu untersuchen, inwieweit die Schätzung von Tiefeninformationen durch RGB Daten [131, 132] für eine DPG ausreichend ist.

Zukünftig soll eine Portierung von DPG auf Smartphones eine mobile Spirometrie ermöglichen. Dies würde einer Vielzahl von Patienten erlauben, jederzeit, unkompliziert ihre Respirationsparameter zu überwachen. Von einer daraus abgeleiteten besseren Therapie können mehr als 540 Millionen Patienten einer chronischen Atemwegserkrankung [2] profitieren. Darüber hinaus sind kontaktlose und mobile Einsätze der DPG zur situativen Überwachung von Respirationsparametern möglich, wie zum Beispiel in Gefängnissen, nach Tauchgängen, in Katastrophengebieten oder bei Geburtenstationen.

Anhang A

Berechnungen

A.1 Extrahierte Merkmale

Tabelle A.1: Übersicht aller extrahierten Merkmale F . Es werden insgesamt 52 Merkmale der Ruheatmung, 26 Merkmale der forcierten Atmung, 8 Bildbasierte Merkmale, 5 persönliche Merkmale und 288 automatische Merkmale der über die Time Series Feature Extraction Library [98] extrahiert.

#	Name	Beschreibung
1-3	F_{TV,A_i}	Extremwerte der lokalen Minima (Ausatmung) A_i
4-6	$F_{TV,t_{A_i}}$	Zeitpunkte der lokalen Minima (Ausatmung) A_i
7-9	F_{TV,E_i}	Extremwerte der lokalen Maxima (Einatmung) E_i
10-12	$F_{TV,t_{E_i}}$	Zeitpunkte der lokalen Maxima (Einatmung) E_i
13	$F_{TV,MW}$	Mittelwert der Atemvolumina $E_i - A_i$
14-16	$F_{TV,U_i,+}$	Wert zwei Samples nach E_i
17-19	$F_{TV,U_i,-}$	Wert zwei Samples vor E_i
20-22	$F_{TV,euklid, ratio}$	Verhältnis des euklidischen Abstands der Werte einer Sekunde vor zu einer Sekunde nach E_i zum Extremwert E_i
23	$F_{TV,Anstieg,+}$	Größter positiver Anstieg in s_{TV}
24	$F_{TV,t,Anstieg,+}$	Zeitpunkt von #23
25	$F_{TV,Anstieg,-}$	Größter negativer Anstieg in s_{TV}
26	$F_{TV,t,Anstieg,-}$	Zeitpunkt von #25

Fortsetzung auf nächster Seite

Fortsetzung von vorheriger Seite

#	Name	Beschreibung
27	$F_{TV,AF,MW,E}$	Mittelwert der Abstände E_i zu A_i
28	$F_{TV,AF,STD,E}$	Standardabweichung der Abstände E_i zu A_i
29	$F_{TV,V,e}$	Mittelwert der Anstiege E_i zu A_i
30	$F_{TV,AF,MW,A}$	Mittelwert der Abstände A_i zu E_{i-1}
31	$F_{TV,AF,STD,A}$	Standardabweichung der Abstände A_i zu E_{i-1}
32	$F_{TV,V,a}$	Mittelwert der Anstiege A_i zu E_{i-1}
33	$F_{TV,\delta,+}$	Wert des größten Gradienten
34	$F_{TV,t,\delta,+}$	Zeitpunkt des größten Gradienten
35	$F_{TV,\delta,-}$	Wert des kleinsten Gradienten
36	$F_{TV,t,\delta,-}$	Zeitpunkt des kleinsten Gradienten
37	$F_{TV,Peak,STD}$	Standardabweichung der Werte E_i
38	$F_{TV,Sinus_{Ampl}}$	Amplitude einer Sinus-Regression
39	$F_{TV,Sinus_{\omega}}$	Frequenz einer Sinus-Regression
40	$F_{TV,Sinus_{Fehler}}$	Absolute Differenz der Flächeninhalte der Sinus-Regression und s_{TV}
41-43	$F_{TV,FFT,abs,i}$	Größten Beträge im Betragsspektrum
44-46	$F_{TV,FFT,f,abs,i}$	Frequenzen von $F_{TV,FFT,abs,i}$
47-49	$F_{TV,FFT,pow,i}$	Größten Amplituden im Betragsspektrum
50-52	$F_{TV,f,FFT,pow,i}$	Frequenzen von $F_{TV,FFT,pow,i}$
53	$F_{VC,U_i,-}$	Wert zwei Samples vor E_0
54	$F_{VC,Anstieg,+}$	Größter positiver Anstieg in s_{VC}
55	$F_{VC,t,Anstieg,+}$	Zeitpunkt von #54
56	$F_{VC,Anstieg,-}$	Größter negativer Anstieg in s_{VC}
57	$F_{VC,t,Anstieg,-}$	Zeitpunkt von #56
58	$F_{VC,\delta,+}$	Wert des größten Gradienten
59	$F_{VC,t,\delta,+}$	Zeitpunkt des größten Gradienten
60	$F_{VC,\delta,-}$	Wert des kleinsten Gradienten
61	$F_{VC,t,\delta,-}$	Zeitpunkt des kleinsten Gradienten
62-64	$F_{VC,FFT,abs,i}$	Größten Beträge im Betragsspektrum
65-67	$F_{VC,FFT,f,abs,i}$	Frequenzen von $F_{VC,FFT,abs,i}$

Fortsetzung auf nächster Seite

Fortsetzung von vorheriger Seite

#	Name	Beschreibung
68-70	$F_{VC,FFT,pow,i}$	Größten Amplituden im Betragsspektrum
71-73	$F_{VC,f,FFT,pow,i}$	Frequenzen von $F_{VC,FFT,pow,i}$
74	F_{VC,A_i}	Extremwerte der lok. Minima (Ausatmung) A_0
75	$F_{VC,t_{A_i}}$	Zeitpunkte der lok. Minima (Ausatmung) A_0
76	F_{VC,E_i}	Extremwerte der lok. Maxima (Einatmung) E_0
77	$F_{VC,t_{E_i}}$	Zeitpunkte der lok. Maxima (Einatmung) E_0
78	$F_{VC,VC}$	Differenz aus E_0 und A_0
79	$F_{ROI,b}$	Breite der ROI
80	$F_{ROI,h}$	Höhe der ROI
81	$F_{ROI,A}$	Fläche der ROI
82	$F_{ROI,MW}$	Mittelwert innerhalb der ROI
83	$F_{ROI,STD}$	Standardabw. innerhalb der ROI
84	$F_{ROI,MED}$	Median innerhalb der ROI
85	$F_{ROI,MIN}$	Minimum innerhalb der ROI
86	$F_{ROI,MAX}$	Maximum innerhalb der ROI
87	$F_{Subj,Age}$	Alter der Testperson
88	$F_{Subj,Height}$	Größe der Testperson
89	$F_{Subj,Weight}$	Gewicht der Testperson
90	$F_{Subj,Gender}$	Geschlecht der Testperson
91	$F_{Subj,BMI}$	Body-Mass-Index der Testperson
92-235	$F_{TSFEL,TV}$	Autom. Merkmalsextraktion für s_{TV} Bezeichnung und Beschreibung anhand [98]
236-379	$F_{TSFEL,VC}$	Autom. Merkmalsextraktion für s_{VC} Bezeichnung und Beschreibung anhand [98]

Anhang B

Zur Person

B.1 Veröffentlichungen

Veröffentlicht:

- Sohrab, Sebastian; **Wichum, Felix** (2024): Contactless Depth-based Plethysmography. In: American Thoracic Society 2024 International Conference. San Diego.
- Sohrab, Sebastian; **Wichum, Felix** (2024): Kontaktlose Tiefenplethysmographie. In: Kongress der Deutschen Pneumologen. Mannheim.
- Alic, Belmin; Seidl, Roman; **Wichum, Felix**; Wiede, Christian; Seidl, Karsten; Hirtz, Gangolf (2023): Contactless recording of vital parameters to evaluate the quality of sleep in the nursing environment. In: Biomedical Engineering / Biomedizinische Technik 68 (s1), S. 1–261. DOI: 10.1515/bmte-2023-2001.
- **Wichum, Felix**; Grewal, Navraj; Wiede, Christian; Grabmaier, Anton; Seidl, Karsten (2023): Detection of Ventricular Tachycardia Using Artificial Intelligence. In: Biomedical Engineering / Biomedizinische Technik 68 (s1), S. 1–261. DOI: 10.1515/bmte-2023-2001.
- Kobel, Svenja Nicola; **Wichum, Felix**; Önel, Hüseyin; Wiede, Christian; Seidl, Karsten (2023): Optical Non-Contact-Based Vital Sign Monitoring for Care Robots. In: Biomedical Engineering / Biomedizinische Technik 68 (s1), S. 1–261. DOI: 10.1515/bmte-2023-2001.

- **Wichum, Felix**; Wiede, Christian; Seidl, Karsten (2022): Depth-Based Measurement of Respiratory Volumes: A Review. In: *Sensors* (Basel, Switzerland) 22 (24). DOI: 10.3390/s22249680.
- Lazzari, Nico de; **Wichum, Felix**; Götte, Miriam; David, Corinna; Seidl, Karsten; Tewes, Mitra (2022): Entwicklung einer KI-gestützten Bewegungstherapie bei onkologischen Palliativpatienten. In: *B&G Bewegungstherapie und Gesundheitssport* 38 (05), S. 208–215. DOI: 10.1055/a-1909-5766.
- Wuerich, Carolin; **Wichum, Felix**; El-Kadri, Omar; Ghantawi, Kusay; Grewal, Navraj; Wiede, Christian; Seidl, Karsten (2022): Blood Pressure Estimation based on Electrocardiograms. In: *Current Directions in Biomedical Engineering* 8 (2), S. 53–56. DOI: 10.1515/cdbme-2022-1015.
- **Wichum, Felix**; Lazzari, Nico de; Götte, Miriam; David, Corinna; Wiede, Christian; Seidl, Karsten; Tewes, Mitra (2022): Development of an AI-supported exercise therapy for advanced cancer patients. In: *Current Directions in Biomedical Engineering* 8 (2), S. 169–172. DOI: 10.1515/cdbme-2022-1044.
- **Wichum, Felix**; Wiede, Christian; Seidl, Karsten (2022): Kontaktlose Messung der Atemvolumina. In: *Biosignale*. Dresden.
- Hoyer, Ingo; Utz, Alexander; Ludecke, Andre; Richter, Mike; **Wichum, Felix**; Gembaczka, Pierre et al. (2022): Detection of atrial fibrillation with an optimized neural network on a RISC-V-based microcontroller for efficient integration into ECG patches. In: *IEEE Medical Measurements & Applications*. June 22-24, 2022, Giardini Naxos - Taormina, Messina, Italy : conference proceedings. 2022 IEEE International Symposium on Medical Measurements and Applications (MeMeA). Messina, Italy, 6/22/2022 - 6/24/2022. Institute of Electrical and Electronics Engineers; IEEE Instrumentation and Measurement Society. Piscataway, NJ: IEEE, S. 1–6.
- **Wichum, Felix**; Hassel, Jacqueline; Wiede, Christian; Seidl, Karsten (2022): Contactless Measurement of Respiratory Volumes: A Calibration Free Method based on Depth Information. In: *Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies*. 15th International Conference on Bio-inspired Systems and Signal

Processing. Online Streaming, 09.02.2022 - 11.02.2022: SCITEPRESS - Science and Technology Publications, S. 150–158.

- Wuerich, Carolin; **Wichum, Felix**; Wiede, Christian; Grabmaier, Anton (2021): Contactless Optical Respiration Rate Measurement for a Fast Triage of SARS-CoV-2 Patients in Hospitals. In: Proceedings of the International Conference on Image Processing and Vision Engineering. International Conference on Image Processing and Vision Engineering. Online Streaming, 2021.04.28 - 2021.04.30: SCITEPRESS - Science and Technology Publications, S. 29–35.
- **Wichum, Felix**; Wiede, Christian; Seidl, Karsten (2021): Vital Signs and Sensors for Post-Exertional Malaise Prevention. In: Current Directions in Biomedical Engineering 7 (2), S. 371–374. DOI: 10.1515/cdbme-2021-2094.

Eingereicht:

- **Wichum, Felix**; Fabricius, Julian; Lou, Zhouyi; Wiede, Christian; Seidl, Karsten (2024): Enhancing Depth-based Plethysmography through Machine Learning and Region of Interest Masks.
- **Wichum, Felix**; Wiede, Christian; Seidl, Karsten (2024): Classification of Respiratory Diseases.

Literatur

- [1] A. Storm, M. Schumann, J. Marschall, S. Hildebrandt-Heene und H.-D. Nolting, Hrsg. *Gesundheitsreport 2022. Analyse der Arbeitsunfähigkeitsdaten: Risiko Psyche: Wie Depressionen, Ängste und Stress das Herz belasten*. Bd. Band 39. Beiträge zur Gesundheitsökonomie und Versorgungsforschung. Heidelberg: medhochzwei, 2022. ISBN: 978-3-86216-919-1.
- [2] J. B. Soriano, P. J. Kendrick, K. R. Paulson, V. Gupta, E. M. Abrams u. a. „Prevalence and attributable health burdens of chronic respiratory diseases, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017“. In: *The Lancet Respiratory Medicine* 8.6 (2020), S. 585–596. ISSN: 22132600. DOI: 10.1016/s2213-2600(20)30105-3.
- [3] World Health Organization. *World health statistics 2023: monitoring health for the SDGs*. Hrsg. von World Health Organization. Geneva, 2023.
- [4] S. Silbernagl und A. Despopoulos. *Taschenatlas der Physiologie*. 5., komplett überarb. und neu gestaltete Aufl. Flexibles Taschenbuch. Stuttgart und New York: Thieme, 2001. ISBN: 3-13-567705-2.
- [5] W. Piper. „Krankheiten der Atmungsorgane“. In: *Innere Medizin*. Hrsg. von W. Piper. Springer-Lehrbuch. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, S. 181–254. ISBN: 978-3-642-33107-7. DOI: 10.1007/978-3-642-33108-4_2.
- [6] S. Al Dahouk und W. Karges. „Pneumologie“. In: *Innere Medizin*. Hrsg. von W. Karges und S. a. Dahouk. Springer-Lehrbuch. Berlin und Heidelberg: Springer, 2014, S. 107–146. ISBN: 978-3-642-41617-0. DOI: 10.1007/978-3-642-41618-7_3.

- [7] J. Y. Choi und C. K. Rhee. „Diagnosis and Treatment of Early Chronic Obstructive Lung Disease (COPD)“. In: *Journal of Clinical Medicine* 9.11 (2020). DOI: 10.3390/jcm9113426.
- [8] M. R. Miller, J. Hankinson, V. Brusasco, F. Burgos, R. Casaburi u. a. „Standardisation of spirometry“. In: *The European respiratory journal* 26.2 (2005), S. 319–338. ISSN: 0903-1936. DOI: 10.1183/09031936.05.00034805.
- [9] B. L. Graham, I. Steenbruggen, M. R. Miller, I. Z. Barjaktarevic, B. G. Cooper u. a. „Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement“. In: *American Journal of Respiratory and Critical Care Medicine* 200.8 (2019), e70–e88. DOI: 10.1164/rccm.201908-1590ST.
- [10] C. P. Criée, S. Sorichter, H. J. Smith, P. Kardos, R. Merget u. a. „Body plethysmography—its principles and clinical use“. In: *Respiratory Medicine* 105.7 (2011), S. 959–971. DOI: 10.1016/j.rmed.2011.02.006.
- [11] C. Wiede. „Untersuchung robuster Verfahren zur kontaktlosen, optischen Vitalparameterbestimmung zum Einsatz im Bereich Ambient Assisted Living“. Dissertation. Chemnitz: Technische Universität Chemnitz, 2018.
- [12] C. Wuerich, R. Rademacher, C. Wiede und A. Grabmaier. „PTT-based Contact-less Blood Pressure Measurement using an RGB-Camera“. In: *Current Directions in Biomedical Engineering* 7.2 (2021), S. 375–378. DOI: 10.1515/cdbme-2021-2095.
- [13] C. Wuerich, F. Wichum, C. Wiede und A. Grabmaier. „Contactless Optical Respiration Rate Measurement for a Fast Triage of SARS-CoV-2 Patients in Hospitals“. In: *Proceedings of the International Conference on Image Processing and Vision Engineering*. SCITEPRESS - Science and Technology Publications, 2021, S. 29–35. ISBN: 978-989-758-511-1. DOI: 10.5220/0010408400290035.
- [14] F. Wichum, C. Wiede und K. Seidl. „Depth-Based Measurement of Respiratory Volumes: A Review“. In: *Sensors (Basel, Switzerland)* 22.24 (2022). DOI: 10.3390/s22249680.

- [15] A. Breitbarth, T. Schardt, C. Kind, J. Brinkmann, P.-G. Dittrich u. a. „Measurement accuracy and dependence on external influences of the iPhone X TrueDepth sensor“. In: *Photonics and education in measurement science 2019*. Hrsg. von M. Rosenberger, P.-G. Dittrich und B. Zagar. Proceedings of SPIE. Bellingham, Washington, USA: SPIE, 2019, S. 7. ISBN: 9781510629813. DOI: 10.1117/12.2530544.
- [16] K. Hornik. „Some new results on neural network approximation“. In: *Neural Networks* 6.8 (1993), S. 1069–1072. ISSN: 08936080. DOI: 10.1016/S0893-6080(09)80018-X.
- [17] I. Döbel, M. Leis, M. M. Vogelsang, D. Neustroev, H. Petzka u. a. *Maschinelles Lernen: Eine Analyse zu Kompetenzen, Forschung und Anwendung*. Hrsg. von Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V. München, 2018.
- [18] M. Goyal, T. Knackstedt, S. Yan und S. Hassanpour. „Artificial intelligence-based image classification methods for diagnosis of skin cancer: Challenges and opportunities“. In: *Computers in biology and medicine* 127 (2020), S. 104065. DOI: 10.1016/j.compbiomed.2020.104065.
- [19] A. S. Panayides, A. Amini, N. D. Filipovic, A. Sharma, S. A. Tsaftaris u. a. „AI in Medical Imaging Informatics: Current Challenges and Future Directions“. In: *IEEE Journal of Biomedical and Health Informatics* 24.7 (2020), S. 1837–1857. ISSN: 2168-2194. DOI: 10.1109/JBHI.2020.2991043.
- [20] V. Soleimani, M. Mirmehdi, D. Damen, M. Camplani, S. Hannuna u. a. „Depth-based whole body photoplethysmography in remote pulmonary function testing“. In: *IEEE Transactions on Biomedical Engineering* 65.6 (2018), S. 1421–1431. DOI: 10.1109/TBME.2017.2778157.
- [21] M. Stücker, A. Struk, P. Altmeyer, M. Herde, H. Baumgärtl u. a. „The cutaneous uptake of atmospheric oxygen contributes significantly to the oxygen supply of human dermis and epidermis“. In: *The Journal of physiology* 538.Pt 3 (2002), S. 985–994. ISSN: 0022-3751. DOI: 10.1113/jphysiol.2001.013067.
- [22] G. Aumüller, G. Aust, A. Doll, J. Engele, J. Kirsch u. a. *Anatomie*. Duale Reihe. Stuttgart: Thieme, 2006. ISBN: 3-13-136041-0.

- [23] R. Larsen und T. Ziegenfuß. „Physiologie der Atmung“. In: *Beatmung*. Hrsg. von R. Larsen und T. Ziegenfuß. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, S. 19–58. ISBN: 978-3-540-40775-1. DOI: 10.1007/978-3-662-06009-4_2.
- [24] U. Bungeroth. *BASICS Pneumologie*. 2., überarb. Aufl. Basics. München: Elsevier Urban & Fischer, 2010. ISBN: 978-3-437-42237-9.
- [25] J. Haueisen. *Verfahren der Biomedizinischen Messtechnik: Respiratorische Messverfahren: Skript zur Vorlesung*. Ilmenau, 2013.
- [26] H. Rehouma, R. Noumeir, S. Essouri und P. Jouvét. „Advancements in Methods and Camera-Based Sensors for the Quantification of Respiration“. In: *Sensors (Basel, Switzerland)* 20.24 (2020). DOI: 10.3390/s20247252.
- [27] W. Oczenski, Hrsg. *Atmen - Atemhilfen: Atemphysiologie und Beatmungstechnik ; 53 Tabellen*. 8., überarb. Aufl. Stuttgart: Thieme, 2008. ISBN: 978-3-13-137698-5.
- [28] M. Classen, V. Diehl und K. Kochsiek. „Lungen- und Atemwegserkrankungen“. In: *Innere Medizin*. Hrsg. von M. Hallek, M. Böhm und W. Schmiegel. Munich: Urban & Fischer, 2009, S. 275–376. ISBN: 9783437428319. DOI: 10.1016/B978-3-437-42831-9.10010-5.
- [29] G. Sharma und J. Goodwin. „Effect of aging on respiratory system physiology and immunology“. In: *Clinical interventions in aging* 1.3 (2006), S. 253–260. ISSN: 1176-9092. DOI: 10.2147/ciia.2006.1.3.253.
- [30] B. Sasko, U. Thiem, M. Christ, H.-J. Trappe, O. Ritter u. a. „Size matters: An observational study investigating estimated height as a reference size for calculating tidal volumes if low tidal volume ventilation is required“. In: *PLoS ONE* 13.6 (2018), e0199917. DOI: 10.1371/journal.pone.0199917.
- [31] S. W. Littleton. „Impact of obesity on respiratory function“. In: *Respirology* 17.1 (2012), S. 43–49. DOI: 10.1111/j.1440-1843.2011.02096.x.
- [32] A. Lomauro und A. Aliverti. „Sex differences in respiratory function“. In: *Breathe (Sheffield, England)* 14.2 (2018), S. 131–140. ISSN: 1810-6838. DOI: 10.1183/20734735.000318.

- [33] P. Myrianthefs und G. Baltopoulos. „A higher tidal volume may be used for athletes according to measured FVC“. In: *TheScientificWorldJournal* 2013 (2013), S. 526138. DOI: 10.1155/2013/526138.
- [34] J. A. Verschakelen und M. G. Demedts. „Normal thoracoabdominal motions. Influence of sex, age, posture, and breath size“. In: *American Journal of Respiratory and Critical Care Medicine* 151.2 Pt 1 (1995), S. 399–405. DOI: 10.1164/ajrccm.151.2.7842198.
- [35] C. Buess, R. M. Schlegelmilch und R. Kramme. „Lungenfunktionsdiagnostik“. In: *Medizintechnik*. Hrsg. von R. Kramme. Springer Reference Technik. Berlin: Springer, 2017, S. 159–180. ISBN: 3662487713.
- [36] A. P. Silvatti, K. J. Sarro, P. Cerveri, G. Baroni und R. Barros. „A 3D kinematic analysis of breathing patterns in competitive swimmers“. In: *Journal of Sports Sciences* 30.14 (2012), S. 1551–1560. DOI: 10.1080/02640414.2012.713976.
- [37] I. M. Rodrigues, G. Bernardina, K. J. Sarro, G. Baroni, P. Cerveri u. a. „Thoracoabdominal breathing motion pattern and coordination of professional ballet dancers“. In: *Sports Biomechanics* 18.1 (2019), S. 51–62. DOI: 10.1080/14763141.2017.1380223.
- [38] F. C. Wiegandt, D. Biegger, J. F. Fast, G. Matusiak, J. Mazela u. a. „Detection of breathing movements of preterm neonates by recording their abdominal movements with a time-of-flight camera“. In: *Pharmaceutics* 13.5 (2021). DOI: 10.3390/pharmaceutics13050721.
- [39] S. Matsuoka, T. Yamashiro, G. R. Washko, Y. Kurihara, Y. Nakajima u. a. „Quantitative CT assessment of chronic obstructive pulmonary disease“. In: *Radiographics* 30.1 (2010), S. 55–66. DOI: 10.1148/rg.301095110.
- [40] D. E. O’Donnell. „Assessment of bronchodilator efficacy in symptomatic COPD: is spirometry useful?“ In: *Chest* 117.2 Suppl (2000), 42S–7S. DOI: 10.1378/chest.117.2_suppl.42s.
- [41] M. Wohlgemuth, E. L. van der Kooi, J. C. Hendriks, G. W. Padberg und H. T. Folgering. „Face mask spirometry and respiratory pressures in normal subjects“. In: *The European respiratory journal* 22.6 (2003), S. 1001–1006. ISSN: 0903-1936. DOI: 10.1183/09031936.03.00028103.

- [42] J. Boudarham, D. Pradon, H. Prigent, I. Vaugier, F. Barbot u. a. „Opto-electronic vital capacity measurement for restrictive diseases“. In: *Respiratory Care* 58.4 (2013), S. 633–638. DOI: 10.4187/respcare.01916.
- [43] H. Takamoto, H. Nishine, S. Sato, G. Sun, S. Watanabe u. a. „Development and Clinical Application of a Novel Non-contact Early Airflow Limitation Screening System Using an Infrared Time-of-Flight Depth Image Sensor“. In: *Frontiers in Physiology* 11 (2020). DOI: 10.3389/fphys.2020.552942.
- [44] A. B. DuBois, S. Y. Botelho und J. H. Comroe. „A new method for measuring airway resistance in man using a body plethysmograph: values in normal subjects and in patients with respiratory disease“. In: *The Journal of clinical investigation* 35.3 (1956), S. 327–335. ISSN: 0021-9738. DOI: 10.1172/JCI103282.
- [45] S. Giancola, M. Valenti und R. Sala. *A Survey on 3D Cameras: Metrological Comparison of Time-of-Flight, Structured-Light and Active Stereoscopy Technologies*. Cham: Springer International Publishing, 2018. ISBN: 978-3-319-91760-3. DOI: 10.1007/978-3-319-91761-0.
- [46] S. van der Jeught und J. J. Dirckx. „Real-time structured light profilometry: a review“. In: *Optics and Lasers in Engineering* 87 (2016), S. 18–31. ISSN: 01438166. DOI: 10.1016/j.optlaseng.2016.01.011.
- [47] G. Chen, C. Wiede und R. Kokozinski. „Data Processing Approaches on SPAD-Based d-TOF LiDAR Systems: A Review“. In: *IEEE Sensors Journal* 21.5 (2021), S. 5656–5667. ISSN: 1530-437X. DOI: 10.1109/JSEN.2020.3038487.
- [48] I. Guyon, S. Gunn, M. Nikravesh und L. A. Zadeh, Hrsg. *Feature extraction: Foundations and applications*. Bd. 207. Studies in fuzziness and soft computing. Berlin, Heidelberg und New York: Springer, 2006. ISBN: 3540354875.
- [49] C. M. Bishop. *Pattern recognition and machine learning*. Corrected at 8th printing 2009. Information science and statistics. New York, NY: Springer, 2009. ISBN: 978-1-4939-3843-8.

- [50] Y. LeCun, B. Boser, J. Denker, D. Henderson, R. Howard u. a. „Handwritten Digit Recognition with a Back-Propagation Network“. In: *Advances in Neural Information Processing Systems*. Hrsg. von D. Touretzky. Bd. 2. Morgan-Kaufmann, 1989.
- [51] S. Hochreiter und J. Schmidhuber. „Long short-term memory“. In: *Neural computation* 9.8 (1997), S. 1735–1780. ISSN: 0899-7667. DOI: 10.1162/neco.1997.9.8.1735.
- [52] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink und J. Schmidhuber. „LSTM: A Search Space Odyssey“. In: *IEEE transactions on neural networks and learning systems* 28.10 (2017), S. 2222–2232. DOI: 10.1109/TNNLS.2016.2582924.
- [53] C. Cortes und V. Vapnik. „Support-vector networks“. In: *Machine Learning* 20.3 (1995), S. 273–297. ISSN: 0885-6125. DOI: 10.1007/BF00994018.
- [54] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer und P. N. Suganthan. „Ensemble deep learning: A review“. In: *Engineering Applications of Artificial Intelligence* 115 (2022), S. 105151. ISSN: 09521976. DOI: 10.1016/j.engappai.2022.105151.
- [55] T. K. Ho. „Random decision forests“. In: *Proceedings of 3rd International Conference on Document Analysis and Recognition*. IEEE Comput. Soc. Press, 1995, S. 278–282. ISBN: 0-8186-7128-9. DOI: 10.1109/ICDAR.1995.598994.
- [56] L. Feitosa, M. de Britto, A. Aliverti, J. B. Noronha und A. D. de Andrade. „Accuracy of optoelectronic plethysmography in childhood exercise-induced asthma“. In: *Journal of Asthma* 56.1 (2019), S. 61–68. DOI: 10.1080/02770903.2018.1424196.
- [57] N. S. de Faria Júnior, I. R. Santos, I. S. Dias, J. J. Urbano, R. K. Da Palma u. a. „Opto-electronic plethysmography: Noninvasive and accurate measurement of the volume of the chest wall and its different thoraco-abdominal compartments“. In: *Medical Science Technology* 54.1 (2013), S. 147–150. DOI: 10.12659/mst.889664.

- [58] H. Meric, L. Falaize, D. Pradon, D. Orlikowski, H. Prigent u. a. „3D analysis of the chest wall motion for monitoring late-onset Pompe disease patients“. In: *Neuromuscular Disorders* 26.2 (2016), S. 146–152. DOI: 10.1016/j.nmd.2015.11.003.
- [59] W. L. Ripka, L. Ulbricht und P. M. Gewehr. „Application of a photogrammetric kinematic model for prediction of lung volumes in adolescents: A pilot study“. In: *BioMedical Engineering Online* 13.1 (2014). DOI: 10.1186/1475-925X-13-21.
- [60] K. Oh, C. S. Shin, J. Kim und S. K. Yoo. „Level-Set Segmentation-Based Respiratory Volume Estimation Using a Depth Camera“. In: *IEEE Journal of Biomedical and Health Informatics* 23.4 (2019), S. 1674–1682. DOI: 10.1109/JBHI.2018.2870859.
- [61] S. Ostadabbas, N. Sebkhi, M. Zhang, S. Rahim, L. J. Anderson u. a. „A Vision-Based Respiration Monitoring System for Passive Airway Resistance Estimation“. In: *IEEE Transactions on Biomedical Engineering* 63.9 (2016), S. 1904–1913. DOI: 10.1109/TBME.2015.2505732.
- [62] V. Soleimani, M. Mirmehdi, D. Damen, J. Dodd, S. Hannuna u. a. „Remote, Depth-Based Lung Function Assessment“. In: *IEEE Transactions on Biomedical Engineering* 64.8 (2017), S. 1943–1958. DOI: 10.1109/TBME.2016.2618918.
- [63] W. Imano, K. Kameyama, M. Hollingdal, J. Refsgaard, K. Larsen u. a. „Non-Contact Respiratory Measurement Using a Depth Camera for Elderly People“. In: *Sensors (Basel, Switzerland)* 20.23 (2020). DOI: 10.3390/s20236901.
- [64] J. M. Harte, C. K. Golby, J. Acosta, E. F. Nash, E. Kiraci u. a. „Chest wall motion analysis in healthy volunteers and adults with cystic fibrosis using a novel Kinect-based motion tracking system“. In: *Medical and Biological Engineering and Computing* 54.11 (2016), S. 1631–1640. DOI: 10.1007/s11517-015-1433-1.
- [65] S. J. Cala, C. M. Kenyon, G. Ferrigno, P. Carnevali, A. Aliverti u. a. „Chest wall and lung volume estimation by optical reflectance motion analysis“. In: *Journal of applied physiology (Bethesda, Md. : 1985)* 81.6

- (1996), S. 2680–2689. ISSN: 8750-7587. DOI: 10.1152/jappl.1996.81.6.2680.
- [66] P. S. Addison, P. Smit, D. Jacquet, A. P. Addison, C. Miller u. a. „Continuous non-contact respiratory rate and tidal volume monitoring using a Depth Sensing Camera“. In: *Journal of Clinical Monitoring and Computing* 36(3).36 (2022), S. 657–665. ISSN: 1573-2614. DOI: 10.1007/s10877-021-00691-3.
- [67] S. Kreiss, L. Bertoni und A. Alahi. *OpenPifPaf: Composite Fields for Semantic Keypoint Detection and Spatio-Temporal Association*. 2021. DOI: 10.48550/arXiv.2103.02440.
- [68] MMPose Contributors. *MMPose: OpenMMLab Pose Estimation Toolbox and Benchmark*. 2020.
- [69] D. Osokin. „Real-time 2D Multi-Person Pose Estimation on CPU: Lightweight OpenPose“. In: *arXiv preprint arXiv:1811.12004*. 2018.
- [70] A. Kendall, M. Grimes und R. Cipolla. „PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization“. In: S. 2938–2946.
- [71] V. Soleimani, M. Mirmehdi, D. Damen, S. Hannuna, M. Camplani u. a. „Remote pulmonary function testing using a depth sensor“. In: *IEEE Biomedical Circuits and Systems Conference: Engineering for Healthy Minds and Able Bodies, BioCAS 2015 - Proceedings* (2015). DOI: 10.1109/BioCAS.2015.7348445.
- [72] G. Bradski. „The OpenCV Library“. In: *Dr. Dobb’s Journal of Software Tools* 25.11 (2000), S. 120–125.
- [73] C. Arrieta, C. Sing-Long, S. Uribe, M. E. Andia, P. Irarrazaval u. a. „Level set segmentation with shape prior knowledge using intrinsic rotation, translation and scaling alignment“. In: *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2015, S. 1568–1571. ISBN: 978-1-4799-2374-8. DOI: 10.1109/ISBI.2015.7164178.
- [74] T. M. Seppanen, J. Kananen, N. Kai, O.-P. Alho und T. Seppanen. „Accurate measurement of respiratory airflow waveforms using depth data“. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society*.

- Annual International Conference 2015* (2015), S. 7857–7860. DOI: 10.1109/embc.2015.7320213.
- [75] V. Soleimani. „Remote Depth-based Photoplethysmography in Pulmonary Function Testing“. Dissertation. Bristol: University of Bristol, 2018.
- [76] S. Transue, P. Nguyen, T. Vu und M.-H. Choi. „Real-Time Tidal Volume Estimation Using Iso-surface Reconstruction“. In: *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*. IEEE, 2016, S. 209–218. ISBN: 978-1-5090-0943-5. DOI: 10.1109/CHASE.2016.72.
- [77] F. Bellemare, A. Jeanneret und J. Couture. „Sex differences in thoracic dimensions and configuration“. In: *American Journal of Respiratory and Critical Care Medicine* 168.3 (2003), S. 305–312. DOI: 10.1164/rccm.200208-876oc.
- [78] H. Kaneko und J. Horie. „Breathing movements of the chest and abdominal wall in healthy subjects“. In: *Respiratory Care* 57.9 (2012), S. 1442–1451. DOI: 10.4187/respcare.01655.
- [79] I. Vogiatzis, A. Aliverti, S. Golemati, O. Georgiadou, A. Lomauro u. a. „Respiratory kinematics by optoelectronic plethysmography during exercise in men and women“. In: *European Journal of Applied Physiology* 93.5-6 (2005), S. 581–587. DOI: 10.1007/s00421-004-1249-4.
- [80] C. Sharp, V. Soleimani, S. Hannuna, M. Camplani, D. Damen u. a. „Toward respiratory assessment using depth measurements from a time-of-flight sensor“. In: *Frontiers in Physiology* 8.FEB (2017). DOI: 10.3389/fphys.2017.00065.
- [81] C. Massaroni, E. Carraro, A. Vianello, S. Miccinilli, M. Morrone u. a. „Optoelectronic Plethysmography in Clinical Practice and Research: A Review“. In: *Respiration; international review of thoracic diseases* 93.5 (2017), S. 339–354. DOI: 10.1159/000462916.
- [82] M.-C. Yu, J.-L. Liou, S.-W. Kuo, M.-S. Lee und Y.-P. Hung. „Noncontact respiratory measurement of volume change using depth camera“. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. An-*

- nual International Conference 2012 (2012)*, S. 2371–2374. DOI: 10.1109/embc.2012.6346440.
- [83] J. Kempfle und K. van Laerhoven. „Breathing In-Depth: A Parametrization Study on RGB-D Respiration Extraction Methods“. In: *Frontiers in Computer Science* 3 (2021). DOI: 10.3389/fcomp.2021.757277.
- [84] Z. Zoumot, A. Lomauro, A. Aliverti, C. Nelson, S. Ward u. a. „Lung Volume Reduction in Emphysema Improves Chest Wall Asynchrony“. In: *Chest* 148.1 (2015), S. 185–195. DOI: 10.1378/chest.14-2380.
- [85] V. Soleimani, M. Mirmehdi, D. Damen und J. Dodd. „Markerless Active Trunk Shape Modelling for Motion Tolerant Remote Respiratory Assessment“. In: *Proceedings - International Conference on Image Processing, ICIP (2018)*. DOI: 10.1109/ICIP.2018.8451202.
- [86] B. A. Reyes, N. Reljin, Y. Kong, Y. Nam und K. H. Chon. „Tidal Volume and Instantaneous Respiration Rate Estimation using a Volumetric Surrogate Signal Acquired via a Smartphone Camera“. In: *IEEE Journal of Biomedical and Health Informatics* 21.3 (2017), S. 764–777. DOI: 10.1109/JBHI.2016.2532876.
- [87] V. Soleimani, M. Mirmehdi, D. Damen, S. Hannuna und M. Camplani. „3D Data Acquisition and Registration Using Two Opposing Kinects“. In: *2016 Fourth International Conference on 3D Vision (3DV)*. IEEE, 2016, S. 128–137. ISBN: 978-1-5090-5407-7. DOI: 10.1109/3DV.2016.21.
- [88] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih u. a. „The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis“. In: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 454.1971 (1998), S. 903–995. ISSN: 1364-5021. DOI: 10.1098/rspa.1998.0193.
- [89] R. Courtney, M. Cohen und J. Reece. „Comparison of the Manual Assessment of Respiratory Motion (MARM) and the Hi Lo Breathing Assessment in determining a simulated breathing pattern“. In: *International Journal of Osteopathic Medicine* 12.3 (2009), S. 86–91. ISSN: 17460689. DOI: 10.1016/j.ijosm.2008.10.002.

- [90] H. Chen, Y. Cheng, D. Liu, X. Zhang, J. Zhang u. a. „Color structured light system of chest wall motion measurement for respiratory volume evaluation“. In: *Journal of biomedical optics* 15.2 (2010), S. 026013. DOI: 10.1117/1.3368680.
- [91] R. Priori, A. Aliverti, A. L. Albuquerque, M. Quaranta, P. Albert u. a. „The effect of posture on asynchronous chest wall movement in COPD“. In: *Journal of applied physiology (Bethesda, Md. : 1985)* 114.8 (2013), S. 1066–1075. ISSN: 8750-7587. DOI: 10.1152/jappphysiol.00414.2012.
- [92] S. Motamedi-Fakhr, R. C. Wilson und R. Iles. „Tidal breathing patterns derived from structured light plethysmography in COPD patients compared with healthy subjects“. In: *Medical devices (Auckland, N.Z.)* 10 (2017), S. 1–9. ISSN: 1179-1470. DOI: 10.2147/MDER.S119868.
- [93] K. Ashutosh, R. Gilbert, J. H. Auchincloss und D. Peppi. „Asynchronous breathing movements in patients with chronic obstructive pulmonary disease“. In: *Chest* 67.5 (1975), S. 553–557. DOI: 10.1378/chest.67.5.553.
- [94] V. Soleimani, M. Mirmehdi und D. Damen. *A Dataset for Depth-Based Whole Body Photoplethysmography in Remote Pulmonary Function Testing*. 2018. DOI: 10.5523/bris.1tqzx39mzkw832msuvy3obktqi.
- [95] V. Kumar. „Feature Selection: A literature Review“. In: *The Smart Computing Review* 4.3 (2014). DOI: 10.6029/smartcr.2014.03.007.
- [96] M. Llobera. „Building Past Landscape Perception With GIS: Understanding Topographic Prominence“. In: *Journal of Archaeological Science* 28.9 (2001), S. 1005–1014. ISSN: 03054403. DOI: 10.1006/jasc.2001.0720.
- [97] C. Wuerich, E.-M. Humm, C. Wiede und G. Schiele. „A Feature-based Approach on Contact-less Blood Pressure Estimation from Video Data“. In: *2022 30th European Signal Processing Conference (EUSIPCO)*. IEEE, 2022, S. 1343–1347. ISBN: 978-90-827970-9-1. DOI: 10.23919/EUSIPCO55093.2022.9909563.
- [98] M. Barandas, H. Liu, T. Schultz und H. Gamboa. „TSFEL: Time Series Feature Extraction Library“. In: *SoftwareX* 11 (2020), S. 100456.

-
- [99] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion u. a. „Scikit-learn: Machine Learning in Python“. In: *Journal of Machine Learning Research* 12 (2011), S. 2825–2830.
- [100] B. H. Menze, B. M. Kelm, R. Masuch, U. Himmelreich, P. Bachert u. a. „A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data“. In: *BMC bioinformatics* 10 (2009), S. 213. DOI: 10.1186/1471-2105-10-213.
- [101] M. Korobov. *ELI5*. 2021.
- [102] P. Probst, M. Wright und A.-L. Boulesteix. „Hyperparameters and Tuning Strategies for Random Forest“. In: (2018). DOI: 10.48550/arXiv.1804.03515.
- [103] D. Basak, S. Pal und D. Patranabis. „Support Vector Regression“. In: *Neural Information Processing – Letters and Reviews* 11 (2007).
- [104] O. Sagi und L. Rokach. „Ensemble learning: A survey“. In: *WIREs Data Mining and Knowledge Discovery* 8.4 (2018). ISSN: 1942-4787. DOI: 10.1002/widm.1249.
- [105] K. Jamieson und A. Talwalkar. *Non-stochastic Best Arm Identification and Hyperparameter Optimization*. 2015. DOI: 10.48550/arXiv.1502.07943.
- [106] J. M. Bland und D. G. Altman. „Statistical methods for assessing agreement between two methods of clinical measurement“. In: *Lancet (London, England)* 1.8476 (1986), S. 307–310. ISSN: 0140-6736.
- [107] M. K. Chowdary, J. Anitha und D. J. Hemanth. „Emotion Recognition from EEG Signals Using Recurrent Neural Networks“. In: *Electronics* 11.15 (2022), S. 2387. DOI: 10.3390/electronics11152387.
- [108] M. Kass, A. Witkin und D. Terzopoulos. „Snakes: Active contour models“. In: *International Journal of Computer Vision* 1.4 (1988), S. 321–331. ISSN: 0920-5691. DOI: 10.1007/BF00133570.
- [109] F. Wichum, C. Wiede und K. Seidl. „Kontaktlose Messung der Atemvolumina“. In: *Biosignale*. 2022.
- [110] P. Senin. *Dynamic Time Warping Algorithm Review*. Honolulu, USA, 2008.

- [111] H. Hwang, K. Lee und E. C. Lee. „A Real-Time Remote Respiration Measurement Method with Improved Robustness Based on a CNN Model“. In: *Applied Sciences* 12.22 (2022), S. 11603. DOI: 10.3390/app122211603.
- [112] K. M. Abubeker und S. Baskar. „B2-Net: an artificial intelligence powered machine learning framework for the classification of pneumonia in chest x-ray images“. In: *Machine Learning: Science and Technology* 4.1 (2023), S. 015036. DOI: 10.1088/2632-2153/acc30f.
- [113] J. R. Landis und G. G. Koch. „The Measurement of Observer Agreement for Categorical Data“. In: *Biometrics* 33.1 (1977), S. 159. ISSN: 0006341X. DOI: 10.2307/2529310.
- [114] N. A. Köhler, C. Nöh, M. Geis, S. Kerzel, J. Frey u. a. „Influence of Ambient Factors on the Acquisition of 3-D Respiratory Motion Measurements in Infants—A Preclinical Assessment“. In: *IEEE Transactions on Instrumentation and Measurement* 72 (2023), S. 1–10. ISSN: 0018-9456. DOI: 10.1109/TIM.2023.3271724.
- [115] R. A. Jörres. *Lungenfunktion/Spirometrie in der Praxis: Durchführung und Interpretation*. München, 17.10.2015.
- [116] M. Feurer, A. Klein, K. Eggenberger, J. Springenberg, M. Blum u. a. „Efficient and Robust Automated Machine Learning“. In: *Advances in Neural Information Processing Systems* 28 (2015). 2015, S. 2962–2970.
- [117] M. Feurer, A. Klein, K. Eggenberger, J. T. Springenberg, M. Blum u. a. „Auto-sklearn: Efficient and Robust Automated Machine Learning“. In: *Automated Machine Learning*. Hrsg. von F. Hutter, L. Kotthoff und J. Vanschoren. The Springer Series on Challenges in Machine Learning. Cham: Springer International Publishing, 2019, S. 113–134. ISBN: 978-3-030-05317-8. DOI: 10.1007/978-3-030-05318-5_6.
- [118] P. Gijsbers, E. LeDell, J. Thomas, S. Poirier, B. Bischl u. a. *An Open Source AutoML Benchmark*. 2019.
- [119] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja u. a. *MediaPipe: A Framework for Building Perception Pipelines*. 2019.

-
- [120] F. Zhang, V. Bazarevsky, A. Vakunov, A. Tkachenka, G. Sung u. a. *MediaPipe Hands: On-device Real-time Hand Tracking*. 2020. DOI: 10.48550/arXiv.2006.10214.
- [121] N. Ganapathy, R. Swaminathan und T. M. Deserno. „Deep Learning on 1-D Biosignals: a Taxonomy-based Survey“. In: *Yearbook of medical informatics* 27.1 (2018), S. 98–109. DOI: 10.1055/s-0038-1667083.
- [122] S. Kiranyaz, T. Ince, O. Abdeljaber, O. Avci und M. Gabbouj. „1-D Convolutional Neural Networks for Signal Processing Applications“. In: *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, S. 8360–8364. ISBN: 978-1-4799-8131-1. DOI: 10.1109/icassp.2019.8682194.
- [123] D. Chen. *Orbbec Astra Pro (UVC support) Datasheet*. 2017.
- [124] J. Jiao, L. Yuan, W. Tang, Z. Deng und Q. Wu. „A Post-Rectification Approach of Depth Images of Kinect v2 for 3D Reconstruction of Indoor Scenes“. In: *ISPRS International Journal of Geo-Information* 6.11 (2017), S. 349. DOI: 10.3390/ijgi6110349.
- [125] A. Papi, F. Blasi, G. W. Canonica, L. Morandi, L. Richeldi u. a. „Treatment strategies for asthma: reshaping the concept of asthma management“. In: *Allergy, asthma, and clinical immunology : official journal of the Canadian Society of Allergy and Clinical Immunology* 16 (2020), S. 75. ISSN: 1710-1484. DOI: 10.1186/s13223-020-00472-8.
- [126] L. Tani, D. Rand, C. Veelken und M. Kadastik. „Evolutionary algorithms for hyperparameter optimization in machine learning for application in high energy physics“. In: *The European Physical Journal C* 81.2 (2021). ISSN: 1434-6044. DOI: 10.1140/epjc/s10052-021-08950-y.
- [127] K. He, X. Zhang, S. Ren und J. Sun. „Deep Residual Learning for Image Recognition“. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2016, S. 770–778. ISBN: 978-1-4673-8851-1. DOI: 10.1109/CVPR.2016.90.
- [128] M. Tan und Q. V. Le. „EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks“. In: (2019). DOI: 10.48550/arXiv.1905.11946.

- [129] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens und Z. Wojna. „Rethinking the Inception Architecture for Computer Vision“. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2016, S. 2818–2826. ISBN: 978-1-4673-8851-1. DOI: 10.1109/CVPR.2016.308.
- [130] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones u. a. *Attention Is All You Need*. 2017. DOI: 10.48550/arXiv.1706.03762.
- [131] N. U. Islam und J. Park. „Depth Estimation From a Single RGB Image Using Fine-Tuned Generative Adversarial Network“. In: *IEEE Access* 9 (2021), S. 32781–32794. DOI: 10.1109/ACCESS.2021.3060435.
- [132] Z. Chen, V. Badrinarayanan, G. Drozdov und A. Rabinovich. „Estimating Depth from RGB and Sparse Sensing“. In: *Computer Vision – ECCV 2018*. Hrsg. von V. Ferrari, M. Hebert, C. Sminchisescu und Y. Weiss. Bd. 11208. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018, S. 176–192. ISBN: 978-3-030-01224-3. DOI: 10.1007/978-3-030-01225-0_11.