

**SENSORFUSION FÜR DAS ROBUSTE UND KONTAKTLOSE
MONITORING DER VITALPARAMETER VON FRÜHGEBORENEN
AUF DER NEONATOLOGISCHEN INTENSIVSTATION**

**Vom Promotionsausschuss der
Technischen Universität Hamburg**
zur Erlangung des akademischen Grades

Doktor-Ingenieurin (Dr.-Ing.)

genehmigte Dissertation (Monografie)

von
Johanna Gleichauf, MSc.

aus
Nürnberg

2025

Erstprüfer: Prof. Dr.-Ing. habil. Alexander Kölpin
Zweitprüfer: Prof. Dr.-Ing. Tobias Knopp
Drittprüferin: Prof. Dr. rer. hum. biol. Christine Niebler

Mündliche Prüfung: 02. September 2025

Creative Commons Lizenzvertrag

Der Text steht, soweit nicht anders gekennzeichnet, unter der Creative-Commons-Lizenz Namensnennung 4.0 (CC BY 4.0). Das bedeutet, dass er vervielfältigt, verbreitet und öffentlich zugänglich gemacht werden darf, auch kommerziell, sofern dabei stets der Urheber, die Quelle des Textes und o. g. Lizenz genannt werden. Die genaue Formulierung der Lizenz kann unter <https://creativecommons.org/licenses/by/4.0/legalcode.de> aufgerufen werden.

DOI: <https://doi.org/10.15480/882.16153>

 <https://orcid.org/0000-0002-2702-1077>

Gendererklärung

Aus Gründen der Lesbarkeit wird bei Personenbezeichnungen und personenbezogenen Substantiven und Pronomen die männliche Sprachform verwendet. Im Sinne der Gleichbehandlung gelten die entsprechenden Begriffe als geschlechtsneutral. Die verwendete Sprachform beinhaltet keinerlei Wertung und impliziert keine Benachteiligung aufgrund des Geschlechts.

Kurzfassung

Herkömmliches kontaktbehaftetes Monitoring der Vitalparameter von Frühgeborenen beinhaltet das Risiko, Patienten zu schädigen. Daher ergibt sich die Frage, ob es möglich ist, die Vitalparameter von Neu- und Frühgeborenen sicher und ausreichend genau, aber kontaktlos zu erfassen. Es wurde ein berührungsloses System entwickelt, das Atemrate, Herzrate und Körpertemperatur der Neonaten überwacht. Zum Einsatz kommen verschiedene Sensoren, wie eine RGB-Kamera, eine Thermokamera, mehrere Time-of-Flight (ToF) Kameras und ein Radarsensor, die miteinander kombiniert oder deren Daten fusioniert werden. Die jeweils von den Sensoren erfasste Region of Interest (ROI) wird automatisch über ein mit fusionierten Bildern trainiertes Neuronales Netz (YOLOv3) bestimmt. Die zur Berechnung der Vitalparameter benötigten Algorithmen wurden entwickelt bzw. vom Stand der Technik weiterentwickelt. Als beste Verfahren erwiesen sich in der Validierungsphase die Auswertung von ToF-Kameradaten für die Erfassung der Atemrate und die Temperaturdetektion mit Thermokamera und Referenztemperatursensor. Die Messung der Herzrate mittels RGB-Kamera liefert nicht die klinisch notwendige Genauigkeit. Es konnte gezeigt werden, dass unter optimalen Bedingungen die kontaktlose Vitalparameterdetektion von Atemrate und Körpertemperatur mit guten Ergebnissen möglich ist.

Abstract

The current monitoring of vital signs of preterm neonates bears the risk to injure the little patients. Therefor, the question arises if it is possible to determine the vital parameters of neonates and preterm neonates precisely and robustly via non-contact methods. A non-contact system was developed for monitoring the respiratory and heart rates as well as the core temperature of the neonates. Different sensors such as an RGB camera, a thermal camera, several time-of-flight cameras and a radar sensor are combined or their data are fused. The region of interest detected by each sensor is automatically determined by a neural network (YOLOv3) which was trained with fused images. The necessary algorithms to calculate the vital parameters were developed or existing algorithms improved. The best approach found during the validation phase is the evaluation of ToF camera data for determining the respiratory rate and the temperature detection via thermal camera with reference temperature sensor. The evaluation of the heart rate using an RGB camera does not deliver the clinical necessary precision. It was shown that under optimal conditions a non-contact detection of the respiratory rate and the body temperature is possible with good results.

Inhaltsverzeichnis

Abkürzungsverzeichnis	xii
1 Einleitung und Fragestellung	1
1.1 Motivation	1
1.2 Problemstellung	2
2 Grundlagen	3
2.1 Medizinische Grundlagen	3
2.1.1 Atemfrequenz	3
2.1.2 Herzfrequenz	4
2.1.3 Körpertemperatur	4
2.1.4 Physiologische Veränderungen	5
2.1.4.1 Oxygenierung des Blutes	5
2.1.4.2 Herzrate	5
2.1.4.3 Körpertemperatur	5
2.2 Technische Grundlagen	6
2.2.1 Grundlagen der Signalverarbeitung	6
2.2.1.1 Fourier-Transformation	6
2.2.1.2 Filter	7
2.2.2 Robot Operating System	7
2.2.3 Neuronale Netze	7
2.2.3.1 Künstliche Neuronale Netze	8
2.2.3.2 Convolutional Neural Networks	9
2.2.3.3 Residual Neural Networks	12
2.2.3.4 RetinaNet	13
2.2.3.5 Neuronales Netz YOLOv3	14
3 Stand der Technik	16
3.1 Überwachung von Frühgeborenen	16
3.1.1 Überwachung der Atmung	16
3.1.1.1 Sauerstoffsättigungs-Messung	16
3.1.1.2 Transkapnode	17
3.1.1.3 Impedanz-Pneumographie	17
3.1.1.4 Beatmungsgeräte	18
3.1.2 Überwachung des Herzens	18
3.1.3 Blutdruck	18
3.1.4 Temperaturmessung	18
3.1.5 Einflüsse auf das derzeit übliche Monitoring	19
3.1.5.1 Leitfähigkeit der Elektroden	19

3.1.5.2	Bewegung des Kindes	19
3.2	Kontaktlose Überwachung von Vitalparametern	19
3.2.1	Atemüberwachung mittels visueller Sensorik	19
3.2.1.1	Atemüberwachung mittels Infrarotkamera	19
3.2.1.2	Atemüberwachung mittels RGB-Kamera	20
3.2.1.3	Detektion der Brustkorbbewegung	20
3.2.1.4	Einsatz von Tiefenkameras	21
3.2.2	Herzrate mittels visueller Sensorik	22
3.2.2.1	Eulerian Video Magnification	22
3.2.2.2	Herzrate mittels Infrarotkamera	23
3.2.3	Radarsensorik zur Vitalparameterdetektion	23
3.2.3.1	UWB-Radar	23
3.2.3.2	FMCW-Radar	24
3.2.3.3	CW-Radar	25
3.2.4	Frequenzanalyse zur Herz- bzw. Atemfrequenzbestimmung	27
3.2.4.1	Herzfrequenzbestimmung	27
3.2.4.2	Atemfrequenzbestimmung	27
3.2.5	Einsatz von Deep Learning bei der Vitalparameter-Detektion	28
3.2.6	Sensorfusion zur Vitalparameter-Detektion	28
3.2.6.1	Thermo-RGB-Fusion zur Vitalparameter-Detektion	28
3.2.6.2	Fusion Time-of-Flight-Kamera und Radar-Sensor	31
3.2.7	Detektion des Gesichts	31
3.2.7.1	Gesichtsdetektion basierend auf Neuronalen Netzen und Bild- verarbeitung	31
3.2.7.2	Detektion von Körperarealen bei Neonaten basierend auf Neuronalen Netzen und Bildverarbeitung	32
3.2.7.3	Gesichtsdetektion basierend auf fusionierten Daten	35
3.2.7.4	Neuronale Netze mit Fusionsdaten trainiert	35
3.2.8	Kontaktlose Temperaturmessung	35
3.2.9	Zusammenfassung	38
4	Theoretischer Ansatz	39
4.1	Anforderungsanalyse	39
4.2	Konzept	40
4.3	Sensoraufbau	41
4.3.1	Kameras und Radar-Sensor	41
4.3.1.1	RGB-Kamera	43
4.3.1.2	Thermokamera	43
4.3.1.3	Tiefenkamera	43
4.3.1.4	Six-Port-Radarsensor	44
4.3.1.5	Temperatur- und Feuchtigkeits-Sensor	48
4.3.2	Sensorbox Konzept	48
4.3.3	Referenzmessung	48
4.3.3.1	Klinischer Goldstandard	48
4.3.3.2	Baby-Thorax-Simulator	49

4.4	Kalibrierung	51
4.4.1	Intrinsische Kalibrierung	51
4.4.2	Extrinsische Kalibrierung	52
4.5	Anforderungen an Algorithmen	52
4.5.1	Detektion von Gesichts- und Körperarealen	52
4.5.2	Kontaktlose Atemdetektion	52
4.5.3	Kontaktlose Herzrattendetektion	53
4.5.4	Kontaktlose Temperaturdetektion	53
4.6	Datengewinnung	53
5	Entwurf	54
5.1	Sensorkonfiguration	54
5.1.1	RGB-Kamera	54
5.1.2	Thermokamera	54
5.1.3	Mikrowellen-interferometrischer Radarsensor	55
5.1.4	Tiefenkameras	57
5.1.4.1	Structured Light Kamera	57
5.1.4.2	ToF Kamera	57
5.1.5	Temperatur- und Feuchtigkeits-Sensor	58
5.1.6	Aufbau der Sensorbox	59
5.1.7	Systemrechner	61
5.1.8	Referenzmessung	61
5.1.8.1	Klinischer Goldstandard	61
5.1.8.2	SimBaby als Referenz für Structured light camera	62
5.1.8.3	Simulationspuppe Paul	62
5.1.8.4	Baby-Thorax-Simulator	63
5.2	Kalibrierung	64
5.2.1	Intrinsische Kalibrierung	64
5.2.1.1	Intrinsische Kalibrierung der RGB-Kamera	64
5.2.1.2	Intrinsische Kalibrierung der Thermokamera	66
5.2.2	Extrinsische Kalibrierung	68
5.2.2.1	Kreis-Detektion	68
5.2.2.2	RGB-ToF und Thermal-ToF-Fusion	68
5.3	Algorithmik	70
5.3.1	Verwendete Algorithmen	70
5.3.1.1	Eulerian Video Magnification	70
5.3.1.2	Algorithmen zur Frequenzanalyse	70
5.3.2	Detektion von Gesichts- und Körperarealen basierend auf Neuronalen Netzen	71
5.3.2.1	Neuronale Netze	71
5.3.3	Kontaktlose Atemdetektion	71
5.3.3.1	RGB-Kamera	71
5.3.3.2	Thermokamera	72
5.3.3.3	Structured-Light-Kamera	72
5.3.3.4	ToF-Kamera	73

5.3.3.5	Radar-Sensor	76
5.3.3.6	Synchronisierung ToF-Kamera und Radar-Sensor	77
5.3.4	Kontaktlose Herzrattendektion	77
5.3.4.1	RGB	78
5.3.4.2	Radar	80
5.3.5	Kontaktlose Temperaturdetektion	81
5.3.5.1	Kalman-Filter	81
5.3.5.2	Differenz periphere und Körperkerntemperatur	81
6	Validierung: Experimente und Ergebnisse	83
6.1	Machbarkeitsstudie	83
6.1.1	Datengewinnung	83
6.1.2	Goldstandards der Vitalparameterdetektion	84
6.2	Atemrattendektion mit Structured-Light-Kamera	84
6.2.1	Messaufbau	84
6.2.2	Goldstandard	84
6.2.3	Datengewinnung	84
6.2.4	Messergebnisse	85
6.3	Synchrone Auswertung der Atemrate mit ToF-Kamera und Radar-Sensor	88
6.3.1	Messaufbau	88
6.3.2	Goldstandard	88
6.3.3	Baby-Thorax-Simulator	88
6.3.3.1	Datengewinnung	88
6.3.3.2	Ergebnisse der synchronen Auswertung Baby-Thorax-Simulator	89
6.3.4	Simulationspuppe Paul	96
6.3.4.1	Datengewinnung	96
6.3.4.2	Ergebnisse Simulationspuppe Paul	97
6.4	Detektion von Körperarealen mit Neuronalen Netzen	99
6.4.1	Training der Neuronalen Netze	99
6.4.2	Neuronale Netze Trainingsverläufe	108
6.4.2.1	RetinaNet	108
6.4.2.2	YOLOv3	108
6.4.3	Ergebnisse Testdatensätze	109
6.4.3.1	Ergebnisse mit Testdatensätzen 1 und 2 nach Trainingsphase 1	110
6.4.3.2	Ergebnisse mit Testdatensätzen 1 und 2 nach Trainingsphase 2	111
6.5	Atemdetektion mit Time-of-Flight-Kamera und Neuronalem Netz	112
6.6	Atemdetektion mit Thermokamera und Neuronalem Netz	114
6.7	Herzrattendektion mit RGB-Kamera und Neuronalem Netz	115
6.8	Herzrattendektion mit Radar-Sensor	117
6.8.1	Software Simulation	117
6.8.2	Hardware Simulation	118
6.8.2.1	Messaufbau	118
6.8.2.2	Goldstandard	118
6.8.2.3	Datengewinnung	118
6.8.2.4	Ergebnisse am Baby-Thorax-Simulator	118

6.9 Temperaturdetektion mit Thermokamera und Neuronalem Netz 119

7 Diskussion 121

7.1 Atemratendetektion mit Structured-Light-Kamera 121

7.2 Atemratendetektion mit ToF-Kamera und Radar-Sensor 121

7.2.1 Einordnung der Ergebnisse 121

7.2.2 Vergleich mit dem Stand der Technik 122

7.3 Detektion von Körperarealen mit Neuronalen Netzen 123

7.3.1 Diskussion der Trainings-Ergebnisse 123

7.3.1.1 Training mit Fusions-Daten vs. Training mit RGB- oder Thermo-Daten 123

7.3.1.2 Vergleich RetinaNet und YOLOv3 124

7.3.2 Vergleich mit dem Stand der Technik 125

7.4 Atemdetektion mit ToF-Kamera und Neuronalem Netz 126

7.4.1 Einordnung der Ergebnisse 126

7.4.2 Vergleich mit dem Stand der Technik 126

7.5 Atemdetektion mit Thermokamera und Neuronalem Netz 127

7.5.1 Einordnung der Ergebnisse 127

7.5.2 Vergleich mit dem Stand der Technik 127

7.6 Herzratendetektion mit RGB-Kamera und Neuronalem Netz 128

7.6.1 Einordnung der Ergebnisse 128

7.6.2 Vergleich mit dem Stand der Technik 128

7.7 Herzratendetektion mit Radar-Sensor 129

7.7.1 Einordnung der Ergebnisse 129

7.7.2 Vergleich mit dem Stand der Technik 129

7.8 Temperaturdetektion 129

7.8.1 Einordnung der Ergebnisse 129

7.8.2 Vergleich mit dem Stand der Technik 130

7.9 Fazit 130

8 Zusammenfassung 131

Literaturverzeichnis 132

Abbildungsverzeichnis 153

Abkürzungsverzeichnis

AMV	Atemminutenvolumen
AP	Average Precision
CNN	Convolutional Neural Networks
CW	Continuous Wave
DFT	Diskrete Fourier-Transformation
EAR	Edge Aware Resample
EKG	Elektrokardiogramm
EVM	Eulerian Video Magnification
FFT	Fast Fourier Transformation
FIR	Finite Impulse Response
FMCW	Frequency Modulated Continuous Wave
FOV	Field of View
FPN	Feature Pyramid Net
fps	frames per second
GPU	Graphics Processing Unit
HMM	Hidden Markov Modell
HOG	Histogram of oriented gradients
HSMM	Hidden Semi-Markov Modell
ICNIRP	International Commission on Non-Ionizing Radiation Protection
IPG	Impedanz-Pneumographie
IQ	in-phase and quadrature
IR	Infrared
LMEDS	Least Median of Squares
LSTM	Long Short-Term Memory
LWIR	Long wave Infrared
MLS	Moving Least Square
NCC	normalized cross correlation
NIR	Near Infrared
PCL	Point Cloud Library
PDE	Partial Differential Equations
PPGi	Photoplethysmographic imaging
PSD	Power Spectral Density
RDS	Respiratory Distress Syndrome
RMSE	Root Mean Square Error
ROI	Region of Interest
ROS	Robot Operating System
SGD	Stochastic Gradient Descent
SNR	Signal to Noise Ratio

SSW	Schwangerschaftswoche
SWIR	Shortwave Infrared
ToF	Time of Flight
UDP	User Datagram Protocol
UWB	Ultra Wide Band
VG	Voxel Grid
WHO	World Health Organization

1 Einleitung und Fragestellung

1.1 Motivation

Jedes Jahr kommen weltweit ca. 15 Millionen Frühgeborene zur Welt. Allein in Deutschland liegt die Rate bei 9,0 %, d.h. im Jahr 2008 wurden 215.634 Frühgeborene geboren. [MM] Die World Health Organization (WHO) definiert Frühgeborene als Neonaten, die vor der 37. Schwangerschaftswoche (SSW) geboren werden. Je nach Alter werden die Neonaten in folgende Kategorien eingeteilt [WHO]:

- extrem Frühgeborenes (weniger als 28 Wochen)
- sehr Frühgeborenes (28 bis 32 Wochen)
- leicht bis spät Frühgeborenes (32 bis 37 Wochen)

Die Ursachen für Frühgeburten liegen bei Mehrlings-Schwangerschaften, Infektionen, chronischen Veranlagungen, wie Diabetes und Bluthochdruck oder genetischen Einflüssen [WHO]. Um die Überlebenschancen zu erhöhen, benötigen Frühgeborene eine weitreichende und komplexe medizinische Überwachung auf der neonatologischen Intensivstation. Über verschiedenste Kabel, Katheter, Elektroden und Schläuche sind die Kinder mit den Überwachungsgeräten verbunden. Elektrokardiogramm (EKG), Pulsoxymeter, Beatmungsgeräte und Blutdruckmessgeräte sind nur einige Geräte, die zum Monitoring eingesetzt werden. Alle diese Geräte haben eines gemeinsam: Die Messverfahren benötigen Kontakt zum Körper des Neugeborenen. Da die Frühchen unterentwickelt sind, kann dies zu fatalen Folgen führen, wie Allergien, Druckstellen und Hautirritationen bis hin zum Ablösen der Haut beim Entfernen von EKG-Elektroden. Meist können die Neugeborenen eine gesunde Körpertemperatur von 37 °C nicht selber halten und müssen daher in einem Brutkasten gelagert werden, der durch eine angemessene Wärme und Feuchtigkeit das Kind unterstützt.



Abb. 1.1: Frühgeborenes auf der Intensivstation [mam]

1.2 Problemstellung

Die bisherigen Monitoring-Methoden verlangen direkten Kontakt mit der Haut des Neonaten und gehen mit den obengenannten Problemen wie Druckstellen, Exzema, Erschwernis für die Pflege, etc. einher. Zudem ist es möglich, dass Ungenauigkeiten durch das Schwitzen des Kindes auftreten oder dass die Bewegung des Kindes zu Fehlmessungen und Fehlalarmen führt. Daher wäre eine kontaktlose Überwachungs-Methode sehr hilfreich. Bestehende kontaktlose Monitoring-Methoden sind nur für Erwachsene ausgelegt, nicht unter allen Bedingungen robust und liefern somit nicht die erforderliche Zuverlässigkeit und Genauigkeit. Daraus leitet sich die Fragestellung für diese Arbeit ab:

Ist es möglich, eine kontaktlose Monitoring-Methode zu entwickeln, die bei Neu- und Frühgeborenen robuste und ähnlich genaue Ergebnisse für die Vitalparameterdetektion von Herz- und Atemrate, sowie Körpertemperatur liefert wie die kontaktbehafteten Goldstandardverfahren?

Diese Arbeit ist folgendermaßen aufgebaut: In Kapitel 2 werden medizinische und technische Grundlagen vorgestellt, die für die restliche Arbeit Voraussetzung sind. Kapitel 3 gibt einen Überblick über die bisherigen Monitoring-Methoden auf der Neonatologischen Intensivstation und über bestehende kontaktlose Überwachungs-Ansätze. Im nächsten Schritt werden ein theoretischer Ansatz (Kapitel 4) und ein Konzept als Entwurf des gesamten Systems vorgestellt (Kapitel 5). In Kapitel 6 werden die durchgeführten Experimente und Ergebnisse präsentiert. Die Ergebnisse werden daraufhin diskutiert (Kapitel 7) und zuletzt zusammengefasst (Kapitel 8).

2 Grundlagen

Zunächst werden medizinische und technische Grundlagen vorgestellt, die zum Verständnis der restlichen Arbeit benötigt werden.

2.1 Medizinische Grundlagen

Um den Gesundheitszustand der Frühgeborenen abschätzen zu können, werden kontinuierlich die Vitalparameter des Kindes gemessen. Vitalparameter sind messbare Größen verschiedener Körperfunktionen. Dazu zählen insbesondere die Herzfrequenz, die Atemfrequenz, der Blutdruck und die Körpertemperatur. [MMA]

2.1.1 Atemfrequenz

Die Atemfrequenz eines Menschen variiert je nach Alter. Die untenstehende Tabelle, ein angepasster Auszug der Checklisten Pädiatrie und Neonatologie [Bit10], fasst diese zusammen.

Tab. 2.1: Normwerte Atemfrequenz [Bit10]

Lebensalter	Atemrate (Anzahl der Atemzüge pro Minute)	Frequenz in Hz
Erwachsener	9 - 36	0,15 - 0,6
Kleinkind	16 - 25	0,27 - 0,417
Säugling	20 - 30	0,33 - 0,5
Neugeborenes	25 - 40	0,67 - 0,83
Frühgeborenes	48 - 60	0,8 - 1,0

Bei Frühgeborenen ist die Lunge meist noch unterentwickelt, was eine eigenständige Atmung erschwert. Aus diesem Grund wird häufig eine künstliche Beatmung benötigt.

2.1.2 Herzfrequenz

Je nach Belastung und Wachzustand unterscheiden sich die Normwerte der Herzfrequenz [Bit10], die in Tabelle 2.2 zu finden sind.

Tab. 2.2: Normwerte Herzfrequenz in Beats pro Minute (BPM) und Hz im Schlaf, Wachzustand und bei Belastung [Bit10]

Alter	im Schlaf		im Wachzustand		bei Belastung	
	Hf (BPM)	Hf (Hz)	Hf (BPM)	Hf (Hz)	Hf (BPM)	Hf (Hz)
Frühgeborene	100 - 150	1,67 - 2,5	100 - 200	1,67 - 3,33	< 220	< 3,67
Neugeborene	80 - 160	1,33 - 2,67	100 - 200	1,67 - 3,33	≤ 220	≤ 3,67
1. Woche - 3. Monat	80 - 200	1,33 - 3,33	100 - 220	1,67 - 3,67	≤ 220	≤ 3,67
3. Monat - 2. Lebensjahr	70 - 120	1,17 - 2	80 - 150	1,33 - 2,5	< 200	< 3,3

2.1.3 Körpertemperatur

Bei Untersuchungen von Dawkins et al. wurde festgestellt, dass eine Wärmeproduktion im braunen Fettgewebe stattfindet, wenn das Kind Kälte ausgesetzt ist. Zudem ist der Nackenbereich des Neonaten wegen des braunen Fettgewebes wärmer als alle anderen Körperteile. [DS65] Dieses ist gut in Abbildung 2.1 zu sehen.

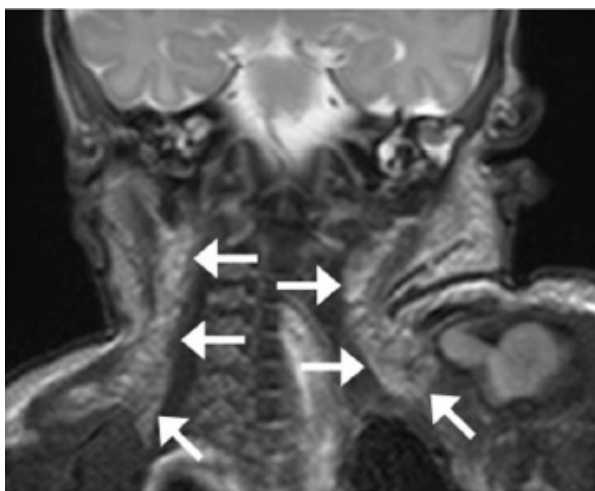


Abb. 2.1: MRT-Bild eines zwei Tage alten Neugeborenen. Das (warme) braune Fettgewebe ist mit den Pfeilen markiert. [Hu+14]

Der Unterschied zwischen peripherer (Haut-) Temperatur und Körperkerntemperatur liegt bei Neu- und Frühgeborenen zwischen 0,5 °C und 1,5 °C. [Jor10]. Andrew Lyon beschreibt, dass

die Differenz für Frühgeborene zwischen 0,5 °C und 1,0°C liegt [Lyo04]. Lyon et al. haben untersucht, wie sich die Differenz je nach Alter des Kindes verhält. Tabelle 2.3 zeigt den Zusammenhang [Lyo+97].

Tab. 2.3: Differenz zwischen peripherer und Körperkerntemperatur in Abhängigkeit des Alters [Lyo+97]

Alter (Tage)	Differenz peripherer Temp. - Körperkerntemp. (°C)
1	0,4
2	0,8
3	0,9
4	1,0
5	1,0

Ein typisches Problem bei Frühgeborenen ist, dass die Regulierung der Körpertemperatur nicht wie bei einem ausgereiften Neugeborenen funktioniert. Der Inkubator unterstützt das Kind dabei, eine konstante Körpertemperatur von 37 °C zu halten. Daher müssen insbesondere die Eltern wissen, dass die Inkubatorklappen nicht zu lange geöffnet werden dürfen. [Spa09] Die Temperatur des Neonaten wird kontinuierlich überwacht, da ein Anstieg oder Abfall der Temperatur ein Indikator für Infektionen oder Krankheiten sein kann [EBB05]. Zudem führt die Temperaturabnahme um 1 °C unter 36 °C zu einer um 28% höheren Sterblichkeit [EFC].

2.1.4 Physiologische Veränderungen

Es gibt verschiedene Einflüsse, die physiologische Veränderungen und somit Änderungen der Vitalparameter der Früh- und Neugeborenen hervorrufen können. Dazu zählen Schmerz, Stress, Stimulation (positive wie auch negative), Lärm, Licht, Temperatur und Feuchtigkeit. Die Aufzählung ist erweiterbar.

2.1.4.1 Oxygenierung des Blutes

Der Sauerstoffpartialdruck kann während einer Stresssituation wie z.B. trachealer Intubation oder Absaugung absinken [LLP80]. Bei Neugeborenen ohne Erkrankungen wird in diesen Situationen normalerweise ein Anstieg vermerkt [Wol87].

2.1.4.2 Herzrate

Die Herzrate kann je nach Art des Reizes und Kontextes entweder ansteigen oder abfallen. Beispielsweise wird in einer Stresssituation wie Blutabnahme der Sympathikus die Herzrate erhöhen. Erhält das Kind jedoch zuvor Glukose, wodurch der Parasympathikus das Kind beruhigt bis es schläft, sinkt die Herzrate auch insgesamt ab. Somit ist bei Neonaten keine grundsätzlich gültige Aussage zu treffen, wie das Kind auf eine Reizsituation reagiert. [Spa09]

2.1.4.3 Körpertemperatur

Es gibt mehrere Krankheitsbilder (wie z.B. Infektionen, periphere Vasokonstriktion, bestimmte Medikamente oder nur eine nasse Windel [Mit96]), die eine Veränderung der Körpertemperatur

(symptomatisch) mit sich bringen. Die zwei typischsten Verhalten bei Neu- und Frühgeborenen werden im Folgenden vorgestellt.

Hyperthermie Eine Hyperthermie beschreibt eine zu hohe Körperinnentemperatur, die z.B. durch einen Hitzschlag ausgelöst werden kann [HJ11]. Bei Neu- und Frühgeborenen kann diese auch im Rahmen der Therapie einer Hypothermie auftreten. Um die Neonaten wieder aufzuwärmen, werden neben Wärmematten auch Wärmefolien verwendet. In einigen Studien wurde erkannt, dass diese im schlimmsten Fall zu einer Hyperthermie führen [McC+].

Hypothermie Eine Hypothermie beschreibt das Gegenteil der Hyperthermie. Hierbei ist der Körper unterkühlt. [HJ11] Liegt eine Asphyxie, also ein Sauerstoffmangel während der Geburt vor, kann Hypothermie als Therapie angewandt werden. Dies verlangsamt den Ablauf potentiell schädigender Kaskaden im zentralen Nervensystem, sodass die Organe, die zu wenig Sauerstoff bekommen haben, ohne größere Schäden weiter funktionieren. [FMH] Nach der Kühlung wird das Kind dann wieder langsam auf eine normale Körperinnentemperatur erwärmt.

2.2 Technische Grundlagen

An dieser Stelle sollen einige technische Grundlagen vorgestellt werden, die Voraussetzung für die vorliegende Arbeit sind.

2.2.1 Grundlagen der Signalverarbeitung

In diesem Abschnitt werden die benötigten Verfahren zur Verarbeitung der Signale beschrieben.

2.2.1.1 Fourier-Transformation

Um Daten aus dem Zeitbereich in ein Frequenzspektrum zu wandeln, eignet sich die Fourier-Transformation. Hierbei wird das Signal in Sinuswellen verschiedener Amplituden und Frequenzen aufgeteilt. Die Fourier-Transformation ist mathematisch folgendermaßen definiert [FB08]:

$$(2.1) \quad x(f) = Fx(t) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt$$

mit f als Frequenz und $x(t)$ als Signal im Zeitbereich. Die Fourier-Transformierte eines Zeitsignals wird als Spektrum bezeichnet. Liegen diskrete Werte vor, muss die Diskrete Fourier-Transformation (DFT) angewandt werden. Die schnelle Fourier-Transformation (Fast Fourier Transformation (FFT)) ist eine optimierte Version der DFT mit geringerem Berechnungsaufwand. [17]

Fenster Wenn die FFT auf ein Signal mit einer nicht ganzzahligen Periodenzahl angewandt wird, resultiert meist eine spektrale Streuung. Um dies zu verhindern bzw. den Effekt zu minimieren, kann ein Fensterverfahren verwendet werden. [17] Die Fenstergröße beeinflusst die zeitliche bzw. Frequenzauflösung. Das Fenster kann durch Bins, die kleine Ausschnitte des Spektrums

darstellen, in gleichmäßige Teile unterteilt werden. Durch die Veränderung der Anzahl an Bins im Fenster kann die Auflösung der Frequenz erhöht werden. [IRC]

2.2.1.2 Filter

Ideales Bandpass-Filter Das ideale Bandpass-Filter wird genutzt, um unerwünschte Frequenzen zu dämpfen und nur die gewünschten durchzulassen. Ideale Bandpassfilter sind dazu geeignet, Frequenzen in einem bestimmten Frequenzband zu isolieren oder herauszufiltern. [Ele]

Savitzky-Golay-Filter Savitzky-Golay-Filter sind ein Typ der Finite Impulse Response (FIR)-Filter, die zur Glättung von Signalen verwendet werden [Mat]. Die hier genutzte Implementierung basiert auf Gram-Polynomen wie in der Publikation von Gorry beschrieben [Gor90]. Über die Fenstergröße und den Grad des Polynoms werden die Eigenschaften des Filters parametrisiert. Das Filter setzt im Zeitbereich an und erlaubt es, für verschiedene Frequenzbereiche eine unterschiedliche Glättung einzustellen.

2.2.2 Robot Operating System

Das Robot Operating System (ROS) ist eine Open Source Middleware, die insbesondere bei der Programmierung von mobilen Robotern eingesetzt wird. Sie erlaubt die Kommunikation zwischen mehreren Prozessen durch das Versenden von Nachrichten (Topics). Die Verwendung von ROS erleichtert die Sensordatenverarbeitung. Kamera- oder andere Sensor-Topics können in einem Bagfile-Datensatz aufgenommen, gespeichert und danach wieder abgespielt werden. ROS stellt Bibliotheken und Werkzeuge zur Verfügung, die Softwareentwicklern helfen, Roboterapplikationen zu entwickeln [ROS].

2.2.3 Neuronale Netze

Laut Definition ist ein neuronales Netz eine Kombination mehrerer Neuronen zur Erfüllung einer bestimmten Aufgabe [KSJ95]. Neuronen sind insbesondere aus der Medizin (Neurologie) bekannt (siehe Abbildung 2.2) und haben mindestens einen Ein- und Ausgang. Sobald das Eingangssignal einen Schwellwert übertrifft, "feuert" das Neuron ein Ausgangssignal. Die Informationsweiterleitung erfolgt entweder über Botenstoffe oder durch Potentialänderung. 1943 entstanden erste mathematische Modelle von McCulloch und Pitts [MP90]. Künstliche Neuronale Netze bauen auf dieser Forschung auf.

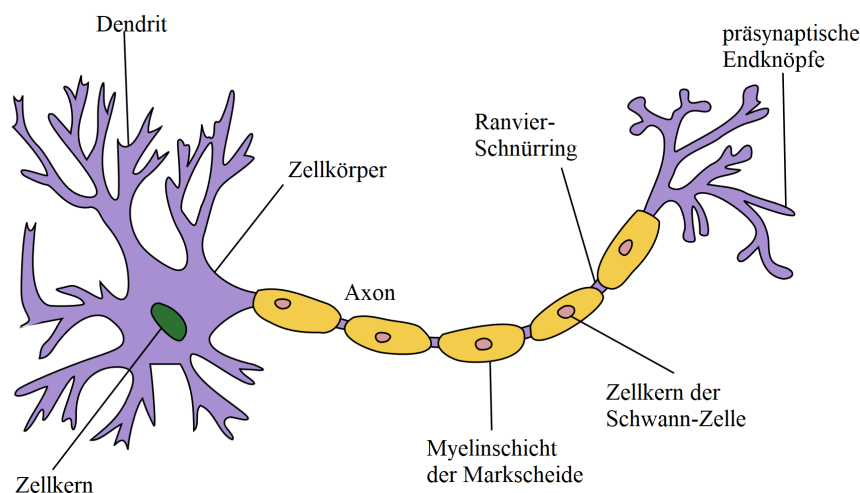


Abb. 2.2: Aufbau eines biologischen Neurons mit Axon. Auf die Dendriten werden Erregungen von anderen Neuronen übertragen und summiert. Abhängig vom Ergebnis der Summation am Axonhügel wird entschieden, ob ein Schwellenpotential überschritten und ein Aktionspotential gebildet wird [wik]. [Jar] [HJ11]

Bei Künstlichen Neuronalen Netzen gibt es eine Aktivierungsfunktion, über die der Schwellwert festgelegt ist. Im Lauf der Zeit wurden verschiedene Aktivierungsfunktionen definiert und entwickelt. Die heute wohl typischste und meist zum Einsatz kommende ist ReLu (Rectified Linear Unit) [GBC16]:

$$(2.2) \quad \text{ReLu}(x) := (x)_+ := \max(x, 0) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases}$$

2.2.3.1 Künstliche Neuronale Netze

Künstliche Neuronale Netze (Artificial Neural Networks) bestehen aus mehreren Schichten (layers) von Neuronen, die über Gewichte (weights) miteinander verbunden sind (siehe Abbildung 2.3). Um das Netz zu trainieren, werden die Gewichte so angepasst, dass die Abweichung zwischen dem Erwartungswert und dem Ergebnis minimal wird. Die Optimierung der Kostenfunktion wird schichtweise über das Verfahren der Backpropagation vorgenommen.

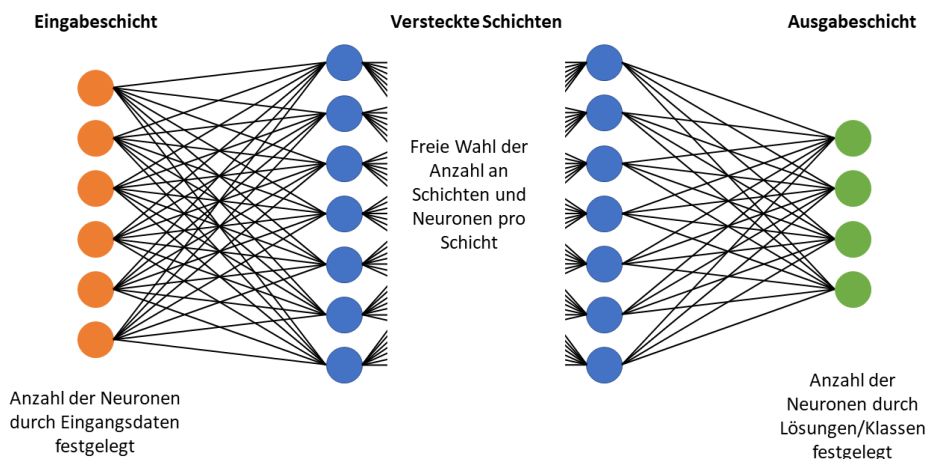


Abb. 2.3: Aufbau eines Künstlichen Neuronalen Netzes mit Ein- und Ausgabeschicht und n versteckten Schichten. Es besteht eine Verbindung jedes Neurons der Vorgängerschicht mit allen Neuronen der Folgeschicht. [Hen23]

Als Optimierungsverfahren kommen meist der Stochastic Gradient Descent (SGD), eine Weiterentwicklung des Gradientenverfahrens, oder Adam zum Einsatz. Adam ist ein Optimierer, welcher eine Kombination von SGD und RMSprop ist. RMSprop ist eine Variante von Adagrad, welche zur Anpassung der Lernrate angewandt wird. Adagrad steht für adaptiver Gradientenalgorithmus und ist ein "Algorithmus zur gradientenbasierten Optimierung" [dat].

2.2.3.2 Convolutional Neural Networks

Für Bildverarbeitungs- und Objektdetektions-Aufgaben werden Convolutional Neural Networks (CNN) eingesetzt. Bei CNN werden die Gewichtsmatrizen (Filter) und der Input über eine Faltung verknüpft (siehe Abbildung 2.4 und Abbildung 2.5). Durch die Faltung wird die Nachbarschaft der Pixel in einem Bild berücksichtigt. Das Filter hat somit ein rezeptives Feld. Durch die spezielle Struktur der Gewichtsmatrizen, die durch die Faltung erhalten wird, werden zudem nicht alle Neuronen des vorhergehenden Layers mit allen Neuronen des nachfolgenden Layers verknüpft. Durch das Weight Sharing (siehe Abbildung 2.6) ergeben sich weniger Verbindungen und somit weniger Parameter, die trainiert werden müssen. Würde ein Fully-Connected Neural Network eingesetzt werden, würde die Nachbarschaft der Pixel nicht berücksichtigt werden und durch die dort vorliegenden Dense Layers (jedes Neuron des vorherigen Layers ist mit jedem Neuron des darauffolgenden Layers verbunden) würde sich eine zu große Anzahl an zu trainierenden Parametern ergeben.

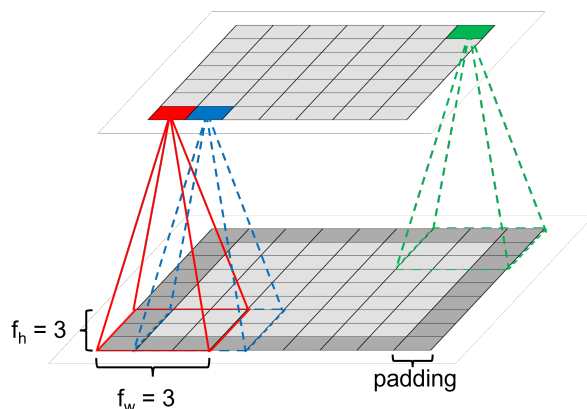


Abb. 2.4: Faltung mit einem 3 x 3 Filter (Schrittweite eins). In der Grafik ist die untere Schicht dargestellt, darüber die nächste Schicht n+1. Das rezeptive Feld ist bunt markiert. Mit Padding ist das Verhalten der Randbereiche gemeint. Meist wird Zero-Padding verwendet. [Hen23]

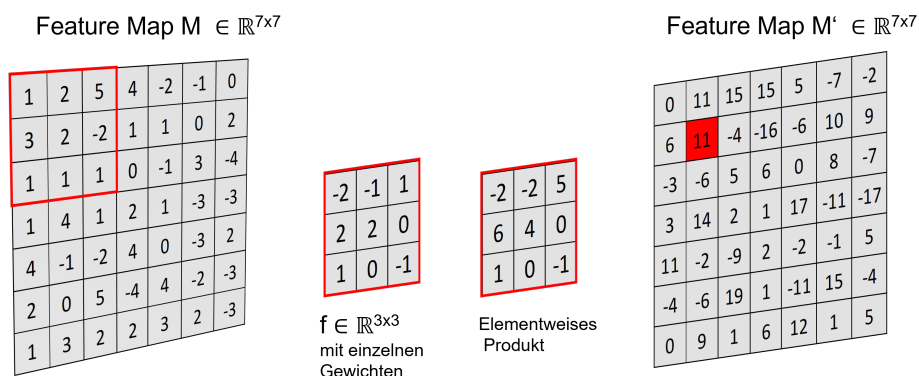


Abb. 2.5: Faltung einer 7 x 7 Feature Map M mit einem 3 x 3 Filter und Schrittweite eins. Die Feature Map M' zeigt das Ergebnis der Faltung. [Hen23]

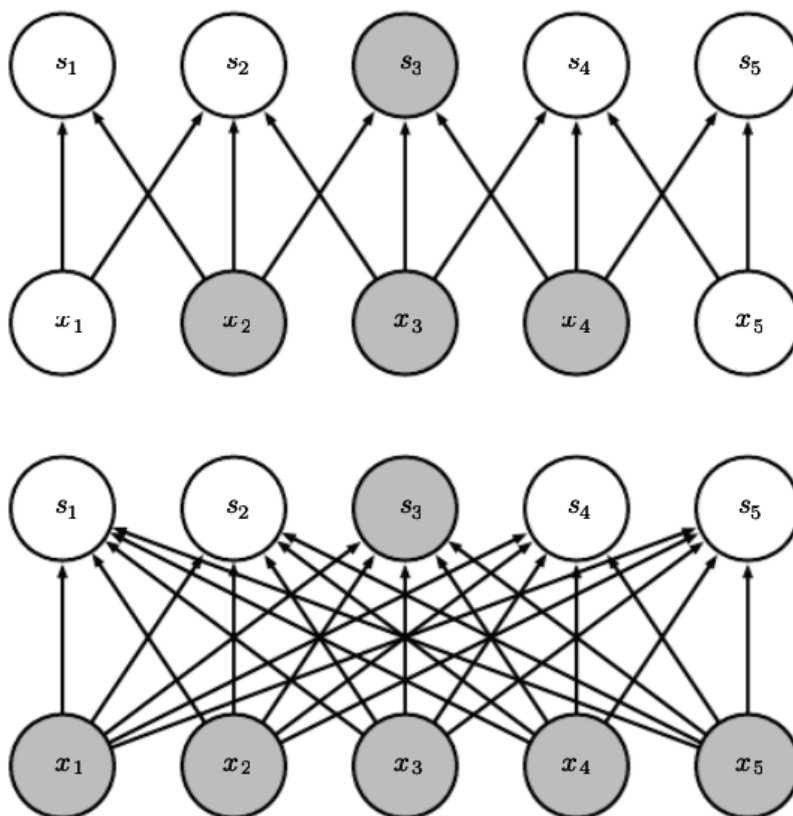


Abb. 2.6: Durch das weight sharing (siehe oben) müssen weniger Parameter trainiert werden. Unten: alle Neuronen sind miteinander über die Gewichte verbunden. [GBC16]

Convolutional Layer sind der Hauptbestandteil von CNN. Außerdem können Pooling Layer zum Einsatz kommen. Diese erlauben die Erfassung mit gleich großen Filtern bei unterschiedlich großen Strukturen.

Pooling Beim Pooling findet eine Datenreduktion statt, die Auflösung wird vermindert. Beim Max-Pooling werden die Schichten in 2×2 große Bereiche aufgeteilt und durch den maximalen Wert ersetzt. Die Grafik in Abbildung 2.7 erläutert das Verfahren.

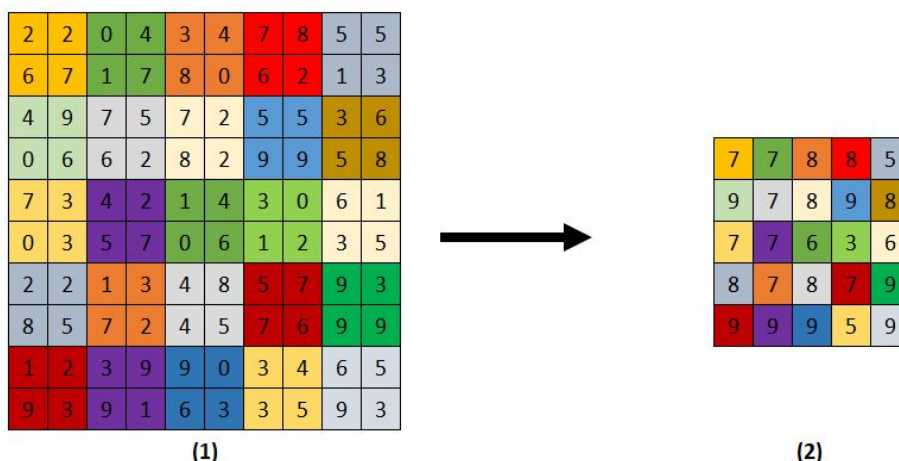


Abb. 2.7: Beispiel für Max-Pooling angewandt auf einen 10 x 10 Layer (li.) mit Aufteilung in 2 x 2 große Bereiche und die Schicht nach dem Pooling (re.) [Hen23]

Batch Normalization Layer kommen bei tiefen Künstlichen Neuronalen Netzen zum Einsatz, um die Leistung, Stabilität und Optimierung zu verbessern. Dabei werden die Aktivierungen der einzelnen Schichten normalisiert mit Hilfe von Anpassung und Skalierung.

Data augmentation Allgemein gesprochen ist bei kleinen Datensätzen Overfitting vorprogrammiert. Overfitting wird auch als Überanpassung beschrieben und bedeutet, dass ein Modell auf spezielle Daten angepasst ist. Data Augmentation (Transformation der vorhandenen Daten) und das Verwenden eines Dropout-Layers verbessern das Ergebnis.

An dieser Stelle sollen drei tiefe CNN vorgestellt werden, die in dieser Arbeit verwendet werden.

2.2.3.3 Residual Neural Networks

Hinter Residual Neural Networks (ResNet) steckt die Idee, dass es einfacher ist, ein Residuum zu approximieren, als die zu approximierende Funktion selbst [He+16a]. Aus der Literatur ist bekannt, dass das ResNet mit mehr Schichten effizienter arbeitet als mit weniger Schichten. Bei tiefen Neuronalen Netzen entwickelt sich unter anderem das Training bei mehr Schichten schlechter aufgrund des Problems des verschwindenden Gradienten (vanishing bzw. exploding gradient Problems) [BSF94] [GB10]. Um diesem Problem entgegenzuwirken, wurden Shortcut Verbindungen eingeführt [He+16a]. Abbildung 2.8 zeigt verschiedene Residual Blöcke. Das full pre-activation identity mapping (Abbildung 2.8 rechts) hat sich durchgesetzt [He+16b].

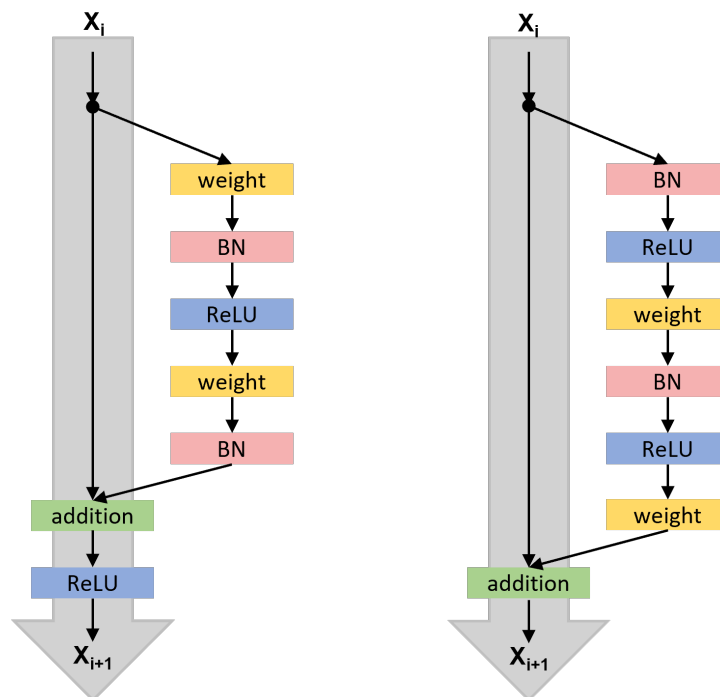


Abb. 2.8: Verschiedene Residual Blöcke. Links der original Residual Block von He. et al. [He+16a] und rechts das full pre-activation identity mapping, welches sich durchgesetzt hat [He+16b]. BN steht hierbei für Batch Normalization.

2.2.3.4 RetinaNet

Ein bekanntes CNN ist das RetinaNet basierend auf der Publikation von Lin et al. [Lin+17b], welches ein Feature Pyramid Net (FPN) Backbone (b) verwendet, das auf einer ResNet Architektur (a) aufsetzt [He+16a]. Das FPN verwendet Features vom letzten ResNet Layer. Danach folgen zwei Subnetze: eines für die Convolutional Objektklassifizierung der Anker-Boxen (c) und eines für die Convolutional Bounding Box Regression der Anker zu den Objekt-Boxen (d). Das RetinaNet verwendet Stochastic Gradient Descent (SGD) als Optimierer und horizontales Flipping der Bilder zur Data Augmentation. Die Struktur des Netzes ist in Abbildung 2.9 zu sehen.

Es ist bekannt, dass in einem Single-Stage-Detector eine starke Unausgewogenheit zwischen Vordergrund- und Hintergrund-Klassen vorliegt [Lin+17b]. Um diesem Problem zu begegnen, verwendet RetinaNet eine spezielle Kostenfunktion des Outputs des Klassifizierungs-Netzwerkes, den Focal Loss. Durch diese Funktion wird weniger loss durch "leichte" negative Fälle (wie der Hintergrund) produziert, sodass sich die Kosten auf die "harten" Fälle, also echte Objekte, die detektiert werden sollen, fokussieren. Dadurch wird die Vorhersage-Genauigkeit erhöht.

In oben genannter Publikation zeigten die Autoren, dass ihr bestes Modell, das auf einem ResNet-101-FPN Backbone basiert, mit einer Average Precision (AP) von 39,1 (bei 5 frames per second (fps)) eine hohe Genauigkeit erzielt, wenn mit dem COCO Datensatz getestet wurde [Lin+14]. Im Vergleich zu anderen Netzen ist das RetinaNet präzise und effizient. In Abbildung 2.10 ist ein Residual Block vom ResNet-50 Backbone, welches in dieser Arbeit verwendet wird, dargestellt.

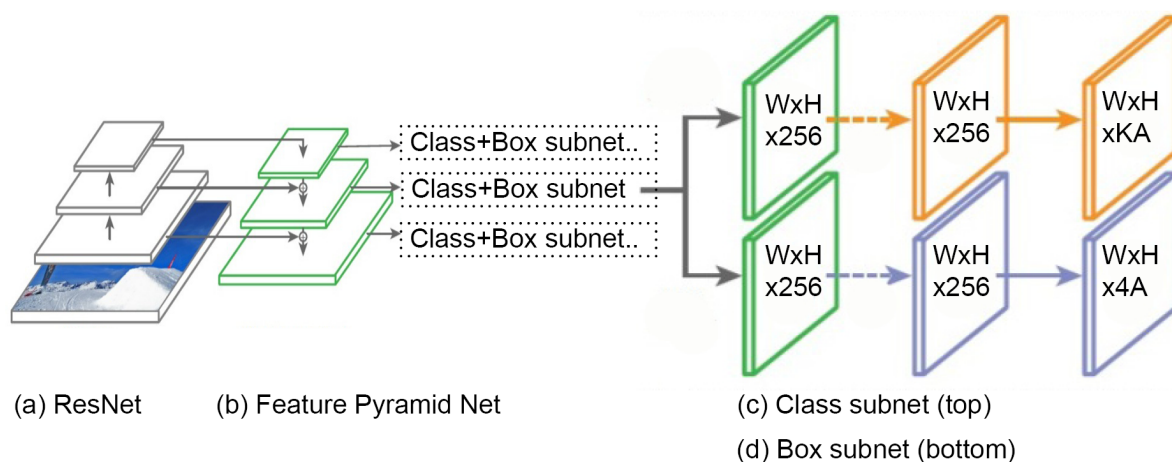


Abb. 2.9: Struktur des RetinaNet [Lin+17a]. Der FPN (b) Backbone setzt auf einer ResNet Architektur (a) auf.

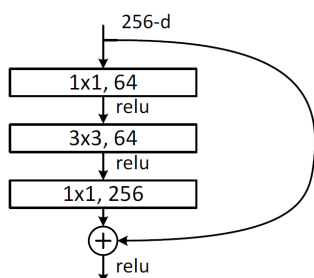


Abb. 2.10: In der Grafik ist ein Residual Block für das ResNet-50 Backbone zu erkennen. In diesem Fall geht die Shortcut Verbindung über drei Layer mit zwei 1×1 Convolutionen zur Reduktion und Erhöhung der Dimensionen und einer 3×3 Convolution. [Lin+17a].

2.2.3.5 Neuronales Netz YOLOv3

YOLOv3 (You only look once) (siehe Abbildung 2.11) verwendet das Darknet 53 als Backbone, welches aus Residual Blocks besteht [RF18]. Es liegen an drei Stellen im Netz Detektionsschichten vor, die mit unterschiedlichen Skalierungen des Bildes arbeiten. Der 82. Layer nimmt die erste Detektion vor. Während der vorherigen 81 Layer wird das Input Bild um einen Faktor 32 runterskaliert. Nach ein paar Convolutional Layers wird der 79. mit dem 61. Layer, der um einen Faktor 2 hochskaliert wurde, verkettet (concatenation). Es folgen weitere Convolutional Layers, bis eine weitere Detektion in Layer 94 stattfindet. Die Vorhersage wird auf einer Feature Map gemacht, die um einen Faktor 16 kleiner ist als das ursprüngliche Bild. Die letzte Detektion findet in Layer 106 statt, mit einem Downsampling-Faktor von 8. Es erfolgt eine tiefe Faltung. Der 91. Layer wird um einen Faktor zwei hochskaliert und verwendet, um mit Layer 36 zu falten. [Gle+23] Letztendlich wird für das Objekt die beste Vorhersage ausgewählt. Zwischen den Detektionsschichten sind Querverbindungen implementiert, um Feature Maps unterschiedlicher Auflösungen zu vereinigen. Als Aktivierungsfunktion wird Leaky Relu verwendet. Außerdem kommt Batch Normalization zum Einsatz. Anstelle von Pooling werden Convolution Layer mit angepassten Strides zum Down- und Upsampling verwendet. Dadurch

wird ein Datenverlust, der bei Pooling aufgetreten wäre, vermieden. YOLOv3 ist eines der schnellsten Objektdetektions-Algorithmen und ist eine gute Wahl, wenn eine Echtzeit-Detektion benötigt wird, bei vernachlässigbarem Verlust der Genauigkeit [Gle+23].

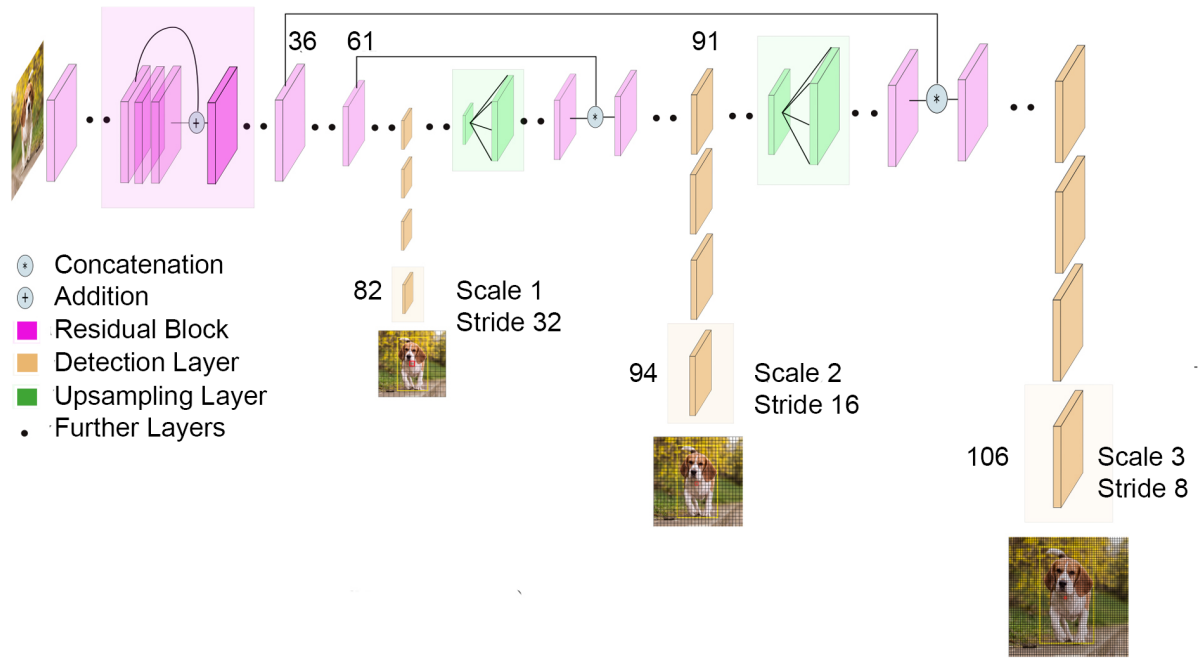


Abb. 2.11: Aufbau von YOLOv3 [Kat]

3 Stand der Technik

Zunächst soll der Stand der Technik der bisherigen Überwachung von Frühgeborenen auf der Intensivstation geschildert werden. Anschließend wird der Stand der Forschung des kontaktlosen Monitorings durch Einsatz von Kamerasensorik und Radartechnik auf der neonatologischen Intensivstation, sowie die Verfahren der Bildverarbeitung und Frequenzanalyse, die der Datenauswertung dienen, vorgestellt.

3.1 Überwachung von Frühgeborenen

Auf der neonatologischen Intensivstation werden Neugeborene bis zum Alter von 33 Wochen betreut. Sobald das Kind seine Körpertemperatur ohne Unterstützung des Inkubators halten kann und es ein Mindestgewicht von 1800 g besitzt, kann das Kind auf die "Intermediate Care Station" verlegt werden. Neugeborene, die vor der 30. SSW auf die Welt kommen, sind auf einen Inkubator angewiesen. Die neonatologische Intensivstation macht es sich zur Aufgabe, das Neugeborene am Leben zu erhalten, zu stabilisieren und für den Zeitraum des Krankenhausaufenthalts ihm und seinen Eltern ein Zuhause zu geben. Um dies zu ermöglichen, ist ein kontinuierliches Monitoring von Vitalfunktionen Voraussetzung, bei dem folgende Verfahren eingesetzt werden:

3.1.1 Überwachung der Atmung

Bei der Atmungsüberwachung interessieren vor allem die Atemrate, die Sauerstoffsättigung (pO_2), der CO_2 -Gehalt (pCO_2), sowie der Druck und das Volumen. Zur Messung des Sauerstoff- und CO_2 -Gehalts werden heutzutage weitgehend nichtinvasive Sensoren eingesetzt, welche die Menge der durch die Haut diffundierenden Gase erfassen [EBB05].

3.1.1.1 Sauerstoffsättigungs-Messung

Im Neonatologiebereich wird die Sauerstoffsättigung meist mit SpO_2 -Messgeräten über einen Clip bestimmt, der Fußgelenk oder Hand umfasst. Die Pulsoxymetrie basiert auf der Transmission (Durchdringung) oder der Reflektion von rotem und infrarotem Licht bei verschiedenen Wellenlängen. Sauerstoffarmes Hämoglobin (Hb) weist eine stärkere Absorption des roten Lichts auf (Wellenlänge 660 nm) als sauerstoffgesättigtes Hämoglobin HbO_2 , wohingegen bei infrarotem Licht das Gegenteil auftritt (siehe Abbildung 3.1).

Diese Eigenschaft wird zur Bestimmung der Sättigung genutzt. Die prozentuale Sauerstoffsättigung kann durch folgende Umrechnung ermittelt werden:

$$(3.1) \quad SpO_2 = 100 \cdot \frac{HbO_2}{(Hb + HbO_2)}$$

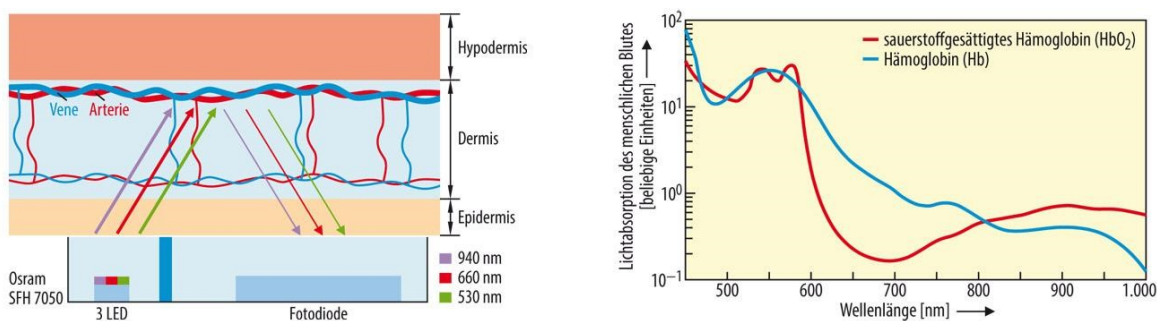


Abb. 3.1: Grundprinzip des reflektorischen, optischen Pulsmessens (li.) und Absorptionsspektrum (re.) [HR14]

Bei einem gesunden Menschen wird eine arterielle O_2 -Sättigung von ca. 94 % - 98 % erwartet. Wenn ein Wert von 92 % unterschritten wird, wird von einer Hypoxämie gesprochen.

3.1.1.2 Transknapode

Auf der neonatologischen Intensivstation wird außerdem eine Transknapode zur Überwachung des CO_2 -Gehalts verwendet. Hierbei wird über die Haut der Partialdruck der im Blut gelösten Kohlensäure (pCO_2) gemessen [Rag95]. Der eingesetzte Sensor ist ein Iridium-ph-Sensor, der transkutan platziert wird [HU09].

3.1.1.3 Impedanz-Pneumographie

Die Impedanz-Pneumographie (IPG) ist eine bekannte Methode, um die Atemrate des Patienten zu bestimmen. Hierfür werden entweder zwei oder vier Elektroden auf dem Brustkörper des Patienten angebracht (siehe Abbildung 3.2) und die Änderung der elektrischen Impedanz, verursacht durch die Atembewegung, gemessen. Durch einen Highfrequency Alternating Current (AC)-Strom wird eine Potentialdifferenz bewirkt, die in Zusammenhang mit dem Widerstand des Gewebes steht. Meist wird die Methode mit zwei Elektroden angewandt, bei der normale Elektrokardiogramm (EKG)-Elektroden eingesetzt werden können. [Gup11]

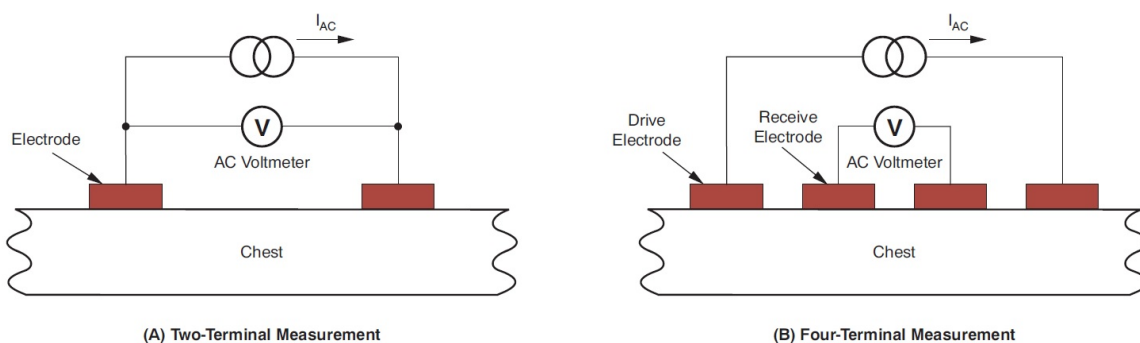


Abb. 3.2: Impedanz-Pneumographie [Gup11]

3.1.1.4 Beatmungsgeräte

Wenn das Kind künstlich beatmet wird, kann das Beatmungsgerät selbst zur Überwachung dienen, da entweder der Druck oder das Volumen geregelt wird. Damit durch die Beatmung der unterentwickelten Lunge des Frühgeborenen keine weiteren Verletzungen zugefügt werden, wird meist der Druck geregelt. Neben den Parametern Volumen und Druck wird auch das Atemminutenvolumen (AMV) bestimmt. Es gibt verschiedene Arten der Beatmung: entweder über einen Tubus durch die Nase oder eine Atemmaske, die an einer Mütze am Kopf des Kindes befestigt wird, um Druckstellen zu vermeiden.

3.1.2 Überwachung des Herzens

Das Standardmessverfahren zur Überwachung des Herzens und zur Erkennung von Fehlfunktionen ist das Elektrokardiogramm (EKG). Der Stromfluss, der durch die Ausbreitung der elektrischen Erregung des Herzens verursacht wird, kann durch das EKG an der Körperoberfläche gemessen werden [HJ11]. EKG-Elektroden bestehen meist aus Polymeren und Elastomeren, die durch Ag-Cl-Gel leitfähig werden [EBB05]. Üblicherweise wird zur kontinuierlichen Überwachung von Frühgeborenen ein 3-Punkt-EKG verwendet.

3.1.3 Blutdruck

Zur Blutdruckmessung wird überwiegend die Riva-Rocci-Methode angewandt. Der Arzt oder Pflegemitarbeiter legt ein Stethoskop in die Ellenbeuge und pumpt die Blutdruckmanschette auf. Während die Luft langsam abgelassen wird, können verschiedene Strömungsgeräusche wahrgenommen werden, die den systolischen und diastolischen Blutdruck abbilden. [HJ11] Elektronische Geräte wenden intern dasselbe Verfahren an. Auf der neonatologischen Intensivstation kann die Riva-Rocci-Methode jedoch nicht eingesetzt werden. Das regelmäßige Aufblasen der Manschette würde das Kind unter Stress setzen. Deswegen wird beim Langzeitmonitoring der mittlere arterielle Druck invasiv gemessen. Hierfür wird nur einmal ein Zugang für einen Katheter gelegt. Bei Frühgeborenen mit einem Gewicht zwischen 1000 g und 2000 g liegt der Normwert des Blutdrucks bei 45 - 50 mmHg (systolisch) [Bit10].

3.1.4 Temperaturmessung

Der am häufigsten kontrollierte Vitalparameter ist die Körpertemperatur. Ein gesunder Mensch weist eine Körpertemperatur von 37 °C auf mit einer Toleranz von +/-0,5 °C. Wenn die Körpertemperatur ansteigt, kann dies ein Anzeichen für Krankheiten oder Infektionen sein. Ein Abfall kann auf einen Schock hindeuten. Häufig wird die Körpertemperatur in Körperhöhlen wie dem Rektum oder der Achselhöhle gemessen, um eine hohe Genauigkeit zu erreichen. Alternativ kann die Temperatur auch im Ohr bestimmt werden. [EBB05] Auf der Frühgeborenenstation wird die Temperatur entweder rektal oder interarteriell gemessen, da eine kontinuierliche Überwachung erwünscht ist. Druckstellen und Ekzeme sind wegen des ständigen Kontaktes nicht komplett vermeidbar.

Zur Überwachung der Körpertemperatur kommen unter anderem Thermosonden im Bereich des Abdomen oder der Fußsohle zum Einsatz. Diese gibt es zum Einmalgebrauch mit Klebepolstern. Hierbei muss beachtet werden, dass im Bereich des Abdomen die Messung durch Wärmestrahler

oder Phototherapie beeinflusst werden kann. Im Allgemeinen kommt es aber nicht nur auf die genauen Messwerte, sondern auf die Tendenz des Temperaturverlaufs an. [Spa09] Anhand dieser Werte können verschiedenste Krankheitsbilder abgeleitet werden.

3.1.5 Einflüsse auf das derzeit übliche Monitoring

An dieser Stelle sollen mögliche Einflüsse auf das derzeit übliche Monitoring und die daraus resultierenden Effekte auf die Messungen beschrieben werden.

3.1.5.1 Leitfähigkeit der Elektroden

Damit beim EKG und der Impedanz-Pneumographie korrekte Messwerte geliefert werden, muss die Leitfähigkeit der EKG-Elektroden gewährleistet sein. Die Leitfähigkeit verringert sich, wenn das Kind schwitzt oder ein Hautödem hat.

3.1.5.2 Bewegung des Kindes

Die Bewegung des Kindes kann die Messungen des EKGs, der Impedanz-Pneumographie und der Temperatur beeinflussen. Durch die Bewegung ergeben sich andere Spannungs-Potenziale auf der Haut, was zu Fehlmessungen durch die EKG-Elektroden führt. Insbesondere bei der Impedanz-Pneumographie ist dies problematisch, da die Frequenzen der Bewegungen meist in einem ähnlichen Frequenzbereich wie die Atemrate selbst liegen. Auch die rektale Temperatursonde kann sich verschieben oder sogar komplett entfernt werden, wodurch falsche oder gar keine Messwerte akquiriert werden.

3.2 Kontaktlose Überwachung von Vitalparametern

Es gibt mehrere Forschungsansätze, die vorschlagen, durch verschiedene Kamera-Messverfahren die Vitalparameter Herzrate, Atemrate und Temperatur kontaktlos zu bestimmen. Auch auf den Einsatz von Radar zur Vitalparameterdetektion soll eingegangen werden.

3.2.1 Atemüberwachung mittels visueller Sensorik

Die folgenden Ansätze basieren auf der Aufnahme von Werten mittels optischen Verfahren.

3.2.1.1 Atemüberwachung mittels Infrarotkamera

Abbas et al. [Abb+11] stellen eine kontaktlose Atmungsüberwachung für Frühgeborene auf der Intensivstation vor, die auf einer gepulsten Thermographie basiert. Um die Atemrate der Spontanatmung zu bestimmen, werden Temperaturprofile der Nasenlöcher der Kinder analysiert. Beim Ein- und Ausatmen können signifikante Temperaturunterschiede erkannt werden. Referenzmessungen wurden aber nicht durchgeführt. [Abb+11]

Murthy et al. stellen ein auf einer Thermokamera basierendes Verfahren zur Überwachung des Atemflusses von erwachsenen Apnoepatienten im Schlaflabor vor. Der Bereich um die

Nase herum wird als Region of Interest (ROI) definiert, überwacht und der Temperaturverlauf aufgezeichnet. [Mur+09]

Pereira et al. nutzen eine Thermokamera, um den interessanten Bereich um die Nase zu bestimmen, zu verfolgen und dann die Atemfrequenz durch Signalverarbeitung zu berechnen. Die Region of Interest wird manuell festgelegt. In einer klinischen Studie mit vier Frühgeborenen zeigte sich, dass nur ein Fehler von 3,24 %, auftritt, im Vergleich zur Atemrate, bestimmt über das EKG (Impedanz-Pneumographie). Der Nachteil dieser Methode ist, dass die ROI manuell festgelegt, und nicht automatisch bestimmt wird. [Per+17]

Auch die Kombination von photoplethysmographischer Bildgebung (PPGI) mit einer Infrarotkamera soll zur Atmungsdetektion einsetzbar sein. PPGI kann genutzt werden, um die Herzrate, die Herzratenvariabilität, die Atemrate und die Atemvariabilität zu bestimmen. Die Infrarotkamera unterstützt in diesem Fall die Bestimmung der Atemrate und der Atemvariabilität in der Nasenregion. Die Untersuchungen wurden nur an Erwachsenen und in Tierversuchen durchgeführt. [Bla+14]

In [BGK15] wird von Bennett et al. gezeigt, dass mit Hilfe des Eulerian Video Magnification (EVM) Algorithmus von Wu et al. (siehe Unterunterabschnitt 3.2.2.1) die Atemfrequenz im Nasenbereich von Erwachsenen mit Hilfe einer Thermokamera bestimmt werden kann. In den jeweils manuell maskierten Regions of Interest wird die Durchschnittsintensität berechnet und über die Zeit dargestellt.

3.2.1.2 Atemüberwachung mittels RGB-Kamera

Im Gegensatz zum oben Gesagten verwenden Davis et al. eine RGB-Kamera, um die Herzrate, Atemfrequenz und die Sauerstoffsättigung (SpO_2) zu bestimmen. Dabei wurde die Durchschnittsintensität der drei Kanäle der Region of Interest (ROI) bestimmt und eine Independent Component Analyse (ICA) angewandt, um eine PPG-Welle (Pulsplethysmographie) zu erhalten. Es konnten saubere Signale der Herzrate, Atemfrequenz und Sauerstoffsättigung generiert werden. Als Referenzmessungen wurden Impedanz-Pneumographie, PPG und ein EKG eingesetzt. Problematisch sind bisher noch Änderungen der Lichtverhältnisse, da die Vitalparameter dann nicht exakt bestimmt werden können. Weitere Einschränkungen sind, dass die ROI manuell festgelegt werden und Daten mit Bewegung des Kindes nicht ausgewertet werden. [Dav+14]

Hassan et al. bestimmen die Herz- und Atemrate anhand von RGB-Kamerabildern des Gesichts, allerdings nur bei Erwachsenen. Hierbei wird der Verlauf des grünen Spektrums in zwei Regions of Interest, dem Nasenbereich und dem Stirnbereich, bestimmt und eine Wavelet-basierte Rauschentfernung durchgeführt. [Has+17]

3.2.1.3 Detektion der Brustkorbbewegung

Um die Atemrate aus der Brustkorbbewegung zu ermitteln, müssen grundsätzlich zwei Aufgaben gelöst werden: Zum einen muss der Brustkorb als Region of Interest erkannt werden und zum anderen die Frequenz der Bewegung bestimmt werden. Folgende Ansätze wurden bereits publiziert:

Janssen et al. beschreiben eine Video-basierte Atmungsüberwachung mit automatischer Detektion der Region of Interest. [Jan+15] In der Publikation von Janssen et al. wird erwähnt, dass laut Bartula et al. die durch die Atmung verursachte Bewegung hauptsächlich in der vertikalen Richtung auftritt [BTM13].

Li et al. bestimmen die Brustkorbbewegung, indem zunächst mit einem Harris-Corner-Detector Features bestimmt werden und diese dann mit dem Lucas-Kanade-Algorithmus verfolgt werden [LYT14]. Die Region of Interest wird hier nicht bestimmt.

Lukac et al. berechnen mit dem Lucas-Kanade Algorithmus den Optischen Fluss und gewichten die Bewegungsvektoren mit dem Signal to Noise Ratio (SNR) [LPC14]. Die Gebiete des maximalen SNR sollen hierbei der relevanten Region of Interest im Brustkörperbereich entsprechen.

Jorge et al. verwenden Temporal Difference für die Brustkorbdetektion im RGB-Kamerabild für Frühgeborene, weil diese Methode nicht so zeitaufwendig wie der Optische Fluss ist. Temporal Difference bedeutet, dass durch Bildsubtraktion die Bildunterschiede in aufeinander folgenden Frames bestimmt werden. [Jor+17]

Die Methode wird in der Publikation von Bai et al. beschrieben [BLY10].

Al-Naji et al. detektieren die Brustkorbbewegung aus RGB-Bildern mittels Euler-Algorithmus und Bildsubtraktion. [AC16] Die Bilder in der Publikation zeigen einen Säugling mit einem stark gemusterten Body. Es ist zu vermuten, dass nur unter diesen Bedingungen das hier vorgestellte Verfahren gut funktioniert.

Einsatz von Structured-Light-Plethysmographie Lauhkonen et al., Ghezzi et al. und Hmeidi et al. zeigen, dass es möglich ist, mit Structured-Light-Plethysmographie die Brustkorbbewegung zu detektieren und die Atmung zu erkennen [LCI19] [Ghe+17] [Hme+18]. Dies wurde bisher an Kindern zwischen 2 und 12 [Hme+18] bzw. 15 Jahren [Ghe+17] gezeigt. Das Structured-Light-Plethysmographie Gerät projiziert ein Muster auf den Brustkorb und den Bauch des Patienten. Mit zwei Kameras wird die Verschiebung der Schnittpunkte des Musters beobachtet.

3.2.1.4 Einsatz von Tiefenkameras

Time of Flight-Kamera Es gibt einige Autoren, die eine Time of Flight-Kamera zur Bestimmung der Atemfrequenz einsetzen. Die meisten Verfahren wurden nur an Erwachsenen getestet wie z.B. das von Penne et al.. Hierzu werden verschiedene Ebenen bestimmt: Brustkorb-, Bauch- und Auflageebene. Es wird die Distanz zur Auflageebene über die Zeit berechnet. [Pen+09]

Martinez und Stiefelhagen beschreiben den Einsatz einer Prime Sense 1080 Tiefenkamera zur Überwachung der Atmung von Erwachsenen im Schlaflabor. Die ROI des Bettes wird automatisch detektiert, die ROI der Brust wird dynamisch bestimmt. Jedes Pixel in der Region of Interest produziert eine Trajektorie. Für jede Trajektorie wird die spektrale Leistungsdichte ermittelt und dessen Maximalwert bestimmt. Bei dieser Frequenz liegt die Atemfrequenz. [MS17]

Procházka et al. verwenden eine Kinect-Kamera, um die Atmung zu detektieren. Ähnlich wie bei Martinez und Stiefelhagen berechnen sie Trajektorien. [Pro+16] Bernacchia et al. verwenden auch eine Kinect, um die Herz- und Atemrate zu bestimmen. Hierbei wird die Region of Interest zunächst manuell bestimmt und es wurde nur an Erwachsenen getestet. [Ber+14] Auch bei Benetazzo kommt eine Kinect-Kamera nur bei Erwachsenen zum Einsatz [Ben+14]. Rihana et al. verwenden eine Kinect 2 zur Atemdetektion von Erwachsenen [Rih+16]. Nguyen et al. kombinieren eine Kinect mit einem Radar-Sensor und bezeichnen dies als Wikispiro, welches ebenfalls nur an Erwachsenen getestet wurde. [Ngu+16]

Auch Coronel et al. setzen eine ToF-Kamera ausschließlich zur Atemdetektion von Erwachsenen ein. [Cor+20]

Procházka et al. verwenden eine Thermo- und eine Kinect-Kamera. Im Thermokamerabild wird die Temperatur des Atemstroms des Erwachsenen analysiert. Die Kinect zeichnet die Brustkorbbewegung auf und dient als Vergleich. Es liegt keine direkte Fusion vor. [AC17]

Wijenayake und Park verwenden eine Asus Xtion PRO RGB-D Kamera und befestigen auf dem Patienten vier weiße Marker, um die Region of Interest zu markieren. [WP17]

Structured-Light-Kameras Es gibt zwei Ansätze, bei denen Structured-Light-Kameras eingesetzt werden, die bereits an Neugeborenen getestet wurden. Die erste Methode benötigt zwei Kameras und berechnet das Atemvolumen. In einem weiteren Schritt wird die Atemrate bestimmt. [Reh+17] [Reh+18] Der Einsatz von zwei Kameras ist allein für die Atemdetektion nicht nötig. Somit ist dieses Verfahren zu komplex.

Der zweite Ansatz wird bisher nur bei offenen Inkubatoren eingesetzt und verlangt eine manuelle Auswahl der Messpunkte durch den Nutzer. Somit wird der Brustkorb nicht automatisch detektiert. Dieser Ansatz wurde bisher nur an Frühgeborenen getestet und ist auf deren Atemfrequenz spezialisiert. [Cen+15]

3.2.2 Herzrate mittels visueller Sensorik

3.2.2.1 Eulerian Video Magnification

2012 wurde der Eulerian Video Magnification Algorithmus vorgestellt, der es ermöglicht, kleine, für das menschliche Auge unsichtbare Farbänderungen und Bewegungen zu verstärken. In der Publikation von Wu et al. wird beschrieben, wie in einem einfachen Farbvideo des Gesichts eines Erwachsenen durch die Verstärkung der Farbänderung der Herzschlag sichtbar gemacht werden kann. Auch kleine Bewegungen, wie z.B. die Brustkorbbewegung eines Neugeborenen, können dadurch verstärkt und erkennbar gemacht werden. Der Eulerian Video Magnification (EVM) Algorithmus wendet als erstes eine spatial decomposition (räumliche Zerlegung) auf das Input-Video an. Die daraus resultierende Laplace-Pyramide wird zeitlich (temporal processing) verarbeitet. Das Signal wird um einen Faktor α verstärkt. Ein passender Verstärkungsfaktor kann durch die folgende Formel ermittelt werden:

$$(3.2) \quad (1 + \alpha) \cdot \delta(t) < \lambda/8$$

mit α als Verstärkungsfaktor, δ als Verschiebefunktion und λ als Cutoff-Wellenlänge. Das Frequenzspektrum, das manipuliert werden soll, wird durch die untere und obere Cutoff-Frequenz beschrieben und hängt somit von dem Signal ab, das verstärkt werden soll. Diese Methode kann die tatsächliche Frequenz z.B. der Brustkorbbewegung nicht ermitteln, sondern nur Farbänderungen sichtbar machen. Zur Frequenzbestimmung wird ein weiterer Algorithmus benötigt. [Wu+12]

Chen et al. beschreiben die kontaktlose Herzratenbestimmung an fünf Neugeborenen basierend auf dem EVM-Algorithmus und einer FFT. Anders als bei Wu et al. wird im EVM anstelle der Laplace-Pyramide eine Gauss-Pyramide verwendet. Das Gesicht wird als Region of Interest im HSV-Farbbereich segmentiert. Nach Anwendung des EVM-Algorithmus wird der grüne Farbkanal zur Extraktion des Blutvolumenpulses genutzt. Ein Bandpass mit Cutoff-Frequenzen zwischen 1,8333 und 2,6667 Hz wird angewandt und die spektrale Leistungsdichte über die FFT bestimmt. Die Frequenz mit der höchsten Leistung wird als Herzrate definiert. Es wird mit dieser Methode ein durchschnittlicher RMSE von 15,1829 erreicht. [Che+20]

3.2.2.2 Herzrate mittels Infrarotkamera

Ähnlich wie bei der Atemrate (in [BGK15] beschrieben) kann auch die Herzrate von Erwachsenen aus Infrarotbildern unter Verwendung des Eulerian Video Magnification Algorithmus bestimmt werden. Hierbei wird ein Verstärkungsfaktor $\alpha = 100$ eingesetzt, um den Herzschlag eines Erwachsenen zu detektieren. [BGK16]

3.2.3 Radarsensorik zur Vitalparameterdetektion

Ein weiteres kontaktloses Messverfahren zur Detektion der Atemrate ist Radar. Es gibt verschiedene Arten von Radarsignalen: Continuous Wave (CW), Frequency Modulated Continuous Wave (FMCW) und Ultra-Wideband (UWB) [Wol]. CW sendet kontinuierlich ein Hochfrequenz-Signal. FMCW ist dahingehend ähnlich, dass das Signal kontinuierlich gesendet wird, aber die Frequenz wird periodisch moduliert in einem vordefinierten Frequenzbereich. Dies ermöglicht die Detektion von absoluten Distanzen. UWB sendet kurze Radarsignal-Pulse oder einen breitbandigen Sweep und empfängt währenddessen das reflektierte Signal. Eine besondere Variante des CW-Radar-Sensors ist der "Microwave Interferometric" Radar-Sensor, der relative Distanzunterschiede im Zeitbereich misst. Außerdem gibt es multiple-input multiple-output (MIMO) Radar-Sensoren, die meist FMCW-Radar mit mehreren Empfängern und Sendern verwenden, um eine höhere Auflösung im Raum über den Winkel zu erreichen (spatial resolution). [Con]

3.2.3.1 UWB-Radar

Es existieren mehrere Verfahren zum Messen der Atemrate mittels UWB, die an Kleinkindern und Säuglingen getestet wurden. Kim et al. haben ein Impuls-Radio Ultra-Wideband (IR-UWB)

Gerät mit einer Mittenfrequenz von 7,29 GHz und einer Bandbreite von 1,5 GHz zur Überwachung der Herz- und Atemrate von sechs Neonaten verwendet [Kim+19]. Es wurden keine Angaben zu einem Referenzsystem gemacht.

Schleicher et al. haben ihre Methode basierend auf IR-UWB mit einem Frequenzband zwischen 3,1 GHz und 10,6 GHz an männlichen Erwachsenen und an einem sieben Wochen alten Säugling getestet und zeigten, dass es möglich ist, den Atemzyklus aufzuzeichnen [Sch+13].

In einem Patent von Tupin wird ein UWB Überwachungsgerät für Neonaten zur Erkennung von Atemnot und Herz-Lungen-Beschwerden vorgestellt [Tup12]. Der Nachteil dieser Methode ist, dass Impedanzwandler-Pads verwendet werden müssen, um den Verlust von UWB-Energie zu verhindern. Der Säugling muss auf diese Pads gelegt werden.

Lee et al. verwenden IR-UWB Radar, um Herz- und Atemrate von 34 Neonaten zu messen. Als Referenz kamen EKG und Impedanz-Pneumographie zum Einsatz. Die mittlere Abweichung bei der Atemdetektion liegt bei 0,17 Atemzüge/Minute und bei -0,23 BPM bei der Herzraten-detektion. [Lee+21]

Park et al. können mit ihrem IR-UWB Radar die Atemrate, die Pulsrate und das R-R Intervall detektieren. Dies wurde an 19 erwachsenen Patienten mit normalem Herzschlag und 14 mit Vorhofflimmern durchgeführt. Es wurde eine gute Übereinstimmung zwischen der mit dem IR-UWB detektierten Atemrate und der manuell vom Arzt gezählten festgestellt (intraclass correlation coefficient [ICC] 0,852). Bei normalem Sinusrhythmus beträgt der ICC der Pulsrate zum EKG 0,985. [Par+19]

3.2.3.2 FMCW-Radar

Es wurden einige Tests mit FMCW-Radar zur Vitalparameter-Detektion von Erwachsenen durchgeführt. Wang et al. haben einen FMCW-Radar-Sensor mit einer Mittenfrequenz von 80 GHz und einer Bandbreite von 10 GHz mit Sägezahn-Modulation verwendet für Slow Chirps (was nur die Messung von Distanzen ermöglicht) [Wan+15]. Die besten Ergebnisse konnten bei einer Lage des Sensors in 1 m Abstand frontal zum Patienten erreicht werden. Der Fehler bei der Detektion der Atemrate beträgt dann 6,89 %, bei der Detektion der Herzrate 8,09 %.

In der Veröffentlichung von Mathews et al. wird ein weiteres kontaktloses Überwachungssystem vorgestellt, das auf einem 35-GHz-FMCW-Radar-Sensor basiert [MSB00]. Die vorgeschlagenen Einsatzgebiete sind das Militär, neonatologische Intensivstationen, Patienten mit traumatischen Verletzungen und Brandwunden. Getestet wurde es allerdings noch nicht an Neonaten. Die Schwierigkeit besteht im Entfernen von Bewegungsartefakten, d.h. der Patient sollte sich ruhig verhalten. Dies kann bei Frühgeborenen nicht sichergestellt werden.

Marnach et al. haben einen 24-GHz-FMCW-Radar-Sensor verwendet, um Messungen an einem Simulator der Atmung eines Säuglings aufzuzeichnen. Als Simulator wurde ein Reflektor auf einem Lautsprecher platziert [MSD19] [SMD19].

Antonilos et al. messen neben der Herz- und Atemrate die Herzratenvariabilität mit einem 122 GHz FMCW-Radar [Ant+20b].

Rudrappa et al. verwenden einen MIMO FMCW-Radar und konnten damit die Herz- und Atemrate von sich bewegenden Erwachsenen detektieren. Als Referenzmessung wurde ein Pulsoxymeter und eine Fitnessuhr eingesetzt. Der mittlere Fehler lag unter 1 %. [RHK20]

3.2.3.3 CW-Radar

Unmodulierter CW-Radar kann ebenfalls zur Detektion von Vitalparametern eingesetzt werden. Li et al. haben einen Doppler-Radar-Sensor entwickelt mit einer Frequenz von 5,8 GHz, der auch für das Home-Monitoring von Neugeborenen und Kindern genutzt werden kann, um plötzlichen Kindstod zu verhindern [Li+09]. Die Anwendung an Frühgeborenen ist nicht explizit genannt. Dieses System wird in einem Monitoring-Prototyp für Säuglinge verwendet. Der Radar-Sensor wird an der Seite des Baby-Bettes positioniert. Das bestehende Problem des Systems ist, dass Bewegungen des Kindes zu ungenauen oder falschen Messungen der Vitalparameter führen, weil Atmung und Herzschlag von Bewegungsartefakten überlagert werden. Dementsprechend besteht Verbesserungspotential. In einem weiteren Schritt wurde das System an einem Säuglings-Simulator verifiziert (high fidelity infant simulator METI, 3–6 Monate) [Yan+09]. In diesem Fall wurde der Radar-Sensor unterhalb des Neugeborenen angebracht. Unmodulierter CW-Radar ist sensibel gegenüber starkem Hintergrund-Clutter, verursacht durch Laufen oder Sprechen von Personen in der Nähe der Messumgebung. Auch diese Bewegungen können die Vitalparameter überlagern und die Detektion erschweren.

Hu et al. haben mit einem CW-Radar gearbeitet zur Überwachung der Herz- und Atemraten von Erwachsenen. Es wurde ein relativer Fehler von 2,53 – 4,83 % für die Interbeat-Intervalle verglichen mit einem EKG erreicht [Hu+14].

Tworzydło hat für die kontaktlose Überwachung von Gefängnis-Insassen ein 24-GHz-CW-Doppler-Radar eingesetzt. Mit einem mittleren Fehler von 6 BPM konnte detektiert werden, ob die Person normal atmet. Wenn der Atem angehalten wurde oder sich bewegende Personen in der Nähe waren, konnte die Atemrate nicht korrekt erkannt werden [Two16].

Rahman, T. et al. haben ein System zum Tracking der Schlaf-Qualität entwickelt, das auf einem 24-GHz-CW-Radar-Sensor basiert und die Herz- und Atemrate, sowie Körperbewegungen tracken soll. Es wurde ein mittlerer absoluter Fehler von 1,98 BPM für die Herzrate und 3,29 BPM für die Atemrate erreicht. [Rah+15]

Droitcour et al. setzen einen 2,4-GHz-CW-Radar zur Messung der Atemrate an 24 Probanden im Alter zwischen 43 und 91 ein. Ihr Ergebnis zeigt, dass ihr System als Alternative zu den jetzigen (Referenz-) Monitoring-Geräten genutzt werden könnte. [Dro06]

Beltrao et al. verwenden zur Atemdetektion einen 24-GHz-CW-Radar-Sensor, der ein Prototyp des IEE's VitaSense® Sensor ist. Getestet wurde an 12 Frühgeborenen mit Impedanz-Pneumographie als Referenzmessart. Die besten Ergebnisse wurden erhalten, wenn das Kind

sich in der Bauchlage befand. Der Gesamt-Mittlere quadratische Fehler (RMSE) ist kleiner 7 BPM, in Bauchlage sogar nur 5 BPM. [Bel+22]

In einer Vergleichsstudie zwischen UWB und CW-Radar zur Atemüberwachung zeigte sich, dass ein CW-Radarmodul von der Sapienza-Universität di Roma Bewegungen der Größe 1 mm aus einer Distanz von 2 m feststellen konnte. Das verwendete FMCW-Radar (IVS-148 von Firma InnoSent) konnte Bewegungen der Amplitude 2 mm aus einer Distanz von 3 m detektieren. Im Vergleich konnte das UWB Modul NVA-R640, welches die Universität Oslo verwendet, nur Bewegungen von 10 mm aus einer Distanz von 3 m erkennen [Pis+14]. Dies zeigt, dass CW-Radar kleinere Bewegungen aus größeren Distanzen als UWB-Radar detektieren kann. Ein Sechs-Tor-CW-Radar-Sensor ist sogar in der Lage, Bewegung im Mikrometer Bereich zu erkennen [Lin+13].

Das BMBF-Projekt Guardian (Förderkennzeichen 16SV7694), welches Six-Port-Radarsensorik zur Überwachung der Vitalparameter in der Palliativmedizin einsetzt, konnte mit dieser Technik schon gute Ergebnisse erzielen. Der 24-GHz-Six-Port-Radarsensor lieferte im Vergleich zum herkömmlichen EKG einen quadratischen Fehler (RMSE) von 18 ms [Wil+17b]. Als weitere Referenzmessgeräte wurden ein Pulsplethysmographie-Gerät (PPG) und ein Spirometer für die Atmungsdetektion verwendet. Template Matching und Kreuz-Korrelation kamen zum Einsatz. Hierfür wurden fünf verschiedene Templates generiert, welche typische Kurven im EKG widerspiegeln und diese per Template Matching abgleichen. [Wil+17b] In einer Folgearbeit wurde gezeigt, dass sogar die lokalen Herztöne gemessen werden können. Es wurde ein Segmentierungsalgorithmus basierend auf einem Hidden Semi-Markov Modell entwickelt und gezeigt, dass im Vergleich zum EKG ein RMSE von 44,2 ms erzielt werden kann. Mit einem Advanced Template Matching wurde nur ein RMSE von 144,9 ms erreicht. [Wil+18]

Schellenberger et al. haben später gezeigt, dass die Atmung mit den Amplituden der ersten und zweiten Herztöne korreliert und dieser Zusammenhang zur Bestimmung der Atemrate ohne durch langsame Körperbewegungen verursachtes Rauschen verwendet werden kann. [Sch+19]

Schellenberger und Shi et al. haben zwei Radar-Datensätze von Vitalparametern veröffentlicht. Es wurden jeweils die Interbeat Intervals (IBI) mit dem Referenz-EKG verglichen. Beim ersten Datensatz waren 68,9 % der IBI Werte innerhalb von 10 ms der Referenz. Bei 90 % der Daten lag die Abweichung innerhalb von 45 ms. [Sch+20] Beim zweiten Datensatz lagen 63,4 % innerhalb der 10 ms und 96 % innerhalb der 50 ms [Shi+20b]. Weitere Möglichkeiten zur Klassifizierung der Herztöne wurden untersucht und beispielsweise Long Short-Term Memory (LSTM) Netzwerke kamen zum Einsatz [Shi+20a] [Sa21].

Vinci et al. haben Six-Port-Radar-Sensoren bereits zur Überwachung von Vitalparametern von Erwachsenen zum Beispiel im Automotive-Sektor eingesetzt. Sie wurden auch im klinischen Umfeld getestet und lieferten gute Ergebnisse. [Vin+15] Der Ansatz entspricht "Microwave Interferometric" Radar-Sensoren. Ein großer Vorteil ist bei ihnen, dass die Distanzänderungen mit Mikrometer-Genauigkeit durch Phasenevaluation gemessen werden können. Aus diesem Grund sollte ein Radar-Sensor basierend auf "Microwave Interferometric" Radar verwendet werden.

Ein Argument dafür, CW-Radar-basierte Sensoren für ein Überwachungssystem zu verwenden, ist, dass sie den Einsatz des Industrial, Scientific and Medical (ISM) Frequenz-Bandes erlauben, das insbesondere für wissenschaftliche und medizinische Anwendungen anerkannt ist. Im Vergleich zu UWB ist CW leichter zu realisieren und erlaubt eine einfachere Schaltkreis-Technologie, kann aber zu Problemen führen bei Mehrfachreflexionen, wenn sich andere bewegende Objekte vor oder hinter dem Patienten befinden. Diese müssen durch Filterung und Signalverarbeitung entfernt werden.

3.2.4 Frequenzanalyse zur Herz- bzw. Atemfrequenzbestimmung

In diesem Kapitel wird auf verschiedene Frequenzanalyse-Methoden eingegangen, die von anderen Autoren bereits zur Herz- bzw. Atemfrequenzbestimmung eingesetzt wurden:

3.2.4.1 Herzfrequenzbestimmung

Aarts et al. setzen eine FFT-Analyse des grünen Farbkanals einer RGB-Kamera ein, um die Herzrate zu bestimmen. Das kamerabasierte Verfahren wird zur Photoplethysmographie von Frühgeborenen eingesetzt. Problematisch ist die Invarianz gegenüber sich ändernden Lichtverhältnissen und auftretenden Bewegungsartefakten. [Aar+13]

In [BGK16] wird der EVM-Algorithmus für Erwachsene mit einem schmalen Bandpassfilter mit Cutoff-Frequenzen von $f_{\text{low}} = 1,33$ Hz und $f_{\text{high}} = 1,50$ Hz eingesetzt. Die Durchschnittsintensität wird über die Zeit abgebildet.

Ein anderes Verfahren, welches nicht klassisch die Frequenz analysiert, ist das Hidden Semi-Markov Modell (HSMM), welches zur Detektion von Herztönen in Phonokardiographie- und Radardaten [Wil+18] verwendet werden kann. Eingesetzt wird es bisher für Erwachsene beispielsweise in der Palliativmedizin. Ein Hidden Markov Modell (HMM) berechnet die Wahrscheinlichkeit, dass sich Daten in einem bestimmten "versteckten Zustand" befinden und trifft eine Aussage über den Zustandswechsel der Daten. Die HMMs werden von der Markov-Eigenschaft geprägt, welche lautet: Der nächste Zustand ist nur von dem Zustand abhängig, der beim aktuellen Zeitschritt eingetreten ist. Dies trifft auf Herztöne zu, da jeder Zustand nur vom vorherigen erreicht werden kann. Um auf die wahrscheinlichste Zustands-Sequenz schließen zu können, wird außerdem der Viterbi-Algorithmus angewandt [Rab89]. Ein Nachteil des HMM ist, dass es keine Aussage zur Dauer eines Zustands macht. Ein HSMM ist ein zeitabhängiger HMM und geht auf diese Problematik ein [Yu10]. [STC16] Problematisch sind diese Modelle, wenn die Zustandsabfolge unterbrochen wird, z.B. bei Arrhythmien.

3.2.4.2 Atemfrequenzbestimmung

Ähnlich wie in [BGK16] kann auch für die Atemfrequenzbestimmung eines Erwachsenen ein Bandpassfilter mit Cutoff-Frequenzen von $f_{\text{low}} = 0,15$ Hz und $f_{\text{high}} = 0,60$ Hz verwendet werden [BGK15].

Klaessens et al. beobachten die Veränderungen der Temperatur um die Nasenlöcher des Neugeborenen während des Atemzyklus. Nachteil dieser Methode ist, dass die ROI manuell festgelegt werden muss und die Detektion zudem nicht echtzeitfähig ist. [Kla+14].

Genauso kann eine kontinuierliche Wavelet Transformation (Debauchines wavelet function) des Temperaturunterschiedes der Ein- und Ausatemphase eingesetzt werden [Xu+14].

Tarassenko et al. verwenden einen FIR-Bandpassfilter für das reflektierte PPG Signal, welches einen Atmungs-Peak beinhaltet, der eine geringere Amplitude als der Herzschlag-Peak hat [Tar+14] [Dav+14].

3.2.5 Einsatz von Deep Learning bei der Vitalparameter-Detektion

Chaichulee et al. verwenden ein Multi-task CNN für die Detektion der Hautareale eines Frühgeborenen für die Atem- und Herzrattendetektion. Das CNN besteht aus einem VGG-16 core network, sowie einem Netz zur Patientendetektion und zur Segmentierung der Hautareale. [Cha+19]

In den detektierten Hautarealen kann das Photoplethysmographic imaging (PPGi)-Signal bestimmt werden [Gre+18] [Jor+18].

3.2.6 Sensorfusion zur Vitalparameter-Detektion

3.2.6.1 Thermo-RGB-Fusion zur Vitalparameter-Detektion

Die Fusion einer Thermo- und einer RGB-Kamera zur Vitalparameter-Detektion von Erwachsenen kann zum Screening von Infektionen am Flughafen eingesetzt werden. Dieses Prinzip von Negishi et al. wurde bisher nur an Erwachsenen getestet. Nur das RGB-Kamerabild wird zur Detektion der Herzrate, sowie zur Gesichtserkennung (Haar-Classifer) und Tracking (Mean-Shift-Algorithmus) verwendet, nicht das fusionierte Bild. Das Thermokamerabild wird zur Detektion der Atemrate eingesetzt. [Neg+18]

Kado et al. fusionieren zur Herzrattendetektion RGB- mit Near Infrared Kamerabildern des Gesichts von Erwachsenen. Als erstes werden die Bilder räumlich aneinander ausgerichtet. In beiden Bildtypen werden mehrere ROI-„Patches“ betrachtet und die Frequenz der Intensitätsänderung (der verschiedenen Farbkanäle) berechnet. Eine Studie an 20 Erwachsenen zwischen 20-60 Jahren zeigte, dass selbst bei unterschiedlichen Lichtverhältnissen und Beleuchtungsarten die eingeführte Methode nur einen absoluten Fehler kleiner 3 BPM für mehr als 90 % der Probanden hat. Die verwendete Kamera ist eine dual-CCD RGB-NIR Kamera (AD-130GE, JAI Ltd., Japan). [Kad+18]

Da die extrinsische Kalibrierung ausschlaggebend für eine robuste Fusion ist, sollen an dieser Stelle verschiedene Ansätze für die Fusion einer RGB- mit einer Thermokamera präsentiert werden. Dabei wird zwischen direkten Ansätzen und indirekten Ansätzen mit Hilfe eines weiteren Sensors unterschieden.

Direkte extrinsische Kalibrierung In der Publikation von St.-Laurent et al. werden die Winkel und das Verhältnis der Field of View (FOV) der beiden Kameras mit eingerechnet [SPM10]. Die Kalibrierung wird durch parallele oder konvergente optische Achsen der Kameras erleichtert.

Shivakumar et al. führen die RGB-Thermo-Kalibrierung über eine Projektion der RGB-Koordinaten in den 3D-Bereich mit Hilfe eines Tiefenbilds durch, welches durch Stereo-Tiefen-Berechnung erzeugt wird [Shi+20c]. Die 3D-Koordinate wird dann in das Thermokamerabild projiziert. Für die intrinsische Kalibrierung wird ein Schachbrettmuster verwendet.

Yang et al. führen zunächst eine Detektion der Kreismittelpunkte des Kalibrierungsmusters im RGB- und Thermokamerabild durch [YSL18]. Der Root Mean Square Error (RMSE) der Koordinaten der detektierten Mittelpunkte wird berechnet. Als Bildregistrierungs-Verfahren wird letztendlich normalized cross correlation (NCC) verwendet.

Voss et al. registrieren zwei Thermokamerabilder von einem Baby-Simulator mit einem RGB-Bild basierend auf Punktkorrespondenzen und dem random sample consensus (RANSAC) Algorithmus und bestimmen daraus eine Homographie-Matrix [Vos+22].

Indirekte extrinsische Kalibrierung Da es schwierig ist, eine hohe Genauigkeit mittels einer direkten extrinsischen Kalibrierung zu erreichen [SPM10], sollen nun Ansätze vorgestellt werden, die eine extrinsische Kalibrierung über einen dritten Sensor ermöglichen.

Krishnan und Saripalli führen eine Cross-calibration einer RGB- mit einer Thermokamera mittels eines LIDARs durch [KS17]. Zunächst wird die RGB-Kamera mit dem Laser kalibriert. Hierfür werden die Transformationen zwischen den Kanten des Kalibrierungsmusters (Schachbrett) im Kamerabild und in der 3D-Punktwolke bestimmt. Analog wird die extrinsische Transformation zwischen Thermokamera und Laser bestimmt, nur mit einem anderen Kalibrierungsmuster, das im Thermokamera-Bild sichtbar ist.

Gleichauf et al. haben im Rahmen eines Projekts mit der Deutschen Bahn eine Thermo-RGB-Laser-Fusion eingesetzt [Gle+20]. Hierbei wurde keine direkte Fusion zwischen RGB- und Thermokamera durchgeführt, sondern nur die Einzelsysteme mit dem 3D-Laserscanner fusioniert. Die Punkte des Laserscanners wurden in die zwei Bildmodalitäten eingeblendet.

Thermo-ToF-Fusion Folgende Ansätze sind bekannt, bei denen eine Thermo- mit einer ToF-Kamera fusioniert werden.

Tisha et al. fusionieren eine Thermokamera mit einer Kinect2. Zunächst wird die intrinsische Kalibrierung vorgenommen und im nächsten Schritt die Homographie beider Kameras berechnet. Die Multiplikation der Homographie mit einem Pixel-Punkt der einen Kamera ergibt den Pixel-Punkt der anderen Kamera [Tis19].

Die extrinsische Kalibrierung von Van Baar et al. basiert auf der Publikation von Yang et al. [Yan+07]. Dort wird eine Fusion einer RGB-Kamera mit einem Tiefenbild vorgenommen. Für die Thermokamera kommt ein Schachbrettmuster mit erhitzten Widerständen zum Einsatz. [Van+12]

Cao et al. fusionieren eine Thermo- mit einer RGB-D-Kamera (Kinect2) unter Verwendung des thermal-guided iterative closest point (T-ICP) Algorithmus [Cao+18]. Es wird eine Thermo-3D-Pointcloud-Registrierung angewandt. Der daraus resultierende Fehler wird mit einem "nonlinear least-squares objective" definiert und mit der Gauss-Newton Methode minimiert.

Pfützner fusioniert eine Thermokamera mit einer Kinect über die infrared sensor frames der Kinect zur Thermokamera und zur Kinect2. Hierbei wird mit dem RGB-Bild der Kinects gearbeitet. [Pfi18]

Ähnlich funktioniert die Methode von Rocco Spremolla, bei der zunächst die intrinsische Kalibrierung der Thermo- und Kinect2-Kamera ermittelt wird und dann die relativen Posen von Thermo / RGB (Kinect) zu IR / Depth von Kinect ermittelt werden. Daraus können die Transformationen berechnet werden. [Roc+16]

RGB-ToF-Fusion Ähnlich wie für die Thermokamera bestehen auch Fusionsansätze zwischen ToF- und RGB-Kameras.

Salinas et al. wollen anstelle der Berechnung der extrinsischen Parameter der ToF- und RGB-Kamera mit einer tiefenabhängigen Homographie Lookup Table arbeiten. Hierfür werden in beiden Modalitäten Punktkorrespondenzen gefunden [Sal+15].

Kim et al. verwenden fünf RGB- und drei ToF-Kameras, welche in einem Halbkreis angeordnet werden. Die intrinsische und extrinsische Kalibrierung können über das Intensitäts-Bild der ToF-Kamera berechnet werden, allerdings können Fehlregistrierungen der Tiefenbilder auftreten. [Kim+09] Das Ziel hierbei ist, eine genaue 3D-Rekonstruktion von Oberflächen zu erreichen.

Fusion unter Verwendung von Neuronalen Netzwerken Es gibt einige Ansätze, die Neuronale Netzwerke verwenden, um Fusions-Bilder zu erstellen. In jedem Fall werden die Bilder aus einer langen Distanz aufgenommen. [Tan+22] Das Netzwerk, das Alexander et al. einsetzen, wird für Anwendungen in der öffentlichen Infrastruktur verwendet und sie sagen, dass die Robustheit durch die Fusions-Bilder steigt [Ale+22].

Jung et al. verwenden ein Neuronales Netz zur Generierung von Fusions-Bildern aus Near Infrared (NIR) und verwaschenen RGB-Bildern [JZF20].

Wang et al. wenden ein Multi-Kanal CNN für die Fusion von Infrarot- und RGB-Bildern der selben Szene und Position an [Wan+22a].

Ein anderer Ansatz verwendet salience detection und ein CNN für die Generierung der Fusions-Bilder [Wan+22b].

Das Problem bei der Verwendung von Neuronalen Netzen für die Sensorfusion ist, dass viele Daten zum Trainieren des Netzwerks benötigt werden.

Es wurde zudem gezeigt, dass die Fusionsansätze nicht für die kurzen Distanzen im Inkubator geeignet sind [Gle+23].

3.2.6.2 Fusion Time-of-Flight-Kamera und Radar-Sensor

Es gibt eine Publikation, bei der eine Kinect2 (Time of Flight (ToF)) mit einem 2,4-GHz-Continuous-Wave-Radar kombiniert wird. Das System von Nguyen et al. wird WiKiSpiro genannt und während des Schlafs bei Erwachsenen eingesetzt. Hierbei wird das Atemvolumen berechnet, wobei der Effekt der Brustkorbbewegung auf die Phase des reflektierten Signals betrachtet wird. Das System funktioniert auch bei großen Bewegungen des Patienten. Diese werden mit einer Tiefenkamera getrackt und die Position und Haltung des Patienten bestimmt. Der Radar-Sensor wird dann so verschoben, dass er wieder orthogonal zum Brustkörper des Patienten positioniert ist. Die Fusion der ToF-Kamera und des Radar-Sensors dient somit nur der korrekten Ausrichtung des Radar-Sensors. Das Atemvolumen wird nur mit dem Radar-Sensor bestimmt. [Ngu+16]

3.2.7 Detektion des Gesichts

Um die Atemfrequenz im Nasenbereich aus einem Thermokamerabild bzw. die Herzrate sowie die Temperatur im Gesicht bestimmen zu können, ist eine Detektion des Gesichts erforderlich. In diesem Kapitel sollen bestehende Algorithmen und Ansätze zur Gesichtsdetektion vorgestellt werden.

3.2.7.1 Gesichtsdetektion basierend auf Neuronalen Netzen und Bildverarbeitung

Es gibt einige Detektions-Algorithmen, die bereits bei der Gesichtsdetektion von Erwachsenen eingesetzt werden, für die Anwendung an Neugeborenen allerdings ungeeignet sind, da sie nur mit Erwachsenenendaten trainiert wurden.

Nachdem der WiderFace Datensatz [Yan+15] im Jahr 2015 veröffentlicht wurde, wurde die Gesichtsdetektion von Erwachsenen immer beliebter. Der Datensatz enthält hauptsächlich Erwachsenengesichter. Von Säuglingen und Kleinkindern sind nur wenige Daten vorhanden, die überdies nicht im klinischen Umfeld aufgenommen wurden. Neuronale Netze werden im Normalfall für die allgemeine Objektdetektion verwendet. Um diese für die Gesichtsdetektion einsetzbar zu machen, müssen Anpassungen vorgenommen werden und die Netze zum Beispiel mit dem WiderFace Datensatz trainiert werden. Bekannte Methoden sind: YOLO5Face [Qi+21], RetinaFace [Den+19], MTCNN [Zha+16], Mask R-CNN [YIS19] und eine Implementierung des Faster R-CNN [JL16]. Da diese nur auf Erwachsenenendaten trainiert wurden und Erwachsene andere Gesichtsproportionen als Neugeborene aufweisen, sind die zuvor genannten Netze für Neugeborene nicht anwendbar.

In ROS gibt es bereits entwickelte Pakete zur Gesichtserkennung. Das Paket *face_detection* basiert auf Haar-Kaskaden, wurde aber nur mit Frontalaufnahmen von Gesichtern von Erwachsenen trainiert [Lud15].

face_detector basiert auf dem OpenCV face detector und auf einer Kaskade von Haar-Features. Die Methode in diesem Paket verwendet Stereo-Tiefeninformationen, um die tatsächliche Größe des Gesichts bestimmen zu können. [PL15] Ein weiteres Paket *opencv_apps_face_detection* [Oka] beruht auf der OpenCV Face Recognition, welche die Eigenface-Methode [Ope] anwendet.

Bisher gibt es nur wenig Forschung im Bereich der Kopf- und Gesichtsdetektion in Thermokamera-Bildern. Ein Ansatz beruht auf der Otsu Schwellwert-Methode, um das Thermokamera-Bild in binäre Bilder umzuwandeln [CYN14]. Cheong et al. verwenden dann die horizontale Projektion der Bilder, um das globale Minimum zu detektieren. Dies hilft dabei, die Höhe und Breite des Kopfes zu identifizieren. [CYN15]

Auch ein Histogramm of oriented gradients (HOG) Detektor kann zur Auswertung von Thermokamera-Bildern verwendet werden, wie von Kopaczka et al. gezeigt [KNM17]. Die Autoren vergleichen auch andere Machine-Learning-Methoden mit klassischen Ansätzen der Bildverarbeitung. Sie behaupten, dass mit Machine-Learning-Methoden bessere Ergebnisse erzielt werden als mit traditionellen Ansätzen.

Silva et al. [Sil+19] verwenden Deep Learning Methoden wie das Neuronale Netz YOLOv3. Sie wenden Transfer Learning an, um ein mit RGB-Bildern trainiertes Netzwerk für die Kopfdetektion in Thermokamera-Bildern einsetzbar zu machen.

Vuković et al. [Vuk+19] haben einen ähnlichen Ansatz. Bei ihnen kommt ein R-CNN zum Einsatz, das zwar gute Ergebnisse liefert, aber nicht in Echtzeit einsetzbar ist.

Die SCRFD DL Architektur von Guo et al. [Guo+21] lieferte gute Ergebnisse bei der WiderFace Challenge. Mucha et al. [MK22] haben diese dahingehend angepasst, dass damit Gesichter in Thermokamera-Bildern erkannt werden können.

Die Autoren Bhowmik et al. beschreiben für die Gesichtsdetektion geeignete Infrarot-Typen, wovon Long wave Infrared (LWIR) mit einer Wellenlänge von 8 -14 nm am besten ist. Die Publikation fasst die verschiedenen Bilddatenbanken mit Thermo- und RGB-Kamera-Bildern von Gesichtern, die insbesondere als Benchmark für die Gesichtsdetektion dienen, zusammen. Die IRIS (Imaging, Robotics and Intelligent System) Thermal / Visible Face Database [vci] ist eine davon. Eine weitere Datenbank ist die Terravic Facial IR Database [Cor]. Beide Datenbanken beinhalten nur Aufnahmen des Gesichts verschiedener Erwachsener, aber keine Bilder von Kindern und Neugeborenen, sie können daher nicht zum Training von Neuronalen Netzen für Detektionen an Frühgeborenen genutzt werden. Bhowmik et al. beleuchten zudem unterschiedliche Fusionsmethoden, die im Bereich der Gesichtsdetektion zum Einsatz kommen können. Diese können in Feature Level Fusion, Decision Level Fusion und Pixel / Data Level Fusion eingeteilt werden. [Kan+11]

3.2.7.2 Detektion von Körperarealen bei Neonaten basierend auf Neuronalen Netzen und Bildverarbeitung

Ohne weitere Anpassungen können die bisher vorgestellten Methoden nicht bei Neu- und Frühgeborenen angewandt werden, da ihre Körperproportionen stark von denen Erwachsener abweichen.

Sitthichok et al. stellen ein selbst entworfenes Multi-Task-CNN vor, das die sichtbaren Bereiche der Haut des Neugeborenen in einem RGB-Bild segmentiert [Sit+17]. Dadurch kann der Torso

und der Kopf sichtbar gemacht werden, was letztendlich der Detektion der Herz- und Atemrate dient.

Das Netzwerk wurde von Green et al. [Gab+18] dahingehend erweitert, dass neben Kopf und Torso auch die Windel detektiert werden kann. Es basiert auf dem Faster R-CNN [Sha+15]. Neben dem Kern-Netzwerk (VGG16 [SZ14]), gibt es drei Zweige für die Patientendetektion, die Haut-Segmentierung mit Bewegungsüberwachung und einen Zweig für die Detektion der Körperregionen. Für die Kopf-Klasse wird eine Average Precision von 98,2 erreicht. Der Datensatz beinhaltet 2269 Bilder von 15 Probanden verschiedener ethnischer Hintergründe und verschiedenen Geschlechts. Alle Aufnahmen wurden in einem Brutkasten während des Aufenthalts auf der NICU gemacht.

Auch das RetinaNet kann zur Detektion des Thorax und des Gesichts von Neonaten in RGB-Bildern verwendet werden [Tsu+17]. Kyrollos et al. [Kyr+21] haben ein Modell mit Transfer Learning trainiert, indem Gewichte angewandt wurden, die zuvor auf dem ImageNet Datensatz [Rus+14] vortrainiert wurden. Mit ResNet50 als Kernnetzwerk wurde eine mean Average Precision von 56,9 erreicht. Die Detektionsergebnisse werden zur Atemdetektion verwendet. Es kamen RGB-Bilder (200 Bilder pro Proband bei 5 Probanden) zum Einsatz, die nicht im Inkubator aufgenommen wurden.

Lu et al. präsentieren ein CNN mit pyramidisch hierarchischen Features [Gua+18]. Zum Extrahieren von Features im normalisierten Bild kann eine Anpassung von VGG16 [SZ14] verwendet werden. Um verschiedene Größen von Gesichts-Features des Neugeborenen vorhersagen und detektieren zu können, werden multi-scale Feature Maps ausgewählt. Die folgenden zwei Zweige haben die Aufgabe der Klassifizierung von Gesichts- und Nicht-Gesichts-Regionen und der Regression für die Position des Gesichts. Der verwendete Datensatz enthält 3600 Bilder mit verschiedenen Perspektiven, Gesten, Gesichtsausdrücken und Hintergründen. Es ist unbekannt, ob die Bilder auf einer NICU aufgenommen wurden. Für die Gesichtsregionen wurde eine Average Precision von 92,7 erreicht.

In der Literatur kommen YOLOv3 [JA18] und YOLOv5 [Joc+20] bei der Detektion von Neonaten zum Einsatz. Nagy et al. verwenden YOLOv3 für die Detektion von Kopf, Torso, Flasche und Hände des Pflegepersonals oder der Eltern [Nag+21]. Wenn dieses mit einem LSTM (Long Short-Term Memory) kombiniert wird, können statische Objekte im RGB-Video detektiert werden. Letztendlich kann damit der Puls und die Atemrate bestimmt werden. Die Autoren haben 96 h gelabelte RGB-Video-Daten, die auf der NICU aufgenommen wurden, verwendet. Die Genauigkeit der Objektdetektion wird in der Publikation nicht genannt. Da die Klassifizierung gute Ergebnisse liefert (97,9 sensitivity, 97,5 specificity und ein F1 score von 98,1), wird auch bei der Objektdetektion eine hohe Genauigkeit erwartet.

Auch Khanam et al. [Kha+21] verwenden YOLOv3 für die Gesichtsdetektion von Frühgeborenen im Inkubator. Es wurde mit Transfer Learning gearbeitet und mit den Gewichten von einem YOLOv3 Netz, das zuvor mit dem MS COCO Datensatz trainiert wurde. Für den Transfer wurden 473 Bilder aus dem Internet verwendet. Angaben zur Genauigkeit der Gesichtsdetektion werden nicht gemacht.

Hausmann et al. trainieren YOLOv5 auf einem eigenen RGB-Datensatz mit gelabelten Gesichtern von Neonaten im klinischen Umfeld und erreichen dabei eine Echtzeit-Gesichtsdetektion [Hau+22]. Sie erhalten eine Genauigkeit von 68,7. Der University of South Florida Multimodal Neonatal Pain Assessment Datensatz (USF-MNPAD-I) [Sal+21] wurde verwendet. Außerdem wurde zum Vergleich mit dem WiderFace Datensatz trainiert, der eine Genauigkeit von 7,37 erreichte.

Dosso et al. [Dos+22] vergleichen verschiedene Gesichtsdetektions-Modelle wie RetinaFace [Den+19] und YOLO5Face [Qi+21], die mit dem WiderFace Datensatz trainiert wurden. Im nächsten Schritt wenden sie Transfer Learning an mit ihrem eigenen Datensatz und erreichen eine Genauigkeit von 81,45. Am besten schnitt ein optimiertes YOLO5Face ab, welches sie NICUFace nennen. In ihrem Datensatz befanden sich 33 Probanden, die über 153 h gefilmt wurden. Eine Thermokamera wurde auch eingesetzt, allerdings wurden die Daten nicht für ihre Veröffentlichung ausgewertet.

Antink et al. [Ant+20a] verwenden einen angepassten ResNet50 Encoder, um verschiedene Körperregionen von Frühgeborenen zu segmentieren. Es werden Segmentierungs-Daten von Erwachsenen aus einem frei verfügbaren Datensatz eingesetzt. Für das Transfer Learning wurde ein zweiter Datensatz mit Near-infrared- (NIR) und RGB-Bildern von Frühgeborenen genutzt. Für die Segmentierung des Kopfes in RGB-Bildern wird eine Genauigkeit (Intersection over Union (IoU)) von 82,0 erreicht. Für den Torso liegt der IoU bei 41,0. In den NIR-Bildern wurden schlechtere Genauigkeiten erreicht, und zwar 62,0 für den Kopf und 14,0 für den Torso. Voss et al. erweitern diesen Ansatz, indem sie eine U-Net Architektur verwenden, um mit RGB-, NIR- und Fusions-Bildern zu trainieren [Vos+23]. Ihre Fusion findet nicht im Bildbereich, sondern auf der Feature-Ebene statt. Sie erreichen mit ihrem Fusions-Modell einen mittleren IoU von 85,0. Die Ergebnisse für RGB liegen bei 85,0. Bei den NIR-Bildern sinkt die Genauigkeit stark ab auf einen mittleren IoU von 75,0. In allen Modalitäten hat die Segmentierung des Torsos am schlechtesten abgeschnitten.

Beppu et al. forschen im Bereich der Detektion verschiedener Körperregionen in Thermokamera-Bildern [Bep+22]. Sie möchten sechs Körperregionen wie z.B. Kopf, Torso, Arme und Beine erkennen können. Dafür wurde YOLOv5 [Joc+20] trainiert. Als Datensatz kamen 4820 Thermokamera-Bilder von 26 verschiedenen Probanden zum Einsatz (1872 h Videomaterial). Die Bilder wurden auf der NICU in einem Inkubator aufgenommen. Für die Kopfdetektion wurden Genauigkeiten von bis zu 94,8 erreicht. Letztendlich wurden die detektierten Regionen zur Messung der Körpertemperatur der Frühgeborenen verwendet.

Es gibt außer Deep Learning Ansätzen auch noch andere Möglichkeiten für die Gesichtsdetektion. Beispielsweise kann eine CIE L^*a^*b Intensitäts-basierte Detektion verwendet werden (beschrieben von Neophytou et al. [NM08] und Fairchild [Fai13]). Awais et al. [Muh+20] wenden diese in ihrem Framework NFAR (Neonatal Face Attributes Recognition) für RGB-Bilder an. Mit NFAR werden Genauigkeiten von bis zu 94,9 erreicht.

Von Rivera et al. [RML14] wird ein Algorithmus zur Segmentierung des Gesichts eines Frühgeborenen aus einem Infrarotkamerabild vorgestellt. Hierbei werden aktive Konturen betrachtet,

um die Region of Interest vom Hintergrund zu separieren und Fuzzy Sets eingesetzt, um zusammengehörige Regionen im Gesicht zu finden. Es wurde eine automatische und eine manuelle Detektion implementiert, verglichen und die hieraus berechneten Temperaturwerte betrachtet. Der manuelle Modus liefert derzeit bessere Ergebnisse. Wenn das zu segmentierende Bild gedreht ist, können Fehler auftreten, die noch behoben werden müssen. Solange keine Rotationsinvarianz besteht, ist der Algorithmus nicht komplett einsetzbar. [RML14]

3.2.7.3 Gesichtsdetektion basierend auf fusionierten Daten

Zwei Fusions-Ansätze für die Gesichtsdetektion basierend auf Thermokamera- und visuellen Bilddaten werden von Bebis et al. vorgestellt. Der eine Ansatz ist ein pixelbasierter, der im Wavelet-Bereich durchgeführt wird, der andere ist merkmalsbasiert im Eigenraum-Bereich. Eine einfache Pixel-zu-Pixel Fusion hat den Nachteil, dass räumliche Informationen verloren gehen. Wenn jedoch an verschiedenen Auflösungs-Ebenen eine Fusion durchgeführt wird, können Merkmale mit verschiedenen räumlichen Ausdehnungen, bei der sie am hervorstechendsten sind, mit einbezogen werden. Beide Ansätze wurden an der Equinox-Datenbank getestet, an LWIR- und RGB-Kamerabildern. Die Datenbank enthält nur Frontalaufnahmen der Gesichter von Erwachsenen. Im Vergleich liefert die Fusion im Wavelet-Bereich bessere Ergebnisse, ist aber rechenintensiver. In der Zukunft sollen Haar-Wavelet-Transformationen angewandt werden. [Beb+06]

Chen et al. schlagen ein Neuronales Netz vor, um Fusions-Bilder zur Gesichtsdetektion zu generieren [Che+22]. Sie schreiben, dass die Erkennung durch die Verwendung von fusionierten Daten verbessert wird. Das Netz ist nur mit Erwachsenendaten trainiert.

Bis zum heutigen Stand gibt es keine Literatur zur Gesichtsdetektion in fusionierten Bildern von Neonaten.

3.2.7.4 Neuronale Netze mit Fusionsdaten trainiert

Es gibt wenige Ansätze, bei denen fusionierte Daten als Input für ein Neuronales Netz verwendet werden. Die Anwendungen liegen in der Umgebungswahrnehmung für das autonome Fahren [Vad+22] und der Erkennung von Fußgängern (Faster R-CNN) [SJP19]. In beiden Fällen werden die Fusionsbilder durch ein Neuronales Netz generiert. Shopovska et al. erwähnen, dass die Fusions-Bilder informativer sind.

Es gibt RGB-T Tracker, die Objekte in RGB-Thermo-Bildern mit Hilfe eines Neuronalen Netzes tracken. Dies ist nicht dasselbe wie Bild-Fusion, da für sie die Fusions-Genauigkeit nicht wichtig ist und nur der Vordergrund des Bildes von Interesse ist [Zha+20] [Zha+19].

3.2.8 Kontaktlose Temperaturmessung

Neben den Temperaturmessungen über eine rektale Sonde, die vermieden werden soll, gibt es Verfahren, die die Temperatur durch Messungen im Ohr [OKC94] [Fra97], über temperatur-sensitive Materialien [Eyc+10], durch Ausrichtung des Temperatur-Messgerätes auf bestimmte Körperteile [Bel99], wie die Stirn [Pom01] oder über visuelle Erfassung [Cra+16] bestimmen. Häufig müssen die Geräte von einer Person nahe an den Patienten gehalten werden.

Bereits 1971 wurden erste Studien durchgeführt, bei denen Infrarot-Thermographie eingesetzt wurde, um thermographische Veränderungen von Neugeborenen zu detektieren. In der Publikation von Topalidou werden alle relevanten Publikationen aufgeführt, bei denen Thermokameras in der Neonatologie eingesetzt werden. Die abgedeckten Themengebiete können folgendermaßen aufgegliedert werden: Oberflächentemperatur und Temperaturverteilungen [VA71] [Tah+72] [RPL72] [PES72] [Bha+76] [PLU77] [CSS14] [Ek+99] [Oya+97] [Chr+03] [SW08] [Ric+10] [HFB12] [KGR11] [Kno+13] [Kur+15] [Bac+18], tiefer liegende Strukturen und Organe [Bha+76] [PLU77], klinische Zustände (Pathologien) [Bha+76] [PLU77] [SW08] [Ric+10] [HFB12] [KGR11], Wärmeverlust [Tah+72] [Ek+99] [Ada+00], Atmungsdetektion [Ada+00] [Abb+11], Sicherheit von Laser-Akupunktur [Kur+15] und Schlafneigungen [Bac+18]. Im Nachfolgenden soll auf die besonders relevanten Publikationen detaillierter eingegangen werden.

1980 wurden erste Infrarot-Farb-Thermographie-Messungen zur Erstellung von Temperaturkarten von Frühgeborenen im Inkubator durchgeführt. Als Referenzdaten wurden Temperaturdaten von einem Thermokoppler herangezogen. Eine Abweichung von nur $\pm 0,6^\circ\text{C}$ konnte dabei erzielt werden. Aus den Messungen wurde ersichtlich, dass die Oberflächen-Temperaturen im Bauchbereich ($36,2^\circ\text{C}$), Rücken ($36,01^\circ\text{C}$) und an der Stirn ($35,83^\circ\text{C}$) am höchsten sind. [CSS14]

Abbas et al. beschreiben in [Abb+12], wie das Mapping der Oberflächentemperatur von Frühgeborenen durch eine Infrarotkamera präziser gemacht werden kann. Externe Einflüsse wie wechselnde Umgebungstemperatur, unterschiedliche Luftgeschwindigkeit und Feuchtigkeit, sowie Temperaturdrifts bzw. Fehler bei Temperaturmessungen mit der Infrarot (IR)-Kamera (methodische Messfehler, Kalibrierungsfehler, elektronische Bahnfehler) müssen beachtet und kompensiert werden. Die vorgeschlagenen Kompensationsmethoden beziehen sich auf die verschiedenen Lagerungsmöglichkeiten, d.h. den geschlossenen Inkubator, das Thermobett und die Känguru-Methode. Abbas et al. stellen eine Formel auf, aus welchen Bestandteilen sich die Temperatur, die beim IR-Detektor ankommt, berechnen lässt:

$$(3.3) \quad T_{detector} = M_{cam}(\Phi_{obj} + \Phi_{refl} + \Phi_{atm})$$

mit $T_{detector}$ als Temperatur, die von der Kamera wahrgenommen wird, M_{cam} ist ein Parameter, der vom atmosphärischen Dämpfungsverhalten, den optischen Komponenten der Kamera und den Eigenschaften des Detektors abhängt. Mit Φ_{obj} als Wärmestromdichte, die vom Objekt emittiert wird, Φ_{refl} als reflektierte Wärmestromdichte und Φ_{atm} als Wärmestromdichte der Umgebungskomponenten.

Die zugehörige medizinische Studie wird in [Hei+13] beschrieben.

Auch zur Überwachung der Thermoregulation von Neugeborenen kann die Infrarot-Thermographie eingesetzt werden [Abb+11]. Die Thermoregulation kann Hinweise auf den Gesundheitszustand des Neugeborenen liefern.

Knobel-Dail et al. bauen auf Forschungsergebnisse von Abbas und Leonhardt [AL14] und Abbas et al. [Abb+11] auf, um die Hypothermie bei Frühgeborenen zu untersuchen und zu überwachen. Trotz der warmen und feuchten Umgebung im Inkubator war es möglich, die Temperatur der sehr

jungen Frühgeborenen zu messen. Als Referenzmessverfahren wird eine Y series Steri-Probe® (Model 499B, Cincinnati Sub-Zero, Cincinnati, OH), die einen Thermistor beinhaltet, eingesetzt. Die Messungen wurden am Fuß und der Bauchgegend (Abdominal) des Kindes durchgeführt. Die Messung mit dem Thermistor am Bauch ergab 36,44 °C, wohingegen mit der Thermokamera ein Wert von 36,57 °C gemessen wurde. Die Genauigkeit ist also hoch. In ihrer Publikation zeigen die Autoren auch, dass es möglich ist, Temperaturmessungen der Hautoberfläche durch eine Infrarotkamera zu erhalten, da die Hautoberfläche ein hohes Emissionsvermögen von 0,98 hat. [Kno+17]

Adams et al. vergleichen die infrarot-thermographische Kalorimetrie der Durchschnitts-Hauttemperatur von Kindern mit der Atmungs-indirekten Kalorimetrie. Zwischen den beiden Verfahren besteht eine Abweichung von 1,3 %. Mit der Kalorimetrie kann die gesamte Energieaufwendung, sowie der Temperaturverlust von frühgeborenen Kindern bestimmt werden. Die Durchschnitts-Oberflächentemperatur des Körpers lässt sich aus mehreren Bestandteilen bestimmen: Wärmestrahlung, Wärmeumwälzung, Wärmeleitung und Verdunstung. Insbesondere wurde bei den Untersuchungen überprüft, wie sich die Durchschnitts-Oberflächentemperatur verändert, wenn das Kind eine Mütze trägt. [Ada+00]

Laxminarayan et al. beschreiben in ihrer Publikation wie aus der Hauttemperatur unter Berücksichtigung der Herzrate und weiterer Parameter die Körperinnentemperatur bei Erwachsenen bestimmt werden kann. [Lax+18]

$$(3.4) \quad \frac{d\Delta HR}{dt} = -\alpha_1 \Delta HR + \beta A_c^4$$

mit ΔHR als Änderung der Herzrate vom Ruhezustand HR_0 , α_1 als Ratenkonstante von HR und A_c als Erhöhung der Herzrate resultierend aus physikalischer Aktivität. $A_c = 0$ im Ruhezustand, $A_c = 1$ bei leichter Aktivität, $A_c = 2$ bei mittelmäßiger Aktivität, $A_c = 3$ bei hoher Aktivität und $A_c = 4$ bei sehr hoher Aktivität.

$$(3.5) \quad \frac{d\Delta T_c}{dt} = -\alpha_2 \Delta T_c + \gamma_1 S(\Delta HR) - \gamma_2 (T_c - T_s)$$

$$(3.6) \quad \frac{d\Delta T_s}{dt} = -\alpha_3 (T_s - T_a) - \alpha_4 (P_s - P_a) + \gamma_2 (T_c - T_s)$$

mit $\Delta T_c = T_c - T_{c0}$ und $\Delta T_s = T_s - T_{s0}$ mit T_{c0} und T_{s0} als initiale Körperinnen- und Körperoberflächentemperatur. P_s ist der Druck von Wasserdampf bei T_s und P_a ist der Druck von Wasserdampf bei einer gegebenen Temperatur T_{a0} und relativen Feuchtigkeit. T_{c0} wird hier auf 37 °C gesetzt und T_{s0} als der Durchschnitt der gemessenen T_s während der ersten zehn Minuten der Datensammlung. α_2 ist die thermoregulatorische Konstante von T_c , γ_1 ist die Rate der Wärmeerhöhung, die durch die metabolische Aktivität (HR) verursacht wird. $S(\Delta HR)$ ist die Standard Sigmoid-Funktion, um den Wärmegewinn auf Null zu reduzieren, wenn HR unter den Ruhezustand abfällt. γ_2 ist die Rate der Wärmeübertragung vom Körperinneren zur Körperoberfläche. α_3 ist die Rate der konvektiven Wärmeübertragung von der Haut zur Umgebung. α_4 ist die Rate des Wärmeverlusts an die Umgebung aufgrund von Schwitzen.

3.2.9 Zusammenfassung

Zusammenfassend kann gesagt werden, dass die existierenden Methoden Lücken hinsichtlich der Lösung der Problemstellung aufweisen. Die meisten Verfahren wurden bisher nur für Erwachsene implementiert. Einige Methoden weisen bei wechselnden Lichtverhältnissen Probleme auf. In der untenstehenden Tabelle (Tabelle 3.1) werden die Publikationen aufgeführt, die Methoden zur Vitalparameter-Detektion am Frühgeborenen beschreiben und mit welcher Genauigkeit im Vergleich zum Goldstandard sie arbeiten. Für den Einsatz von Time-of-Flight-Kamera bzw. Structured-Light-Kamera stehen keine derart geeigneten Veröffentlichungen zur Verfügung. Wie zu sehen ist, gibt es Abweichungen zum Goldstandard. Im folgenden Kapitel soll ein theoretischer Ansatz vorgestellt werden, um diesen Lücken zu begegnen.

Tab. 3.1: Stand der Technik im Vergleich zum Goldstandard

Sensortyp	Autor	Detektion	Abweichung	Referenz	Lücke
IR-Kamera	[Per+17]	Atmung	3,42 %	EKG	manuelle Bestimmung der ROI notwendig
RGB-Kamera	[Dav+14]	HR RR PPG	4 BPM - -	EKG IPG PPG	bei Dunkelheit und sich ändernden Lichtverhältnissen nicht einsetzbar
CW	[Bel+22]	Atmung	5 BPM	IPG	
IR-UWB	[Lee+21]	RR HR	0,17 BPM 0,23 BPM	IPG EKG	
IR-Kamera	[Kno+17]	Temp	0,13 K	Thermistor	bestimmt statt Kerntemperatur Hauttemperatur

4 Theoretischer Ansatz

Der theoretische Ansatz ergibt sich aus den Lücken des Standes der Technik in Kapitel 3. Zunächst sollen die Anforderungen an das System analysiert werden.

4.1 Anforderungsanalyse

Folgende Anforderungen werden an das zu entwickelnde System gestellt.

Anforderungen

- Ein Demonstrator für ein kontaktloses Monitoringsystem soll erstellt werden, der die nötige Sensorik in den Inkubator integriert
- Das System darf andere medizinische Geräte nicht beeinflussen
- Die Robustheit der Vitalparameter-Erkennung muss gewährleistet sein
- Die Vitalparameter Herzschlag, Atmung und Körpertemperatur sollen erkannt werden
- Die jeweilige Region of Interest muss automatisch bestimmt werden können
- Der Komfort für das Kind soll erhöht und dabei Sterilität gewährleistet werden
- Das kontaktlose System darf das Handling beim Wiegen und anderen Untersuchungen nicht beeinträchtigen
- Klinische Anforderungen an die Genauigkeit: ± 5 BPM für die Herzrate, ± 2 BPM für die Atemrate (bei sediertem Kind) und $0,1$ °C für die Körpertemperatur.
- Der Einsatz muss auch bei Dunkelheit und bei sich ändernden Lichtverhältnissen möglich sein
- Die Sicherheit des Kindes muss gegeben sein
- Der Zugang zum Kind durch das Personal muss gewährleistet sein
- Untersuchungen und Eingriffe dürfen nicht behindert werden

4.2 Konzept

Um die oben genannten Anforderungen zu erfüllen, ist folgendes Konzept für ein Gesamtsystem angedacht: Die Idee ist, heterogene Sensorik in einer Sensorbox im Inkubator zu integrieren, um so die Herz- und Atemrate sowie die Körpertemperatur ermitteln zu können. Die benötigte Software muss entwickelt werden. Durch den Einsatz von Sensorfusion soll die Robustheit des Systems maximiert und die höchstmögliche Genauigkeit garantiert werden.

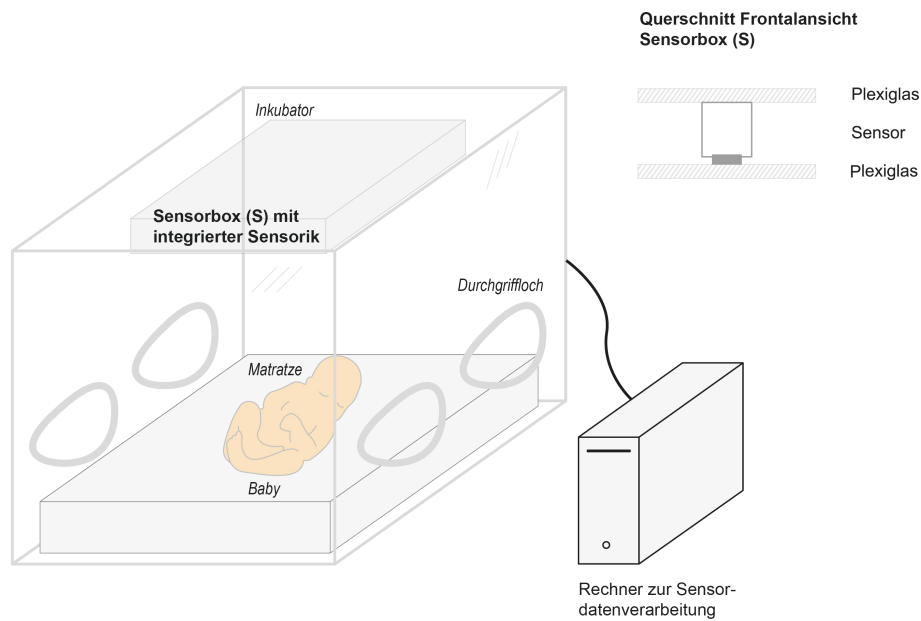


Abb. 4.1: Konzept des Gesamtsystems. Heterogene Sensorik soll in einer Sensorbox im Inkubator integriert werden.



Abb. 4.2: Säuglingsinkubator Isolette 8000 von der Firma Dräger. Bildquelle: Fa. Dräger

4.3 Sensoraufbau

Die Kameras zur Vitalparameterdetektion müssen so beschaffen sein, dass ein Großteil der Inkubatormatratze im FOV der Kamera liegt. Das Inkubatormodell, das auf der neonatologischen Intensivstation in der Kinderklinik in Erlangen eingesetzt wird, ist ein Säuglingsinkubator Isolette 8000 von der Firma Dräger (siehe Abbildung 4.2). Die Größe der Matratze des Isolette 8000 Modells beträgt 40 cm x 76 cm. Beim Isolette 8000 Inkubator wird mit einer Höhe von 40 cm vom Boden zum Inkubatordeckel gerechnet.

4.3.1 Kameras und Radar-Sensor

Zu Beginn wurden verschiedene Kameras und Sensoren darauf untersucht, ob sie die benötigten Anforderungen erfüllen. In Tabelle 4.1 werden die verschiedenen Kameraoptionen und ihre Eigenschaften vorgestellt.

Jeder Sensor hat Stärken und Lücken, die durch geschicktes Kombinieren ggf. ausgeglichen werden können, um die Robustheit gegenüber widrigen Umgebungsbedingungen wie z.B. Dunkelheit zu erhöhen. Dabei kann es sich sowohl um eine echte Sensorfusion, also eine Erstellung von Punktkorrespondenzen handeln, wie auch um eine Synchronisierung z.B. über gemeinsame Timestamps. RGB z.B. hat den Vorteil, dass das 2D-Farbbild reichlich Features wie Kanten etc. enthält, aber den Nachteil, dass bei Dunkelheit keine Daten zu gewinnen sind. Diese Lücke kann durch eine Kombination mit Time-of-Flight-Kamera, Radar-Sensor oder Thermokamera geschlossen werden. Die geringen Features im Thermokamera-Bild können wiederum durch Kombination mit RGB ausgeglichen werden. Da RGB- und Thermokamera beide 2D-Bilder

Tab. 4.1: Übersicht Kameras

Kameratyp	Kamera	Spektrum	Tag	Nacht	Auflösung	Öffnungswinkel
RGB	Webcam Logitech		++	--	HD 1080p	83° in der Diagonale
NoIR	Raspberry Pi Camera		++	Nur mit IR-Beleuchtung		
Low Light Sensitivity	Waveshare OV2710		++	+ (Farbbild)	1920 x 1080	145° in der Diagonale
SWIR	SWIR	900 - 1700 nm	++	+ (kein Farbbild)	640 x 512	von Objektiv abhängig
RGB	JAI GO-2400C-USB Kamera		++	-	1146 x 716	185° x 185°
Thermo	Optris PI 640	7,5 - 13 µm	++	++	640 x 480	90 x 64°
Structured Light	Astra Orbecc Stereo		++	++	640 x 400 (Tiefe)	H67,9° V45,3° D78° ±3.0°
Time-of-Flight	pmd Pico Flexx		++	++	224 x 171	62° ± 45°
Time-of-Flight	pmd Flexx2		++	++	224 x 172	56° ± 44°

liefern, ist eine echte Fusion mittels Punktkorrespondenzen denkbar. Der Radar-Sensor kann Hübe im Mikrometerbereich genauer detektieren als die Time-of-Flight-Kamera. Bei dieser Kombination ist allerdings keine Punktkorrespondenz möglich, sondern es muss eine Synchronisierung erfolgen. Dasselbe gilt für die Kombination von Radar-Sensor mit Thermokamera oder RGB-Kamera.

Wie aus der Tabelle zu erkennen ist, ist die Logitech RGB-Kamera bei Dunkelheit nicht einsetzbar. Dies muss aber möglich sein, da der Inkubator auf der Intensivstation häufig mit Tüchern oder Decken abgedunkelt wird. NoIR-Kameras können nur mit einer zusätzlichen Infrarot-Beleuchtung, die vermieden werden soll, verwendet werden. Bei der Waveshare Low Light Sensitivity Kamera ist beim ersten Test starkes Rauschen aufgefallen. Die Shortwave Infrared (SWIR)-Kamera liefert bei Dunkelheit kein Farbbild. Aus diesen Gründen und weil der Öffnungswinkel des Objektivs geeignet ist, fiel die Entscheidung auf die JAI GO-2400C-USB RGB-Kamera.

Die Optris PI 640 ist eine Thermokamera, mit der bereits gute Erfahrungen gemacht wurden. Sie soll gemeinsam mit der o.g. RGB-Kamera zum Einsatz kommen, da im RGB-Bild mehr Kanten und Features zu sehen sind, als im Thermokamera-Bild.

Die Structured Light bzw. Time-of-Flight Kameras liefern unabhängig von der Beleuchtungssituation eine 3D-Punktwolke.

Die Theorie hinter den genannten Kameras soll nun beschrieben werden.

4.3.1.1 RGB-Kamera

RGB-Kameras sammeln sichtbares Licht und stellen ein Farbbild dar, wie es ein Mensch wahrnehmen würde [inf].

4.3.1.2 Thermokamera

Thermokameras sind 2D-Infrarot-Thermometer, die die emittierte Wärmestrahlung von Objekten detektieren. Im Thermokamera-Bild werden Temperaturgradienten graphisch dargestellt.

4.3.1.3 Tiefenkamera

Es gibt verschiedene Arten von Tiefenkameras. In diesem Kapitel soll die Funktionsweise von Structured Light Kameras und Time-of-Flight Kameras erläutert werden.

Structured-Light-Kamera Structured-Light-Kameras senden ein Lichtmuster aus, welches verzerrt wird, wenn ein Objekt davor liegt [Emm]. Die Verzerrung des Musters kann genutzt werden, um die Distanz zu berechnen.

3D-Time-of-Flight-Kamera 3D-Time-of Flight (ToF)-Kameras können auf zwei verschiedenen Messprinzipien basieren: der Impuls-Time-of-Flight-Methode oder der Phasenverschiebungsmethode.

Impuls-Time-of-Flight-Methode Bei der Impuls-ToF-Methode wird die Distanz d zum Objekt aus der Zeit t , die der Impuls vom Senden zum Empfangen benötigt, berechnet:

$$(4.1) \quad d = \frac{t \cdot c}{2}$$

mit c als Lichtgeschwindigkeit.

Phasenverschiebungsmethode Bei der Phasenverschiebungsmethode wird über die Phasenverschiebung des reflektierten modulierten Signals die Distanz d zum Objekt berechnet:

$$(4.2) \quad d = \frac{\phi}{2 \cdot \pi} \cdot \frac{\lambda}{2}$$

mit Phasenverschiebung ϕ und der Wellenlänge λ des modulierten Signals.

Filterung von 3D-Punktwolken Bei ToF-Kameras können nicht-systematische und systematische Fehler auftreten [May09]. Zu den nicht-systematischen Fehlern werden das Wandlungsrauschen, das Quantisierungsrauschen und das Quantenrauschen gezählt [Lan00]. Außerdem können Mehrfachreflektionen z.B. an Kanten und Lichtstreuung, die beispielsweise auftritt, wenn ein helles Objekt vor einem dunklen Hintergrund steht, zu Fehlern führen. Systematische Fehler, die durch Kalibrierung entfernt werden können, sind beispielsweise Distanz-bezogene Fehler (Circular Distance Errors), Fixed Pattern Noise (FPN) und Amplituden-bezogene Fehler, die z.B. durch eine inhomogene Beleuchtung hervorgerufen werden [May09]. Um die Folgen der nicht-systematischen Fehler zu minimieren, müssen die ToF-Daten gefiltert werden, damit eine präzise Weiterverarbeitung der Punktwolke ermöglicht wird.

Han et al. vergleichen unterschiedliche Arten von Filterung bei 3D-Punktwolken (nicht nur für ToF-Punktwolken). Diese können in vier verschiedene Klassen unterteilt werden: Statistik-basierte, Nachbarschafts-basierte, Projektions-basierte und Partial Differential Equations (PDE)-basierte Filter. Die Autoren untersuchen das Verhalten folgender Filter: Voxel Grid (VG), Edge Aware Resample (EAR), L0 Minimierung und Moving Least Square (MLS). Ihr Ergebnis ist, dass VG bei großen Punktwolken zwar schnell ist, aber der Filtereffekt ungenügend ist. EAR und L0 Minimierung liefern eine gute Filterung, benötigen aber eine hohe Rechenzeit. MLS scheint eine gute Ausgewogenheit zwischen Genauigkeit und Rechenaufwand aufzuweisen. [Han+17]

May empfiehlt ein Median-Filter zur Rauschreduktion von ToF-Punktwolken [May09]. Ein weiterer verwendeter Filter, der an dieser Stelle vorgestellt werden soll, ist der Statistical-Outlier-Removal-Filter.

Median-Filter Das Median-Filter wird typischerweise zur Rauschminimierung eingesetzt. Dabei wird in einem vorgegebenen Fenster der Wert des mittleren Pixels auf den Median der Fensterwerte gesetzt. Das Filter ist robust gegenüber Ausreißern und erhält Kanten. Außerdem eignet es sich gut bei individuellen Pixeln mit extremen Werten. [Ich]

Statistical-Outlier-Removal-Filter Rusu et al. beschreiben in ihrer Publikation ein Filter, das auf der Berechnung der Verteilung der Distanzen zwischen einem Punktwolken-Punkt und seinen Nachbarn basiert. Für jeden Punkt wird die mittlere Distanz zu seinen Nachbarn berechnet. Es wird angenommen, dass eine Gauss-Verteilung zu Grunde liegt und somit werden Punkte, die außerhalb eines Intervalls liegen, welches über die Durchschnittsdistanz und Standardabweichung definiert wird, als Ausreißer angesehen. [Rus+08]

4.3.1.4 Six-Port-Radarsensor

Neben Kameras soll auch ein Radar-Sensor eingesetzt werden. Es gibt unterschiedliche Arten von Radarsignalen: Continuous Wave (CW), Frequency Modulated Continuous Wave (FMCW) und Ultra Wide Band (UWB). Eine weitere Variation, welche CW-Radar benutzt, ist der Six-Port-Radar-Sensor, welcher interferometrische Messungen vornimmt. Ursprünglich wurde das Verfahren 1970 von Engen und Hoer für Leistungsmessungen eingeführt. [Koe+16]. Der Vorteil dieses Verfahrens ist, dass Distanzen mit einer Genauigkeit im Mikrometerbereich gemessen

werden können, wobei geringe Kosten und ein geringer Leistungsverbrauch auftritt. Das Messsystem besteht aus einem Radio Frequency (RF) Front End und einem Direct Current Back End [Wil+17a]. Bei Six-Port-Radarsystemen muss nur die Phasenverschiebung zwischen der gesendeten Referenz und der am Ziel reflektierten Welle gemessen werden [Vin+13]. In diesem Fall muss keine Fouriertransformation eingesetzt werden. Die relative Distanz kann über folgende Formel ermittelt werden [Wil+17a]:

$$(4.3) \quad \Delta x = \frac{\Delta \varphi}{2 \cdot \pi} \cdot \frac{\lambda}{2}$$

mit Δx als relative Distanz, $\Delta \varphi$ als relative Phasenverschiebung und λ als Wellenlänge.

Funktionsweise Ein Analog-Signalgenerator erzeugt das Radiofrequenz-Signal. Das Signal wird durch einen 10 dB Koppler aufgespaltet. Der gedämpfte Teil wird an den Six-Port übertragen. Der größte Teil wird in einen Zirkulator gegeben, welcher das Signal zur Antenne weiterleitet. Die Antenne wird zum Senden und Empfangen genutzt. Das Signal P_2 wird am Objekt reflektiert und wird zum Six-Port geleitet, wo es mit dem Referenzsignal P_1 mit vier verschiedenen, relativen, statischen Phasenverschiebungen überlagert wird. Die Phasenverschiebungen sind Vielfache von $\pi/2$. Es werden vier neue Signale am Ausgang des Six-Ports erhalten [Koe+16]:

$$(4.4) \quad P_3 = 0,5(P_1 + jP_2)$$

$$(4.5) \quad P_4 = 0,5(P_1 - jP_2)$$

$$(4.6) \quad P_5 = 0,5(P_1 + P_2)$$

$$(4.7) \quad P_6 = 0,5(P_1 - P_2)$$

Die Basisband Spannungen B_3, \dots, B_6 können folgendermaßen aus P_3, \dots, P_6 berechnet werden:

$$(4.8) \quad B_3 = |P_3|^2$$

$$(4.9) \quad B_4 = |P_4|^2$$

$$(4.10) \quad B_5 = |P_5|^2$$

$$(4.11) \quad B_6 = |P_6|^2$$

Die vier Basisband Spannungen formen zwei Signalpaare, welche um 180° verschoben sind. Die beiden daraus entstehenden differentiellen Signale haben eine relative Phasenverschiebung von 90° und stellen eine komplexe Zahl dar mit:

$$(4.12) \quad \Re(Z) = B_5 - B_6$$

$$(4.13) \quad \Im(Z) = B_3 - B_4$$

Wenn eine Bewegung des Objekts vorliegt, kann eine variierende Phasenverschiebung zwischen P_1 und P_2 beobachtet werden. Um die relative Phase zwischen den beiden Eingangssignalen zu berechnen, wird ein in-phase and quadrature (IQ)-Verfahren zur Demodulation angewandt:

$$(4.14) \quad \Delta\varphi = \arg(Z) = \arctan \frac{\Im(Z)}{\Re(Z)} = \arctan \frac{B_3 - B_4}{B_5 - B_6}$$

Die relative Distanz zwischen Objekt und Sensor kann über obige Formel 4.3 berechnet werden.

Strahlung Um zu vermeiden, dass die elektromagnetische Strahlung von Radar-Sensoren dem Menschen schadet, geben die International Commission on Non-Ionizing Radiation Protection (ICNIRP), das Institut für Strahlenschutz und die Bundesnetzagentur Grenzwerte für die erlaubte Leistung und Strahlung von Radar-Sensoren vor.

Equivalent Isotropic Radiated Power (EIRP) Die EIRP beschreibt die Sendeleistung einer Antenne in Hauptstrahlrichtung bezogen auf einen isotropen Strahler [Str]. EIRP ist das Produkt aus der eingespeisten Sendeleistung und dem Antennengewinn gegenüber dem isotropen Strahler. Die Bundesnetzagentur legt Grenzwerte für die ausgesendete Leistung bei den jeweiligen Frequenzen fest. Für einen Radar-Sensor mit einer Frequenz von 24 GHz ist eine Leistung von maximal 100 mW bzw. 20 dBm erlaubt [Bun]. Im Vergleich dazu strahlt ein Mobiltelefon meist mit 1 bis 2 W im Frequenzbereich zwischen 450 und 1800 MHz [Bie].

Specific Absorption Rate (SAR) Das ICNIRP gibt Grenzwerte für die spezifische Absorptionsrate vor [ICN]. Dieser Wert beschreibt, wie viel Leistung pro Kilogramm vom Körper aufgenommen wird. Der SAR-Wert lässt sich über folgende Formel berechnen [Mäu]:

$$(4.15) \quad SAR = \frac{1}{2} \cdot \frac{\sigma |E|^2}{\rho}$$

mit σ als elektrische Leitfähigkeit des Gewebes in S/m , ρ als Dichte des Gewebes in kg/m^3 und E als elektrische Feldstärke in V/m .

Je nach Frequenz kann die Strahlung zur schädlichen Erwärmung des Körpers führen. Es wird davon ausgegangen, dass Erwärmungen nur bei Frequenzen unterhalb von 6 GHz auftreten. Es wird ein Grenzwert von 4 W/kg angegeben, der einer Temperaturerhöhung von 1 °C entspricht. Bei Frequenzen über 6 GHz wird keine Erwärmung erwartet (allerdings wurde der Effekt bisher noch nicht untersucht), da die Strahlung nur durch obere Hautschichten absorbiert wird und somit Wärme leicht wieder abgegeben werden kann. [ICN]

Die folgende Tabelle (Tabelle 4.2) beinhaltet die Grenzwerte des Radar-Sensors im Vergleich zu Mobiltelefonen [Bie] und Babyfonen [Exp].

Tab. 4.2: Grenzwerte Strahlung

Gerät	Frequenzbereich	Körper SAR	Lokaler Kopf/Oberkörper SAR
Mobiltelefon	450 - 1800 MHz	0,08 W/kg	2,0 W/kg
Babyfon	1,8 - 2,4 GHz	0,08 W/kg	2,0 W/kg
Radar-Sensor	24 GHz	0,08 W/kg	keine Angabe

Simulation des SAR eines Erwachsenen Für einen Erwachsenen wurde mit dem Simulationsprogramm CST Microwave die absorbierte Strahlung durch den Radar-Sensor berechnet. Der Radar-Sensor wird bei 24 GHz betrieben und ein 15dBi-Horn im Abstand von 20 cm von einer menschlichen Brust angenommen und von einer Sendeleistung von 5 dBm ausgegangen, dann wird ein SAR-Wert von 0,006 W/kg bzw. 0,014 W/kg erhalten, wenn über Massen von 10 g (nach CENELEC, d.h. EU) bzw. 1 g (nach IEEE, d.h. USA) gemittelt wird. Der SAR-Wert für den gesamten Körper wäre noch niedriger. Diese Werte liegen weit unter dem Grenzwert und kommen zustande, da

- der Radar-Sensor bei 24 GHz betrieben wird und durch die Freiraumdämpfung viel Leistung verloren geht (46 dB auf 20 cm)
- wegen der geringen Eindringtiefe nur eine geringe Masse erwärmt wird, was durch die Mittelung (1 g/10 g) dann zu kleinen SAR-Werten führt.

In Abbildung 4.3 ist die Simulation zur Bestimmung des SAR-Wertes eines Erwachsenen zu erkennen.

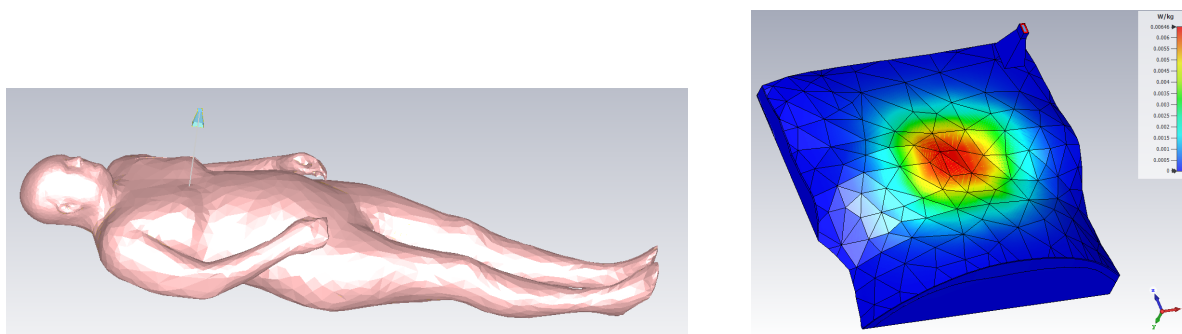


Abb. 4.3: Simulation des SAR-Wertes eines Erwachsenen bei einem Abstand von 20 cm. Der Wert wird über 10 g gemittelt (re.). (Simulationen von Fabian Michler, LTE FAU, erstellt.)

Simulation des SAR eines Neugeborenen Auch für ein Neugeborenes soll der SAR-Wert berechnet werden (siehe Abbildung 4.4). CST Microwave liefert ein Modell für ein Neugeborenes, allerdings nur mit grob gerasterter Haut. Die Haut muss fein gerastert werden, da es insbesondere auf die oberen Hautschichten ankommt. Aus diesem Grund wird erwartet, dass das Simulationsergebnis ungenau ist. Es wurde wieder mit einem Radar-Sensor bei 24 GHz,

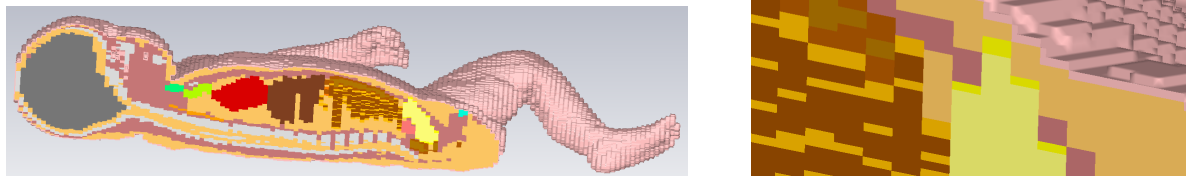


Abb. 4.4: Simulation des SAR-Werts eines Neugeborenen bei einem Abstand von 20 cm und der Detailansicht der Hautschichten (re.). (Simulationen von Fabian Michler, LTE FAU, erstellt.)

einem 15dBi-Horn im Abstand von 20 cm und einer Sendeleistung von 5 dBm gerechnet. Die erhaltenen Werte liegen bei (1 g-Mittelwert) 0,005 W/kg bzw. 0,0003 W/kg (gesamter Körper). Der SAR-Wert für die Brust eines Erwachsenen liegt weit unter dem Grenzwert. Beim Säugling wird über das gesamte Gewicht gerechnet. Aufgrund der hohen Freiraumdämpfung und bei größerer Distanz von 30 bis 40 cm ist davon auszugehen, dass der SAR-Wert bei einem Neugeborenen auch unter dem Grenzwert liegt.

4.3.1.5 Temperatur- und Feuchtigkeits-Sensor

Ein Temperatur- und Feuchtigkeits-Sensor soll für die Bestimmung der Körper- und Umgebungstemperatur integriert werden. Die Genauigkeit eines Fieberthermometers liegt bei 0,1 °C. Die des Sensors muss mindestens genauso hoch sein. Eine weitere wichtige Anforderung ist die der Hygiene. Der Sensor muss geeignet für Wischdesinfektion sein, idealerweise abgedeckt. [Kau22]

4.3.2 Sensorbox Konzept

Der Sensoraufbau soll entwickelt und während einer Machbarkeitsstudie in der Kinderklinik genutzt werden. Die Idee soll an dieser Stelle kurz beschrieben werden: Die gesamte Sensorik soll in ein Gehäuse integriert werden, welches mittels Magneten im Inkubator befestigt wird. Das Gehäuse sollte am besten aus Kunststoff bestehen, damit keine Wärmereflektionen entstehen. Das Material muss für die Kameraobjektive durchsichtig sein und eine Wischdesinfektion erlauben. Die Box stellt somit eine mobile Einheit dar, die in jeden Inkubator einfach eingesetzt werden kann.

4.3.3 Referenzmessung

4.3.3.1 Klinischer Goldstandard

Als klinischer Goldstandard kommen das EKG (für die Überwachung des Herzens und die Impedanz-Pneumographie), die Pulsoxymetrie und eine rektale Temperatursonde zum Einsatz. Die Funktionsweise der Sensoren wurde bereits in Kapitel 2 vorgestellt.

4.3.3.2 Baby-Thorax-Simulator

Der Baby-Thorax-Simulator (vergleiche [Gle+21]) ist insbesondere für Messungen mit dem Radar-Sensor notwendig.

Es wurden zwei mögliche Ansätze betrachtet: ein mechanischer Ansatz (ovale Scheibe) und ein elektro-mechanischer Ansatz (Lautsprecher). Beide Ansätze sollen kurz präsentiert werden.

Mechanischer Ansatz Der mechanische Ansatz simuliert die Atemrate, indem eine Platte durch die Rotation einer ovalen Scheibe nach oben und unten bewegt wird. Die Größe der ovalen Scheibe muss variieren, damit verschiedene Amplituden simuliert werden können. Die Welle wird von einem Motor angetrieben, die Ansteuerung erfolgt über einen Arduino Mikrocontroller. Die Amplitude kann durch Verschiebung der Scheiben angepasst werden. Die Frequenz des simulierten Sinus kann durch Änderung der Geschwindigkeit der Rotation verändert werden. Der Herzschlag wird durch einen Vibrationsmotor, der in die Platte integriert ist, simuliert.

Elektro-mechanischer Ansatz Dieses Konzept (siehe Abbildung 4.5) verwendet einen Lautsprecher, eine Verstärkerschaltung, einen Computer und ein Lasermikrometer. Die Atem- und Herzschlagsignale werden durch ein Software-Programm auf dem Computer generiert. Die Idee ist es, eine Überlagerung von zwei Sinuswellen zu verwenden: eine für die Atmung und ein abfallender Sinus für den Herzschlag. In Abbildung 4.6 kann das Herzschlagsignal (blau) und das Atemsignal überlagert durch den Herzschlag (rot) gesehen werden.

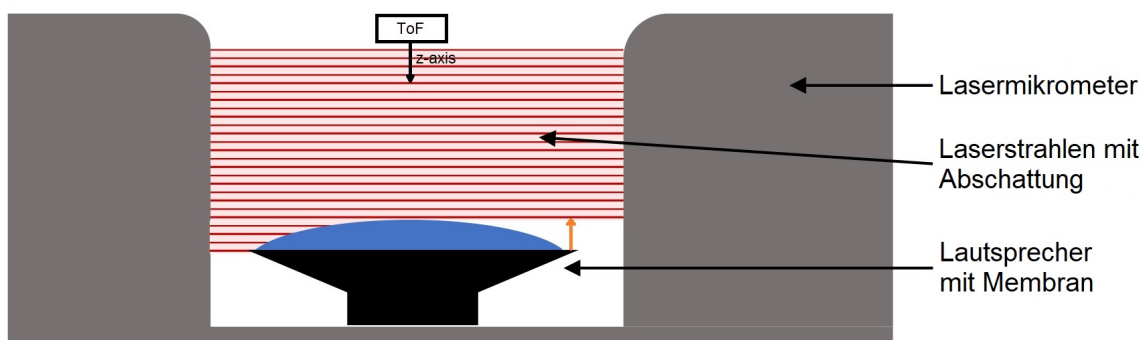


Abb. 4.5: Elektro-mechanischer Ansatz für den Baby-Thorax-Simulator [Gle+]

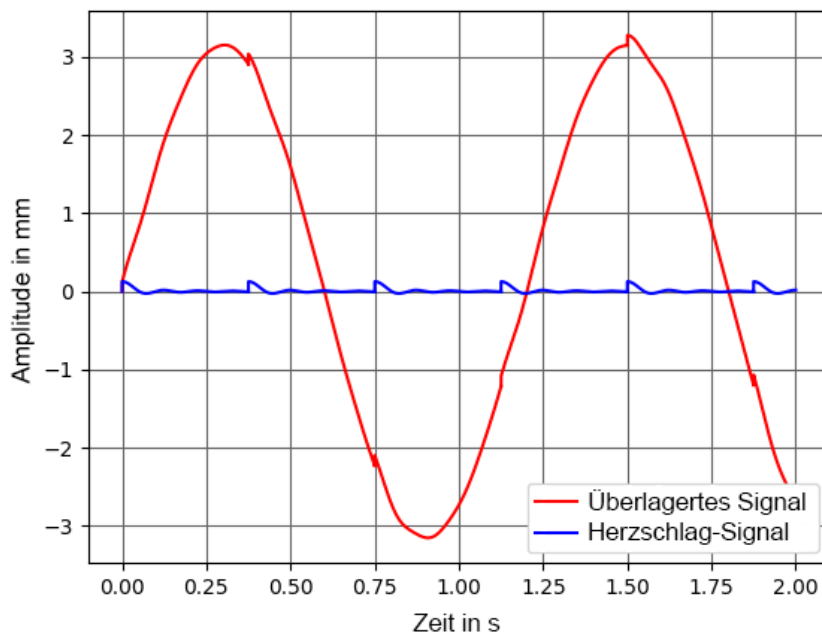


Abb. 4.6: Simulation der Atmung und des Herzschlages durch Überlagerung zweier Sinus-signale [Gle+21]

Da es keine medizinischen Referenzwerte hinsichtlich der Bewegung der Brustwand verursacht durch den Herzschlag von Neonaten gibt, wurde eine Approximation durchgeführt. Die Brustwandbewegung bei Erwachsenen liegt zwischen 0,2 –0,5 mm [RS89] mit Dimensionen des erwachsenen Herzens von 12 cm × 8 cm × 6 cm [Bet+13]. Das Herz eines Neugeborenen hat die Größe einer Walnuss mit Dimensionen von 3,5 cm × 2,5 cm × 2,5 cm [Rei]. Es wird die Bewegung der Herzwand in z-Richtung betrachtet, somit wird angenommen, dass diese bei einem Neugeborenen mit einem Faktor $\frac{2,5 \text{ cm}}{6 \text{ cm}}$ multipliziert wird. Daraus ergibt sich eine approximierte Brustwandbewegung von 0,0834 –0,2085 mm. Der Mittelwert liegt bei 0,146 mm.

4.4 Kalibrierung

Zunächst soll die Theorie hinter der intrinsischen Kalibrierung der Kameras erklärt werden. Anschließend wird das Konzept für die extrinsische Kalibrierung der verschiedenen Sensoren vorgestellt.

4.4.1 Intrinsische Kalibrierung

Zur Rektifizierung werden die tangentialen und radialen Verzerrungskoeffizienten berechnet und entfernt. Voraussetzung hierzu ist, dass die zu kalibrierenden Kameras durch das Lochkameramodell dargestellt werden können.

$$(4.16) \quad \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X/Z \\ Y/Z \\ 1 \end{pmatrix}$$

mit f_x und f_y als Brennweite, u_0 und v_0 als Mittelpunkt des jeweiligen Kamerasensors.

Hierbei muss beachtet werden, dass das Kalibrieremuster aus verschiedenen Positionen, mit unterschiedlichen Abständen und Neigungen aufgenommen wird, damit die Verzerrungskoeffizienten so genau wie möglich berechnet werden können. Die Rektifizierung erfolgt nach folgender Formel:

$$(4.17) \quad \begin{pmatrix} u \cdot (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1 v + p_2 \cdot (r^2 + 2u^2) \\ v \cdot (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1 \cdot (r^2 + 2v^2) + p_2 u \end{pmatrix}$$

mit k_1 , k_2 und k_3 als radiale Verzerrungskoeffizienten und p_1 und p_2 als tangentielle Verzerrungskoeffizienten. r ist die Distanz zwischen u und v .

Intrinsische Kalibrierung Fisheye-Objektiv Bei dieser Methode wird das Fisheye-Kamera-Modell als equidistante Projektion behandelt [stab], die durch folgende Formel beschrieben wird (vgl. [KB06]):

$$(4.18) \quad r = f \theta$$

mit r als Distanz zwischen Bildpunkt und Bildmittelpunkt, f als fokale Länge und θ als Winkel zwischen Hauptachse und einkommendem Strahl. Das Verzerrungs-Modell folgt einer allgemeineren Form:

$$(4.19) \quad r(\theta) = k_1 \theta + k_2 \theta^3 + k_3 \theta^5 + k_4 \theta^7 + k_5 \theta^9 + \dots$$

mit $k_1 - k_5$ als Parameter.

4.4.2 Extrinsische Kalibrierung

Wie in Abschnitt 4.3.1 erwähnt, können durch die Fusion der Daten zweier Sensoren deren Schwachpunkte reduziert werden, was sich besonders im Fall der Kombination von RGB-Kamera mit Thermokamera anbietet. Die höchste Genauigkeit einer direkten Fusion zwischen der Thermo- und der RGB-Kamera ist nur bei der Distanz, für die die Kalibrierung durchgeführt wurde, gewährleistet [SPM10]. Dies ist problematisch, da die Distanz wegen der unterschiedlichen Größen der Neonaten verschieden sein kann. Aus diesem Grund soll eine indirekte Fusionsmethode entwickelt werden, die die Time-of-Flight-Kamera des Sensoraufbaus aus [Gle+21] als dritten Sensor verwendet. Der Vorteil der 3D-ToF-Kamera ist, dass die Tiefe berücksichtigt wird. Die folgenden Schritte werden benötigt, um die indirekte Thermo-RGB-Fusion durchzuführen:

- Detektion der Kreise des Kalibrieremusters im ToF Mono-Bild und der zugehörigen Tiefenpunkte.
- Detektion der Kreise im RGB- und Thermobild.
- Berechnung der Transformation zwischen der RGB- und ToF-Kamera mit Hilfe der Kreismittelpunkte.
- Berechnung der Transformation zwischen der Thermo- und ToF-Kamera mit Hilfe der Kreismittelpunkte.
- Projektion der RGB-Punkte in das Thermo-Kamerabild (an der Position der ToF-Punkte).

4.5 Anforderungen an Algorithmen

Die bisherigen Monitoringgeräte in der Kinderklinik erhalten alle 5 s einen neuen Messwert. Aus diesem Grund müssen die zur Berechnung der Vitalparameter verwendeten Algorithmen mindestens genauso oft einen neuen Wert liefern.

4.5.1 Detektion von Gesichts- und Körperarealen

Die Gesichtsdetektion durch ein Neuronales Netz darf die Generierung von Messwerten nicht verzögern.

ROS-Implementierung Das gesamte System soll unter ROS laufen, weil es die Kommunikation zwischen den einzelnen Sensoren vereinfacht. Somit müssen die Neuronalen Netze in ROS integriert werden, bzw. ein ROS-Interface geschaffen werden.

4.5.2 Kontaktlose Atemdetektion

Die Atemraten von Neu- und Frühgeborenen liegen zwischen 25 und 60 BPM. Zur Sicherheit werden Werte zwischen 5 und 80 BPM betrachtet. Daraus folgt, dass die Abtastrate laut Abtasttheorem mindestens bei zweimal der höchsten Frequenz 2,67 Hz liegen muss.

4.5.3 Kontaktlose Herzraterdetektion

Der Normbereich der Herzrate liegt für Neu- und Frühgeborene zwischen 80 und 220 BPM. Die Abtastrate muss laut Nyquist-Theorem daher mindestens bei 7,34 Hz liegen.

4.5.4 Kontaktlose Temperaturdetektion

Die erforderliche Tempurauflösung sollte größer gleich 0,1 °C sein, entsprechend eines Standard-Fieberthermometers.

4.6 Datengewinnung

Zum Trainieren der verwendeten Neuronalen Netze und zur Validierung des Gesamtsystems werden in ausreichendem Maße reale Daten von Neonaten benötigt.

Ethikantrag und Machbarkeitsstudie Um Daten am Menschen erheben zu dürfen, wird eine Erlaubnis der zuständigen Ethikkommission benötigt. Hierfür wurde ein Ethikantrag für eine Machbarkeitsstudie gestellt (Protokoll Code: 197_20B und Datum der Genehmigung: 28. April 2021, Änderungsprotokoll Code: 20-197_1-B, genehmigt am 22. Dezember 2022).

Im Ethikantrag muss genau angegeben werden, welche Art von Untersuchungen durchgeführt werden sollen, was der Stand der Technik ist, welche Risiken auftreten können und wie mit den akquirierten Daten umgegangen werden soll. Außerdem muss die Studienpopulation, der Studienablauf und die Biometrie festgelegt werden.

Geplant war der Einschluss von 30 reifgeborenen Neonaten mit Gelbsucht, die ansonsten gesund sind. Dies wurde später auf kardio-respiratorisch stabile Frühgeborene ab einem Alter von 34 SSW erweitert. Die Machbarkeitsstudie sollte parallel zur Phototherapie im Inkubator durchgeführt werden. Als Sensoren sollen RGB-, Thermo- und Time-of-Flight-Kameras eingesetzt werden, sowie ein Feuchtigkeits- und Temperatursensor. Als Referenzwerte werden EKG (für Herz- und Atemrate (IPG)), Pulsoxymetrie und eine rektale Temperatursonde verwendet.

5 Entwurf

In diesem Kapitel wird beschrieben, wie das Sensorsystem physisch aufgebaut ist, welche Sensoren zum Einsatz kommen, wie diese kalibriert werden und welche Algorithmen entwickelt wurden, um schließlich die Vitalparameter zu berechnen.

5.1 Sensorkonfiguration

5.1.1 RGB-Kamera

Die eingesetzte RGB-Kamera ist die JAI GO-2400C-USB Kamera (siehe Abbildung 5.1). Diese USB-Vision Kamera wird mit einem Fish-Eye Objektiv von Sony (FUJINON FE185C086HA-1) ausgestattet, damit das FOV groß genug ist. Die RGB-Kamera hat einen Öffnungswinkel von $185^\circ \times 185^\circ$ (Fisheye). Die resultierende Bildgröße liegt bei 1146×716 Pixel. Der verwendete Kameratreiber ist das ROS-Paket `ros_cvb_camera_driver` [HG]. Zu Beginn war deutlich zu erkennen, dass die Bilder einen Grünstich haben. Dieser lässt sich durch ein Whitebalancing entfernen.



Abb. 5.1: Kamerasensoren: Thermokamera Optris PI640 (li.) und JAI GO-2400C-USB Kamera (re.) Quelle: Fa. Optris und Fa. JAI

5.1.2 Thermokamera

Die eingesetzte Thermokamera ist die Infrarotkamera Optris PI 640 mit einem Öffnungswinkel von $90^\circ \times 64^\circ$ (siehe Abbildung 5.1 links). Der passende Öffnungswinkel wurde entsprechend der folgenden Formeln berechnet. Es wird von einer Matratzengröße von $49,3 \times 78,7$ cm ausgegangen und einem Abstand von 38 cm zwischen Matratze und Inkubatordeckel.

$$(5.1) \quad 2 \cdot \arctan \left(0.5 \cdot \frac{78,7\text{cm}}{38\text{cm}} \right) = 92^\circ$$

$$(5.2) \quad 2 \cdot \arctan \left(0.5 \cdot \frac{49,3\text{cm}}{38\text{cm}} \right) = 64^\circ$$

Die Kamera hat eine optische Auflösung von 640 x 480 Pixeln und arbeitet in einem Spektralbereich von 7,5 μm und 13 μm . Die thermische Empfindlichkeit liegt bei 75 mK. Die Systemgenauigkeit beträgt $\pm 2^\circ\text{C}$ bei einer Inkubatortemperatur von größer 28 $^\circ\text{C}$.

Das ROS-Paket *optris_drivers* wird zur Kamera-Ansteuerung verwendet. Damit die Rohdaten fehlerfrei angezeigt werden können, muss die Nodrop-Option gesetzt werden.

Zur korrekten Ausrichtung der Thermokamera im Sensoraufbau wurden Testaufnahmen mit einer Heizmatte gemacht. Die senkrechte Ausrichtung ist besser geeignet und liefert eine größere FOV.

Außerdem muss beachtet werden, dass vor dem Objektiv der Thermokamera keine störenden Materialien vorhanden sind. Wird die Kamera außerhalb des Inkubator-Deckels befestigt, führt dies zu Reflektionen.

5.1.3 Mikrowellen-interferometrischer Radarsensor

Der einzusetzende Radarsensor ist der iSYS-4001 der Firma InnoSens (siehe Abbildung 5.2 links), der speziell für das GUARDIAN-Projekt [BMB] gefertigt wurde. Der Sensor hat kein CE-Zeichen, aber eine Konformitätserklärung und wurde von InnoSens auf EMV getestet. Das Radarmodul ist kein Six-Port, liefert aber ähnliche Ausgangssignale, sodass dieselbe Verarbeitung wie für ein Six-Port verwendet werden kann. Er hat eine Leistung von 100 mW (20 dBm), die gemäß ICNIRP Vorgaben (siehe Kapitel 2) im erlaubten Normbereich liegt.

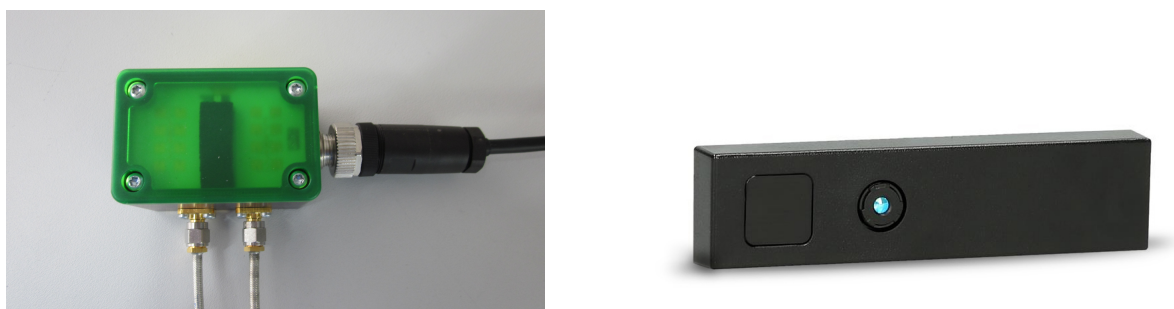


Abb. 5.2: Radarsensor von Fa. InnoSenT (li.) und Time of Flight Kamera CamBoard pico flex von Firma PMD (re.)

Die Radarsensordaten werden über eine Digitalisierungs-Platine ausgelesen. Die Platine ist über Ethernet mit einem Rechner verbunden. Über User Datagram Protocol (UDP) werden Befehle zur Verbindung und Datenübertragung gesendet. Hierfür wird in Matlab auf ein Mex-File zugegriffen, welches die Befehle zur Ansteuerung des Mikrocontrollers beinhaltet. Da das Mex-File nur für Windows bereit stand, musste zunächst ein Mex-File unter Linux generiert

werden. Anstelle von *windows.h* und *winsock.h* müssen folgende Header verwendet werden: *unistd.h*, *sys/types.h*, *sys/time.h*, *sys/socket.h*, *sys/select.h*, *errno.h*, *arpa/inet.h*, *sys/ioctl.h*. Die relevanten Windows-Funktionen wurden durch die jeweiligen Linux-Funktionen ersetzt. In einem ersten Versuch wurden mit Hilfe eines Matlab-Skriptes die Ausgangswerte der einzelnen Kanäle (siehe ?? li.), ein I-Q-Diagramm (siehe ?? re.), sowie der Distanzverlauf (siehe Abbildung 5.5) aufgezeichnet, wenn eine Hand auf den Sensor zu- und wegbewegt wird. Die Distanz lässt sich aus den I- und Q-Kanalwerten über folgende Formel berechnen:

$$(5.3) \quad \Delta x = \frac{\Delta \varphi}{2 \cdot \pi} \cdot \frac{\lambda}{2}$$

$$(5.4) \quad \Delta x = \arctan \left(\frac{I}{Q} \cdot \frac{\lambda}{4 \cdot \pi} \right)$$

mit Δx als relative Distanz, $\Delta \varphi$ als relative Phasenverschiebung und λ als Wellenlänge.

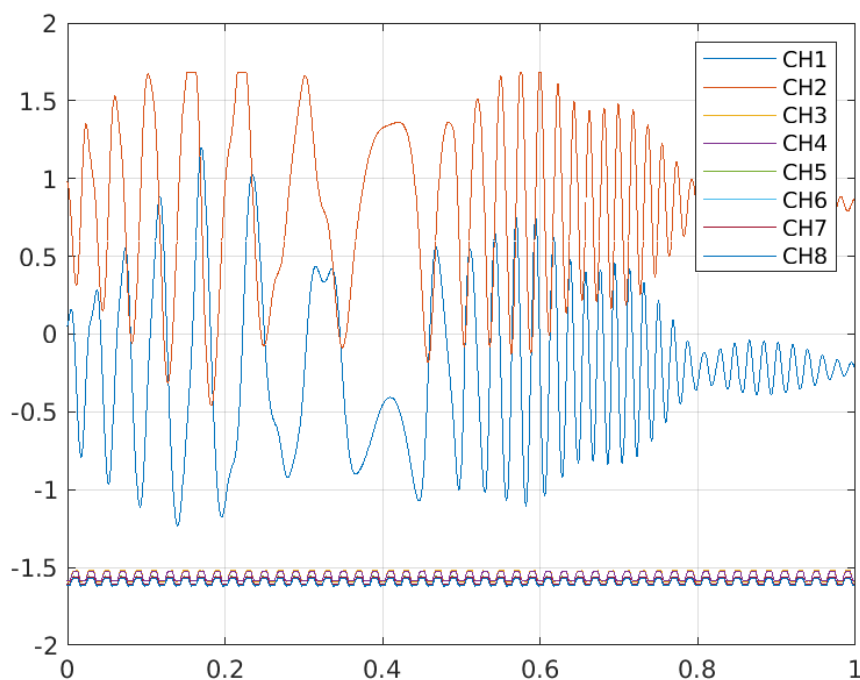


Abb. 5.3: Ausgangskanäle der einzelnen Ausgänge

Das I-Q-Diagramm zeigt mehrere Ellipsen. Damit der Sensor saubere Daten liefert, muss zunächst ein Ellipsen-Fitting durchgeführt werden. Hierfür bietet sich ein Minimal-Least-Square-Fitting an. Dies wurde im Rahmen einer Masterarbeit implementiert [Her21].

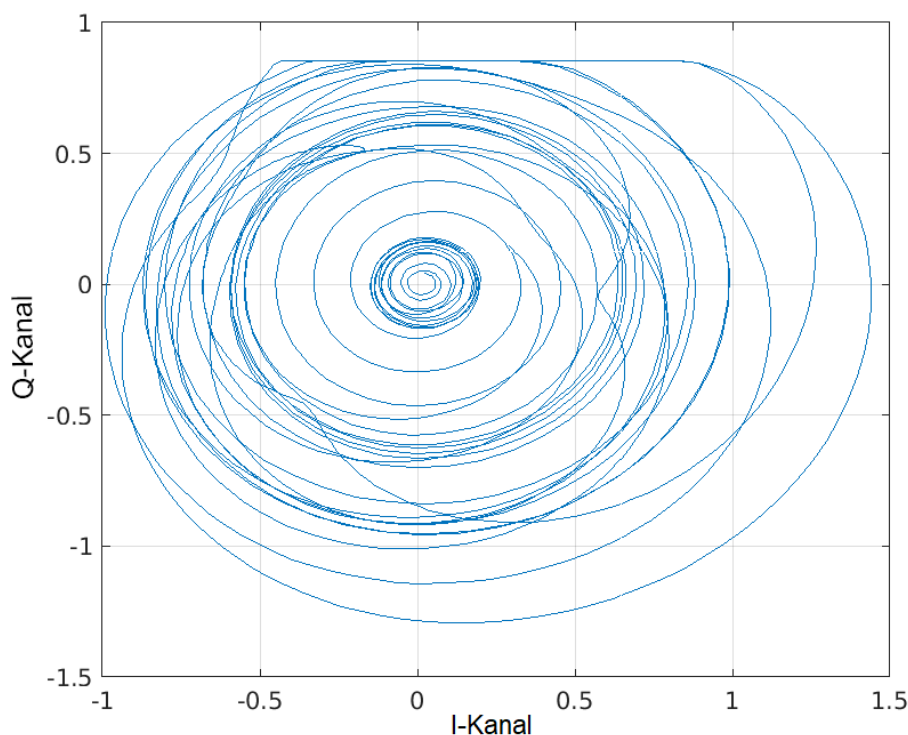


Abb. 5.4: I-Q-Diagramm

5.1.4 Tiefenkameras

5.1.4.1 Structured Light Kamera

Die zum Test eingesetzte Structured Light Kamera ist die Astra Orbecc Stereo. Sie hat eine Tiefenauflösung von 640 x 400 Punkten und Tiefen FOV von $H67,9^\circ$ $V45,3^\circ$ $D78^\circ \pm 3,0^\circ$. Die Kamera kann bei Temperaturen zwischen 10 und 40 ° C betrieben werden und ist somit für die Anwendung im Inkubator geeignet. Der Kameratreiber ist das ROS-Paket *ros_astra_camera*.

5.1.4.2 ToF Kamera

Camboard Pico Flexx Die Time of Flight Kamera Camboard Pico Flexx (siehe Abbildung 5.2) ersetzt die Structured Light Kamera Astra Orbecc Stereo S einen Mindestabstand von 30 cm benötigt, welcher im Inkubator nicht garantiert werden kann. Die CamBoard pico flexx arbeitet bereits ab 10 cm Abstand. Die Kamera hat einen Messbereich von 0,1-4 m, ein FOV von $62^\circ \times 45^\circ$ und eine Auflösung von 224 x 171 (38k) Pixeln. Eine Framerate von bis zu 45 fps (3D Frames) ist gegeben. Der verwendete Treiber ist die Royale SDK, die vom Hersteller zur Verfügung gestellt wird und das ROS-Paket *pico_flexx_driver* [Art]. In dieser Arbeit wird eine Framerate von 35 fps eingestellt. Die Kamera verwendet das Phasenverschiebungsverfahren. Im Aufbau wird diese Kamera über dem Gesicht des Kindes positioniert.

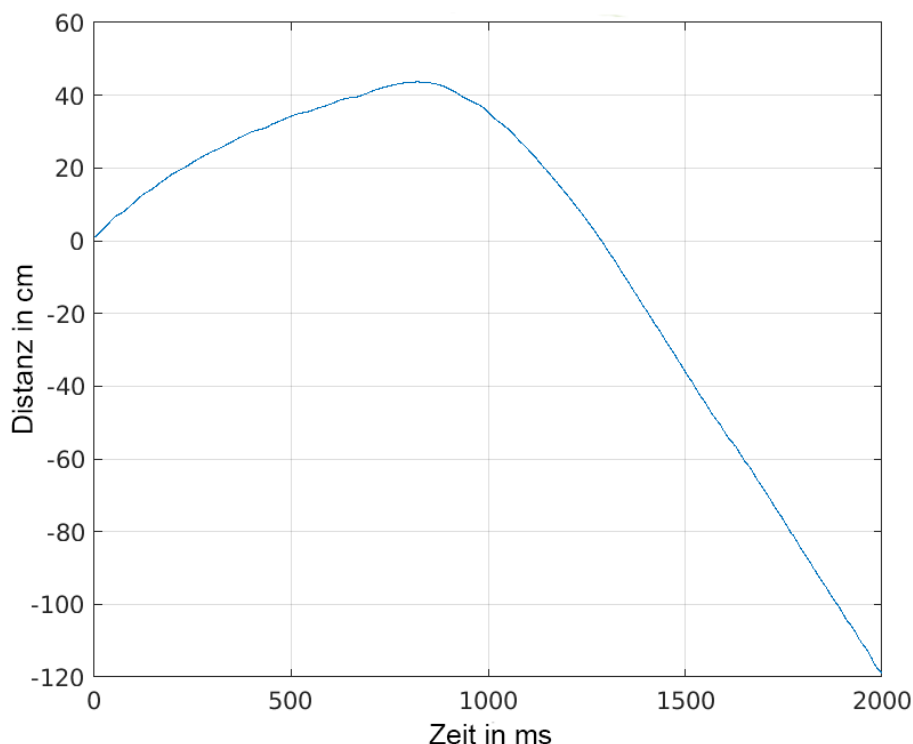


Abb. 5.5: Distanzverlauf über die Zeit, berechnet aus den I-Q-Signalen.

Flexx2 ToF-Kamera Es wurde eine zweite ToF-Kamera, die Flexx2 von Fa. pmd, über dem Torso des Kindes integriert. Die Kameraeigenschaften unterscheiden sich nur leicht von der Pico Flexx. Es liegt eine Auflösung von 224 x 172 Tiefen-Pixeln vor. Der Öffnungswinkel beträgt $56^\circ \times 44^\circ$. Das ROS-Paket *pico_flexx_driver* kann weiterhin verwendet werden, nur muss eine neuere Version der royale SDK (libroyale 4.21.0) eingebunden werden. Als Use Case wird `MODE_5_30FPS` verwendet.

5.1.5 Temperatur- und Feuchtigkeits-Sensor

Die Integration des Temperatur- und Feuchtigkeits-Sensors wurde im Rahmen einer Bachelorarbeit bearbeitet [Kau22]. Hierbei kommt der Sensor PHDC3022DEJT von Fa. Texas Instruments zum Einsatz, der auf einem kapazitivem Messverfahren beruht. Der Feuchtigkeits-Sensor hat einen Arbeitsbereich von 0 % bis 100 %. Die Genauigkeit liegt bei $\pm 1,5$ %. Der Temperatursensor hat einen Arbeitsbereich zwischen -40 °C bis 125 °C und eine Genauigkeit von $\pm 0,1$ °C. [Ins] Es wird mindestens diese Genauigkeit benötigt, da diese auch bei einem typischen Fieberthermometer vorliegen würde. Da die minimale Temperatur im Inkubator bei 35 °C und die maximale bei $38,5$ °C liegt, ist der Sensor gut geeignet. Zur Ansteuerung des Sensors wird der Mikrocontroller NUCLEO G431KB Nucleo-32 von Fa. ST verwendet. Dieser ist ROS-kompatibel und von seinen Maßen gut für den Einbau in die Sensorbox geeignet. Abbildung 5.6 zeigt die Kommunikationsstruktur zwischen ROS und dem Mikrocontroller. Folgende Messwerte werden unter dem Topic mit dem Message-Typ *sensor_data* publiziert:

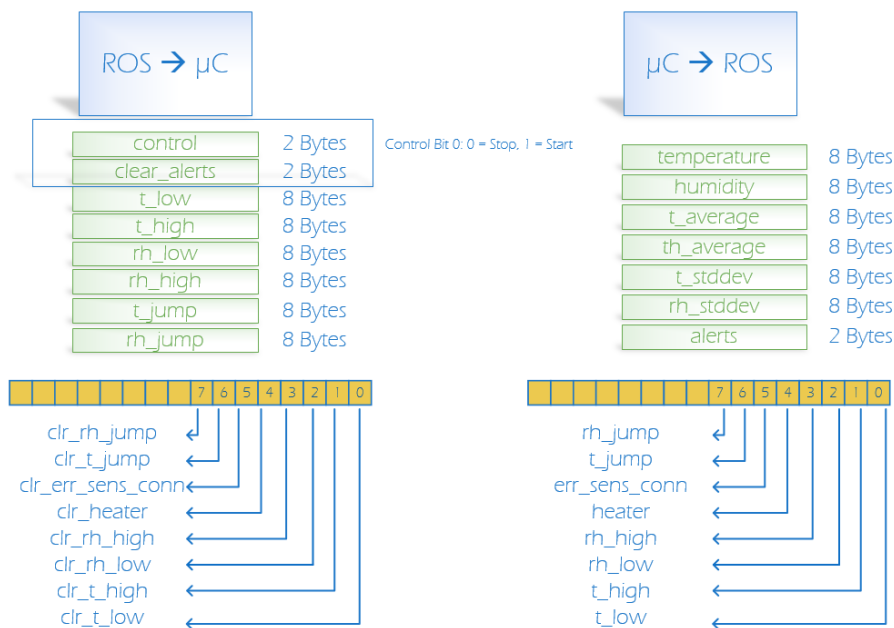


Abb. 5.6: Kommunikationsstruktur zwischen Temperatur- bzw. Feuchtigkeitssensor und Mikrocontroller und ROS [Kau22]

t_average als Mittelwert der letzten zehn Temperatur-Messwerte, *rh_average* als Mittelwert der Feuchtigkeitswerte, *t_stddev* als Standardabweichung der letzten zehn Temperatur-Messwerte, *rh_stddev* als Standardabweichung der Feuchtigkeitswerte und verschiedenen Alarmmeldungen. *rh_jump_alert* und *t_jump_alert* werden aktiviert, wenn ein Sprung der Feuchtigkeit oder der Temperatur vorliegt, wenn beispielsweise die Inkubatorklappe geöffnet wurde. *err_sens_con_alert* zeigt Verbindungsfehler des Sensors an. Der Sensor hat die Funktion, dass bei zu hoher Luftfeuchtigkeit eine Heizung aktiviert wird. In diesem Fall wird der Alarm *heater_alert* gesendet. Währenddessen kann der Sensor keine genauen Werte messen. *rh_high_alert*, *rh_low_alert*, *t_high_alert* und *t_low_alert* erscheinen bei zu hohen bzw. niedrigen Temperatur- und Feuchtigkeitswerten.

5.1.6 Aufbau der Sensorbox

Für den Sensoraufbau, der direkt in den Inkubator integriert werden soll, wird ein Universalgehäuse eingesetzt, welches angepasst werden muss.

Bei einem Test mit der Thermokamera fiel auf, dass das transparente Gehäuse genau wie der Inkubatordeckel nicht wärmedurchlässig ist und die Wärme ähnlich wie Glas reflektiert. Durch Tests verschiedener Materialien wie Polyethylen-Platte, Malerfolie aus low density Polyethylen, Polypropylen Prospekthülle, Folien aus Polyethylenterephthalat bzw. Celluloseacetat, Chalkogenid-Glas und anderen wurde ein geeigneter Schutz für die Thermokamera gewählt: Zinksulfid-Glas (Fa. Edmund Optics) ist infrarot durchlässig und ist nicht Ursache von Reflexionen. Dieses Glas wurde in das bestehende Kameragehäuse eingesetzt und integriert. Wichtig war hierbei, dass möglichst kein Abstand zwischen dem Kameraobjektiv und dem Glas besteht. Das Zinksulfid-Glas für die Thermokamera darf nicht mit Kleber an der Sensorbox befestigt

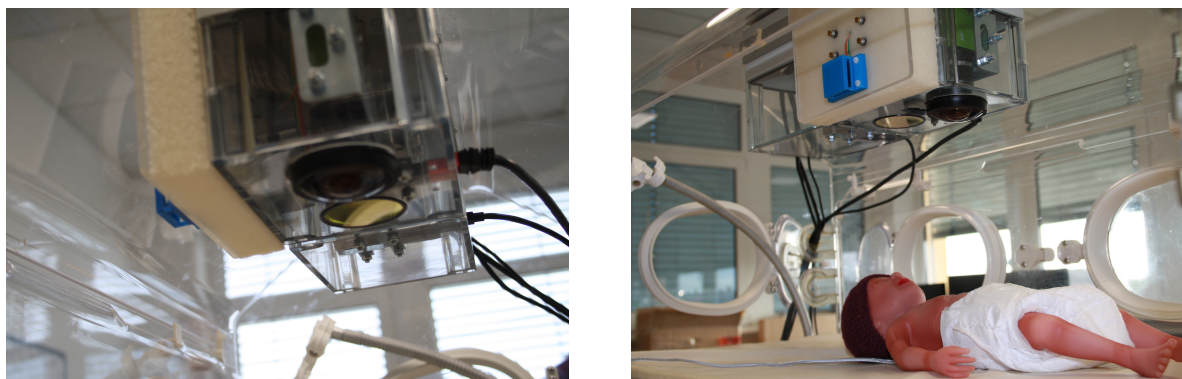


Abb. 5.7: Gesamter Sensoraufbau (li. Ansicht von unten) mit Simulator-Puppe Paul (Fa. Sim-Characters) als Messobjekt (re.). Temperatursensor (vorne blau) mit Isolierung. Die flexx2 ToF-Kamera ist in der Sensorbox ganz links angebracht.

werden. Laut Hersteller könnte beim Einsatz von Kleber das Glas möglicherweise sogar brechen, wenn zu starke Spannungen auftreten. Außerdem kann der Kleber (enthält Polyurethan) zu weiteren Absorptionspeaks im Spektrum führen. Deswegen wurde das Glas in eine Filterfassung geklemmt, welche dann in das Kunststoffgehäuse eingesetzt werden konnte.

Wie zuvor bei der Thermokamera wurde nun auch mit der RGB-Kamera das Verhalten bei Blick durch das durchsichtige Kunststoffgehäuse untersucht. Auch hier treten Reflexionen und Spiegelungen auf. Das Zinksulfid-Glas ist an dieser Stelle nicht geeignet. Deswegen ist das Objektiv der RGB-Kamera ungeschützt eingebaut.

Weitere Anpassungen des Kameragehäuses Alle nötigen Kabeldurchgänge wurden hergestellt. Neodym-Magnete auf der Rückseite des Gehäuses in Verbindung mit einer Stahlplatte an der Oberseite des Inkubators halten das Gehäuse. Es wurde verifiziert, dass die auf der neonatologischen Intensivstation eingesetzten Katheter und Elektroden nicht mit den Magneten interagieren. Die Magnethalterung wurde für ein Gewicht von 1556 g getestet, das entspricht dem 3-fachen Gewicht des Gehäuses und allen Sensoren (als Sicherheitstoleranz).

Wischdesinfektion Zur Wischdesinfektion wird auf der Neonatologischen Intensivstation generell Biguanid verwendet. Für Bildschirme Mikrocid. Das Zinksulfid-Glas darf nur mit Aceton gereinigt werden. Das freiliegende Objektiv der RGB-Kamera kann auf jeden Fall mit alkoholhaltigen Mitteln gereinigt werden, soll aber nicht zu nass werden, da es nicht wasserfest bzw. spritzwassergeschützt ist.

Isolierung für Temperatursensor Es wurde eine Isolierung zwischen Sensorbox und Temperatursensor benötigt, da ansonsten zu viel Wärme abstrahlt. Somit konnten die genauen Korrekturwerte bestimmt werden, um die Inkubatoroberfläche 10 cm über dem Inkubatorboden zu ermitteln (dort wird auch die Temperatur vom Inkubator selbst bestimmt). [Kau22]

5.1.7 Systemrechner

Die Anforderungen an den System- und Trainingsrechner sind folgende: Der Rechner muss für den Einsatz der Titan Xp Grafikkarte geeignet sein, also PCI x16 Slots haben und eine ausreichend hohe Leistung. Die Graphics Processing Unit (GPU) Titan Xp von NVIDIA ist für Deep Learning geeignet und benötigt alleine schon 250 W. Der Rechner Precision 5820 Tower von Dell erfüllt diese Voraussetzungen und wird daher als Systemrechner eingesetzt, sein Netzteil kann eine Leistung von 950 W zur Verfügung stellen. Als Trainingsrechner wird der OMEN HP Computer mit einer GeForce GTX 1080 Ti Grafikkarte von NVIDIA (NVIDIA, Santa Clara, CA, USA) (11 GB Memory) verwendet.

Programme Auf den Rechnern wurde das Betriebssystem Ubuntu 18.04 aufgesetzt. Folgende Programme wurden installiert: NVIDIA Treiber nvidia-410, CUDA 10.0, cuDNN 7.6.3, Python 3, bazel 0.15.0, Tensorflow 1.12.0, Keras, ROS Melodic.

5.1.8 Referenzmessung

5.1.8.1 Klinischer Goldstandard

Als klinischer Goldstandard zur Erfassung der Herz- und Atemrate, der Sauerstoffsättigung und der Temperatur gilt die Messung mittels EKG, Pulsoxymetrie und rektaler Temperatursonde (siehe Abbildung 5.8). Die Sensoren sind an den IntelliVue MX500 Patienten Monitor von Firma Philips angeschlossen.

EKG Das im Monitor integrierte EKG wird zur Überwachung des Herzens und der Atmung (Impedanz-Pneumographie) verwendet. Der Messbereich der Herzfrequenz für Neugeborene liegt zwischen 15 und 350 BPM. Der Hersteller weist darauf hin, dass bei Herzfrequenzen von weniger als 15 Schlägen ein Herzratenwert von 0 angezeigt wird. Es wird eine Genauigkeit von ± 1 % des Messbereichs angegeben. Es kommen drei verschiedene Verfahren zur Herzratenmittelung zum Einsatz: Normalerweise wird die Herzrate durch Mittelung der letzten 12 R-R-Intervalle bestimmt. Bei VES-Salven (mindestens drei aufeinanderfolgende Ventrikuläre Extrasystolen) werden maximal 8 R-R-Intervalle gemittelt. Wenn die Frequenz unter 50 BPM liegt, werden nur die 4 letzten R-R-Intervalle gemittelt. Bei Neugeborenen liegt der Messbereich der Atemrate zwischen 0 und 170 Atemzügen pro Minute. Es wird eine Genauigkeit von ± 1 BPM für Atemraten zwischen 0 und 120 BPM erwartet. Zwischen 120 und 170 BPM wird mit einer Abweichung von ± 2 BPM gerechnet. Die verwendeten EKG-Elektroden sind die Radiolucent NeoLead bzw. Mini NeoLead EKG-Elektroden von Firma NeoTech.

Pulsoxymetrie Es kommen Masimo LNCS Neo-L SpO₂ Klebesensoren für Neugeborene zum Einsatz. Die Messgenauigkeit liegt bei ± 3 %.

Rektale Temperatursonde Die rektale Temperatursonde z.B. von Firma Connect Medizintechnik hat eine Messgenauigkeit von $\pm 0,1$ K im Bereich von 25°C - 45°C.



Abb. 5.8: Sensoren für die Referenzmessungen: EKG-Klebeelektroden (re.), Pulsoxymeter (mi.) und rektale Temperatursonde (li.)

Um zu prüfen, ob die bestehenden Sensoren für die Referenzmessungen Magnetismus aufweisen, wurden eine rektale Temperatursonde, EKG-Elektroden und der Sensor für das Pulsoxymeter einem Neodym-Magneten genähert. Keines der Materialien zeigte ein magnetisches Verhalten oder wurde vom Magneten angezogen. Somit kann die Befestigung des Sensoraufbaus ohne Vorbehalte mit Hilfe von Magneten ausgeführt werden.

Die Referenzmessdaten können direkt am Monitoring-System von Philips über den MIB-Ausgang exportiert werden. Hierfür wird das Tool *ixtrend* der Firma *ixitos* und ein RJ45USB-Kabel benötigt. Das Tool kann direkt auf einer Virtuellen Maschine (Windows) auf dem Systemrechner laufen und vereinfacht das Aufnehmen, Visualisieren und Exportieren der Daten. Da die virtuelle Maschine und der Linuxrechner dieselbe Systemzeit haben, ist eine Zeitsynchronisierung automatisch gegeben. Das Tool *ixtrend* nutzt die Rechnerzeit.

5.1.8.2 SimBaby als Referenz für Structured light camera

Beim SimBaby-Simulator der Fa. Laerdal kann die Atemfrequenz und Atemtiefe eingestellt werden. Der Simulator in der Größe eines 9 Monate alten Säuglings wird als Referenzsystem für die Structured Light Kamera verwendet [GNK20].

5.1.8.3 Simulationspuppe Paul

Der Frühgeborenen-Simulator Paul wurde von SIMCharacters entwickelt und ist laut Hersteller die detailtreueste Simulation eines Frühgeborenen, die derzeit auf dem Markt verfügbar ist [SIM]. Der Simulator hat die Größe eines 27 + 3 Wochen alten Frühgeborenen und wird für das Training von Ärzten, Hebammen und Pflegemitarbeitern genutzt. Physiologische Parameter wie das Gewicht, die Größe und innere Organe sind der Realität nachgeahmt. Eine Funktion des Simulators ist die Atmung. Sie beinhaltet normale Atmung zwischen 0 und 100 BPM sowie pathologische Atmung, wie sie z.B. durch einen Pneumothorax verursacht wird. Das respiratorische System des Simulators besteht aus zwei Testlungen, die an ein Ventilationsystem

angeschlossen sind und kann parametrierbar werden. Auch das Koronar-System kann parametrierbar werden, um pathologisches Verhalten zu simulieren. Der Puls kann nur unter Verwendung von Elektroden gemessen werden. Somit ist es unmöglich, die Herzrate des Simulators mit dem Radarsensor zu detektieren.

5.1.8.4 Baby-Thorax-Simulator

Der Testaufbau für die ToF-Radar-Synchronisierung besteht aus folgenden Komponenten (siehe Abbildung 5.9): Baby-Thorax-Simulator, Laser-Mikrometer als Goldstandard, Pico Flexx ToF-Kamera und Radarsensor-Modul von InnoSenT.

Die Kamera wurde so angebracht, dass die Distanz zum Simulator 23,5 cm beträgt. Der Radarsensor wurde auf derselben Seite parallel zur ToF-Kamera gesetzt, damit der Simulator noch im Messfeld liegt. Um das starke Rauschen und Lichtstreuung bei der ToF-Kamera zu minimieren, wurden Teile des Laser-Mikrometers und des Simulatorsystems mit weißem Papier abgeklebt. Der Baby-Thorax-Simulator und die Sensoren sind mit einem Linux Rechner, auf dem Ubuntu 16.04 installiert ist, über USB bzw. Ethernet verbunden. Das Laser-Mikrometer wird mit einem zusätzlichen Laptop (Windows 10 erforderlich) per Ethernet verbunden. Die Referenzdaten werden als csv-Datei gespeichert und können nachträglich mit den Radar- bzw. ToF-Daten synchronisiert werden. Alle Datentypen werden mit einem Zeitstempel versehen.

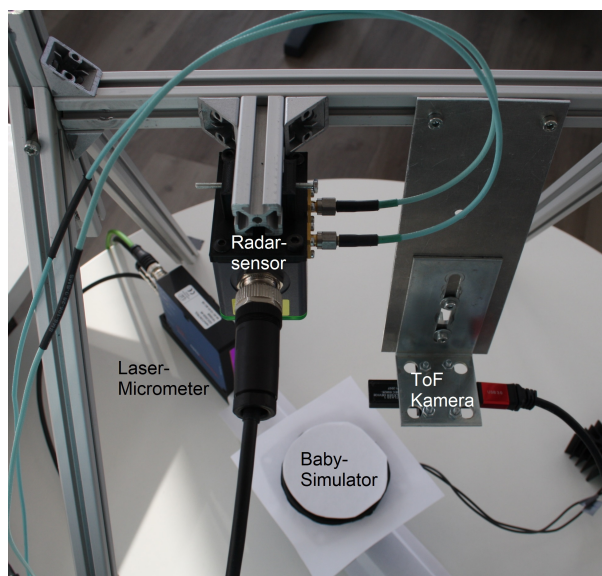


Abb. 5.9: Finaler Testaufbau mit der ToF-Kamera und dem Radarsensor. Der Baby-Thorax-Simulator ist unten im Bild zu sehen [Gle+21]. Das Laser-Mikrometer dient als Goldstandard.

Das Laser-Mikrometer, das für die Referenzmessung eingesetzt wird, ist das optoCONTROL 2520 von MICRO-EPSILON [Mic]. Es hat einen Messbereich von 46 mm und eine maximale Auflösung von 1 μm . Der Laser hat die Laserklasse 1M, was bedeutet, dass keine Gefahr für den Anwender besteht, solange die Laserstrahlen nicht durch ein optisches Instrument verstärkt werden. Empfänger und Sender des Laser-Mikrometers sind mit einem CE2520-1 Kabel

verbunden. Außerdem wird eine 24-V-Spannungsversorgung benötigt. Das optoCONTROL ist per Ethernet an den Rechner angeschlossen. MICRO-EPSILON liefert als Messtool das ODC2520 DAQ Tool mit, das im Modus "Edge light-dark" betrieben wird.

5.2 Kalibrierung

Damit die verschiedenen Kameras korrekte Daten liefern und miteinander fusioniert werden können, muss sowohl eine intrinsische als auch eine extrinsische Kalibrierung durchgeführt werden. Die intrinsische Kalibrierung rektifiziert das Bild, bei der extrinsischen Kalibrierung wird die Transformation zwischen den Sensoren berechnet.

5.2.1 Intrinsische Kalibrierung

5.2.1.1 Intrinsische Kalibrierung der RGB-Kamera

Die JAI Kamera wurde mit dem ROS-Paket *camera_calibration* kalibriert [Oca]. Dort gibt es insbesondere für Kameras mit einer Fish-Eye-Optik ein Kalibrierprogramm. Diesem Programm wird ein bereits zugeschnittenes Bild (1146 x 716) ohne schwarze Ränder übergeben, damit das gesamte Bild genutzt werden kann.

Zur Kalibrierung wird ein Schachbrettmuster mit 5 x 9 Kalibrierpunkten genutzt. Die Quadrate haben eine Größe von 14 x 14 mm. Über den Befehl

```
roslaunch camera_calibration_fisheye cameracalibrator.py -size 5x9 -square 0.014 image:=/  
my_camera/image camera:=/my_camera
```

wird die Kalibrierung gestartet. Mit dem Tool müssen Bilder aus verschiedenen Richtungen und unterschiedlichen Neigungen und Distanzen aufgenommen werden, damit Verzerrungen herausgerechnet werden können.

In Abbildung 5.10 ist auf der linken Seite das unkalibrierte Bild zu sehen. Rechts ist die intrinsische Kalibrierung mit dem zugeschnittenen RGB-Kamerabild unter der Verwendung von 196 Bildern zu erkennen. Wie zu sehen ist, ist der Kisseneffekt noch nicht komplett entfernt.



Abb. 5.10: Intrinsische Kalibrierung mit dem zugeschnittenen RGB-Kamerabild unter der Verwendung von 196 Bildern (rechts). Das Kalibriermuster ist ein 5 x 9 Schachbrettmuster mit Quadraten der Seitenlänge 14 mm. Unkalibriertes Bild (links) [Gle+23].

Manuelle Anpassung der Kalibrierung Im letzten Schritt wurden die Werte der bisher besten Kalibrierung (zugeschnittenes Bild auf 1146 x 716 mit 5 x 9 14 mm Kalibriermuster, Originalbild siehe Abbildung 5.11 (a)) manuell angepasst, sodass der Kisseneffekt weiter abnimmt. Es wurde empirisch ermittelt, wie sich das Bild bei Änderung der Projektionsmatrix sowie der Verzerrungs-Matrix verhält. Die schrittweise Anpassung ist in Abbildung 5.11 und Abbildung 5.12 zu sehen.

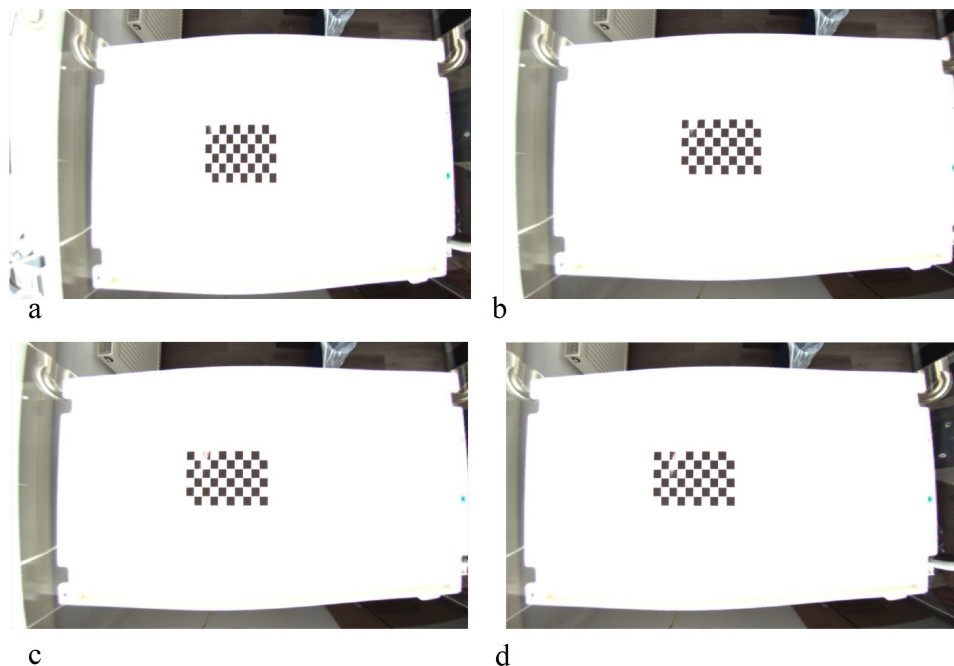


Abb. 5.11: Empirische Anpassung der intrinsischen Kalibrierung der RGB-Kamera.

- (a) Originalbild;
- (b) Änderung des ersten Werts der Projektions-Matrix von 337,75012 zu 400;
- (c) Änderung des ersten Werts der Projektions-Matrix von 337,75012 zu 410;
- (d) Änderung des zweiten Werts der Projektions-Matrix von 567,37183 zu 500.
[Gle+23]

Die finale Kalibrierung ergibt sich somit zu:

$$(5.5) \quad \begin{pmatrix} f_x & 0 & f_y \\ 0 & u_0 & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

mit $f_x = 519,17167$ und $f_y = 521,25487$ als Brennweite,
 $u_0 = 565,71198$ und $v_0 = 354,72952$ als Mittelpunkt des Kamerasensors.

Mit den radialen Verzerrungsfaktoren $k_1 = -0,22825$,
 $k_2 = 0,03338$, $k_3 = 0,00101$
und $k_4 = -0,00106$ und Projektionsmatrix

$$(5.6) \quad \begin{pmatrix} 560 & 0 & 500 \\ 0 & 470 & 354,72952 \\ 0 & 0 & 1 \end{pmatrix}$$

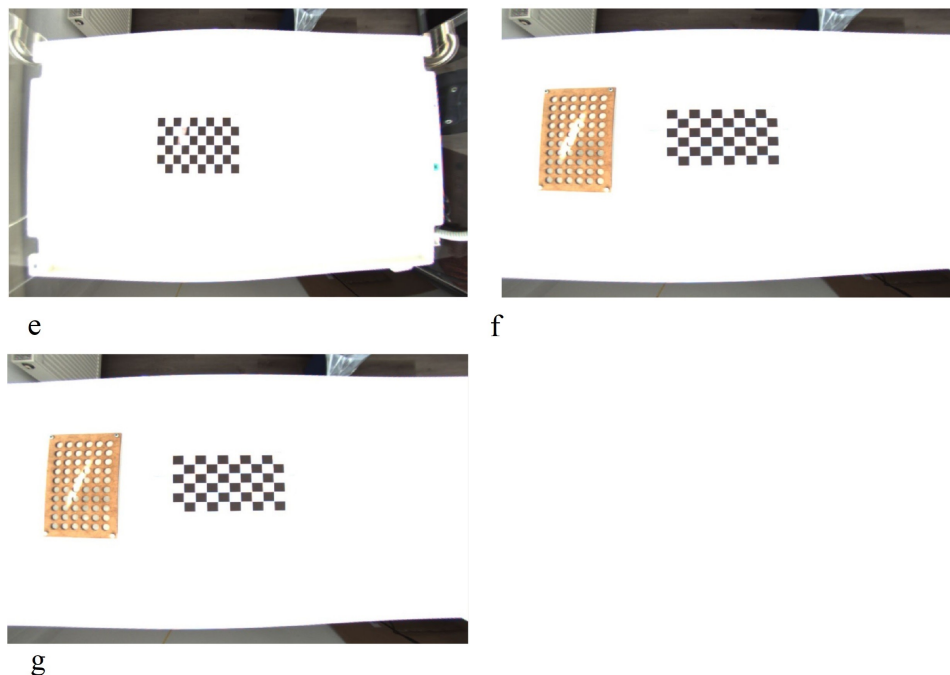


Abb. 5.12: Empirische Anpassung der intrinsischen Kalibrierung der RGB-Kamera.
 (e) Änderung des dritten Werts der Projektions-Matrix von 456,18869 zu 470;
 (f) Projektions-Matrix mit erstem Wert 560, zweiten Wert 500 und drittem Wert 470;
 (g) Änderung der Verzerrungs-Matrix zu $[-0,22825, 0,03338, 0,00101, -0,00106, 0]$ [Gle+23]

5.2.1.2 Intrinsische Kalibrierung der Thermokamera

Das Kalibriermuster für die Kalibrierung der Thermokamera besteht aus einer 2 mm dicken Aluplatte und einem 3 mm dicken High Density Fibreboard, aus welchem 6 x 10 Kreise mit einem Durchmesser von 1 cm mit einem Lasercutter herausgeschnitten wurden. Der Abstand der Mittelpunkte beträgt 13,96 mm.

Die Kalibrierung wurde direkt im Inkubator mit eingebauter Sensorbox durchgeführt.

Für die Kalibrierung kommt das ROS-Paket *camera_calibration* [ros] zum Einsatz. Um die Kalibrierung zu starten, muss folgender Befehl auf der Konsole eingegeben werden:

```
roslaunch camera_calibration cameracalibrator.py --size 10x6 -p 'circles' --square 0.01396 image:=/optris/thermal_binary camera:=/optris
```

Parallel muss das Tool *rqt_reconfigure* gestartet werden, um den Temperatur-Schwellwert des binären Thermokamera-Bildes zu setzen. In diesem Fall werden die Temperaturen in schwarz-weiß kodiert. Je nach Temperatur des Kalibrieremusters variiert der passende Schwellwert, damit alle Kreise des Kalibrieremusters sichtbar sind.

Es wurde darauf geachtet, dass das Kalibriermuster so im Bild liegt, dass die lange Kante horizontal und die kurze Kante vertikal im Bild positioniert ist. Es wurden 144 Bilder eingelesen. Daraus resultierte eine gute Rektifizierung des Bildes (siehe Abbildung 5.13). Die Kisseneffekte an den Kanten sind komplett entfernt und das Kalibriermuster wird als Rechteck im Thermokamerabild dargestellt. Dies wurde auch an verschiedenen Stellen im Bild getestet. Mit Hilfe einer Babypuppe wurde überprüft, dass auch bei kurzen Distanzen zur Kamera keine Verzerrungen auftreten. Die genauen Werte der Kameramatrix und Verzerrung sind untenstehend zu finden.

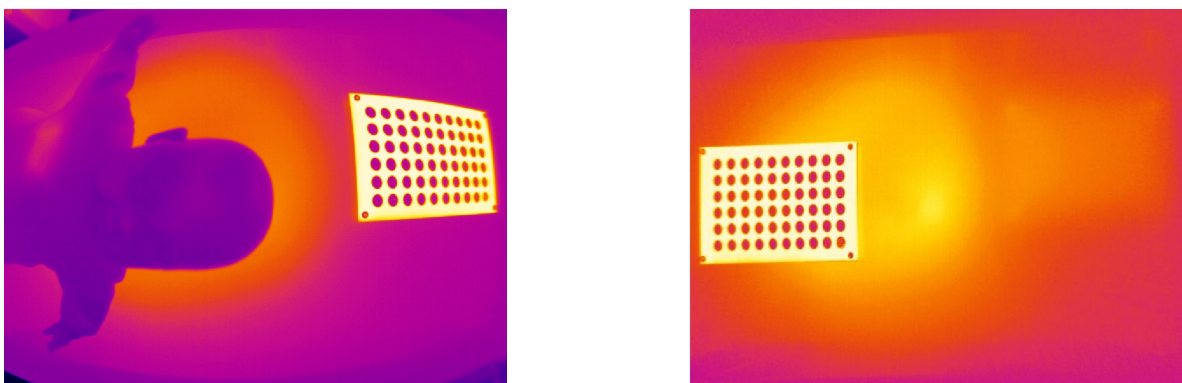


Abb. 5.13: Finale Kalibrierung der Thermokamera mit 144 Bildern. Kalibriertes Bild (re.) und unkalibriertes Bild (li.). Der Kisseneffekt ist komplett entfernt. [Gle+23]

$$(5.7) \quad \begin{pmatrix} f_x & 0 & f_y \\ 0 & u_0 & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

mit $f_x = 457.778477$ und $f_y = 458.128864$ als Brennweite, $u_0 = 305.897364$ und $v_0 = 240.829994$ als Mittelpunkt des Kamerasensors. Mit den radialen Verzerrungsfaktoren $k_1 = -0.433521$, $k_2 = 0.148799$, $k_3 = 0.000553$ und $k_4 = 0.001429$.

5.2.2 Extrinsische Kalibrierung

Die extrinsische Kalibrierung wird in mehreren Schritten durchgeführt, die bereits im Kapitel zu den theoretischen Ansätzen erläutert wurden.

5.2.2.1 Kreis-Detektion

Zur Detektion der Kreismittelpunkte, wurde eine blob detection angewandt. Eine blob detection erkennt zusammenhängende kreisförmige Muster und liefert keypoints zurück, die die Mittelpunkt-Koordinaten der Blobs beinhalten [opea; staa]. Die Ergebnisse können in Abbildung 5.14 gefunden werden.



Abb. 5.14: Blob detection im Thermo (li.), ToF (mi.) und RGB-Kamerabild (re.) Bei dem RGB-Bild wurde zur Erhöhung des Kontrastes ein blaues Papier unter das Kreismuster gelegt.

5.2.2.2 RGB-ToF und Thermal-ToF-Fusion

Die Pixel-Indizes der ToF-Kamerabilder und der Punktwolke gleichen einander aufgrund der internen Fusion. Durch das Ordnen der Punktwolke kann der korrespondierende 3D-Punkt für jedes Pixel bestimmt werden. Um die Transformation zu berechnen, werden die 3D-ToF-Punkte und die 2D-Thermo- und RGB-Punkte der Funktion *SolvePnP* übergeben. Die Funktion berechnet die Orientierung und Position des Objekts aus den gegebenen Objekt-Punkten (3D-ToF-Punkte) und Bildprojektionen (2D-Thermo- und RGB-Punkte), sowie der Kameramatrix und den Verzerrungs-Koeffizienten [oped]. Intern wird der Direct Linear Transform (DLT) Algorithmus durchgeführt, der die Matrix berechnet, die die 3D-ToF-Punkte in die 2D-Bildebene projiziert. Um die Rotationsmatrix aus *rvec* zu erhalten, kann die openCV Funktion *Rodrigues* verwendet werden [opep]. Sie berechnet die zugehörige Matrix aus dem Vektor. Um die berechnete Transformation auf die ToF-Daten anzuwenden und um die Punkte in jedes Bild zu projizieren, kommt die Funktion *projectPoints* zum Einsatz [opeb].

Im Folgenden werden die verwendeten Punktkorrespondenzen zur Berechnung der RGB-ToF-Fusion (26 Korrespondenzen, siehe Abbildung 5.15) und der Thermo-ToF-Fusion (15 Korrespondenzen, siehe Abbildung 5.16) markiert und die projizierten Punkte in die Bilder werden gezeigt. Die passendste Kombination von Punktkorrespondenzen wurde empirisch ermittelt. Um eine optimale Fusions-Genauigkeit zu erhalten, können signifikante Punkte aus dem Video der Neonaten als Punktkorrespondenzen eingebunden werden. Es mussten für jeden Probanden Anpassungen vorgenommen werden.

Sobald die extrinsischen Transformationen bekannt sind, kann die Thermo-**RGB**-Fusion berechnet werden. Für jeden ToF-Punkt in der Punktwolke werden die zugehörigen RGB- und Thermo-Pixel bestimmt (siehe Abbildung 5.17). Dann wird die Farbinformation der RGB-Pixel an der Stelle des zugehörigen Punktes im Thermobild eingefügt, sodass das finale Fusionsbild resultiert. Die Größe des Fusionsbildes entspricht dem des Thermokamerabildes.

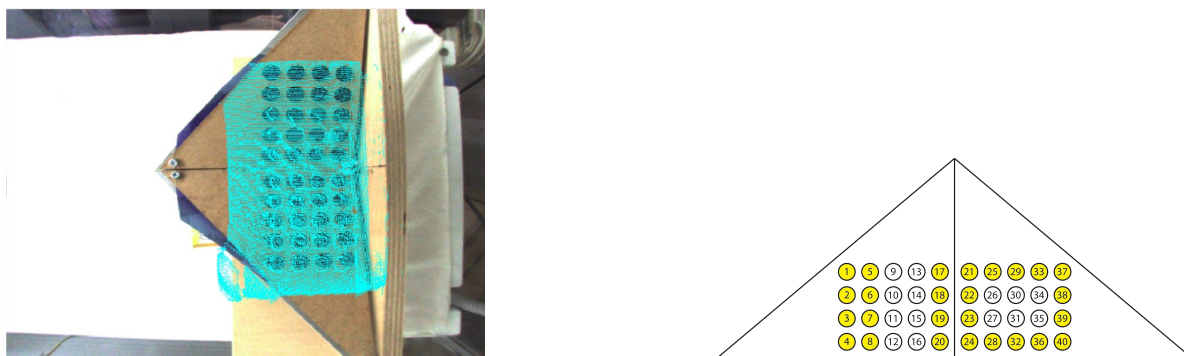


Abb. 5.15: Extrinsische Kalibrierung der RGB- mit der ToF-Kamera; 26 Punktkorrespondenzen wurden verwendet (re.). Die türkisen Punkte sind die projizierten ToF-Punkte (li.).

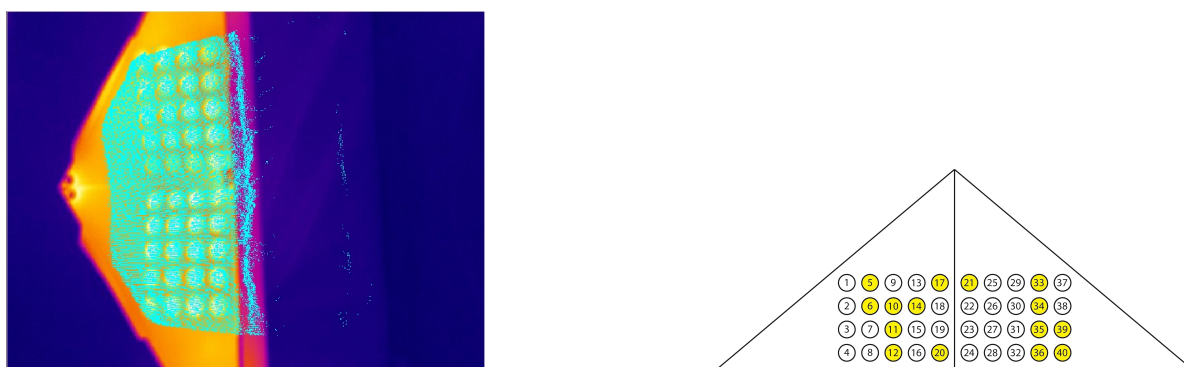


Abb. 5.16: Extrinsische Kalibrierung der Thermo- mit der ToF-Kamera; 15 Punktkorrespondenzen wurden verwendet (re.). Die türkisen Punkte sind die projizierten ToF-Punkte (li.).

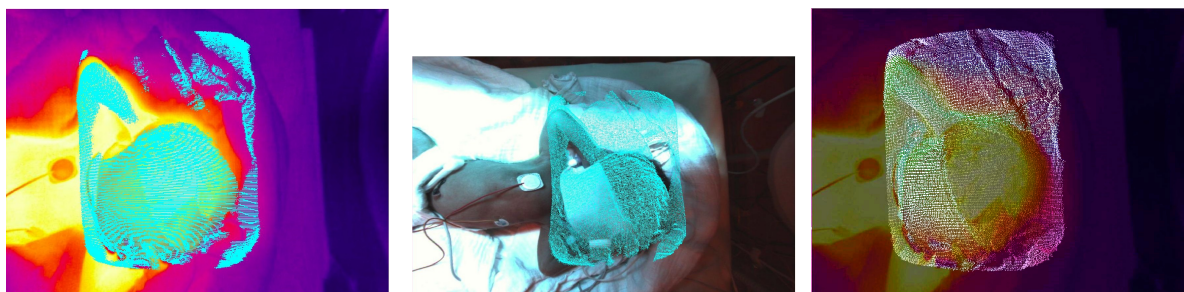


Abb. 5.17: Thermo-ToF-Fusion (li.), RGB-ToF-Fusion (mi.) und Thermo-**RGB**-Fusion Kamera-bild (re.).

5.3 Algorithmik

5.3.1 Verwendete Algorithmen

Zunächst sollen Algorithmen vorgestellt werden, die im Entwurf verwendet werden.

5.3.1.1 Eulerian Video Magnification

Die Eulerian Video Magnification ist ein wichtiger Bestandteil der Herzschlagdetektion, wurde aber zunächst auch für die Implementierung der Atemratendetektion getestet. Die kleinen Bewegungen bzw. Farbänderungen im Gesicht entsprechend der sich ändernden Sättigung des Blutes sollen mit diesem Algorithmus verstärkt werden. Die Funktionsweise der EVM wurde bereits im Kapitel 3 beschrieben. Die Problematik bei der Anwendung des Euler-Algorithmus ist, dass die Parameter aus Veröffentlichungen nicht allgemein verwendbar sind, sondern für den individuellen Anwendungsfall bestimmt werden müssen. Zunächst wurde mit Parametern aus der Veröffentlichung von Wu et al. [Wu+12] getestet. Die individuelle Parametrierung zur Anwendung bei der Herzratendetektion findet sich in Tabelle 5.4.

Bisher bestehen keinerlei offizielle Implementierungen der Eulerian Video Magnification für ROS. Es gibt aber C++-Implementierungen des Algorithmus von Nermin Sevi [Sev13] und von Ramsundar [Ram16]. Die Version von Frau Sevi hat laut Verfasserin die Nachteile, dass sie bei schlechten Lichtverhältnissen nicht gut funktioniert. Um ein erstes praktisches Verständnis für den Algorithmus zu erlangen, wurde die C++-Version von Ramsundar getestet und anschließend angepasst, um die Funktionalität zu verbessern. Anstelle der Laplace-Pyramide wird nun eine Gauss-Pyramide verwendet. Zudem wurde das IIR-Filter der C++ Implementierung durch ein ideales Bandpassfilter ersetzt, wie Wu et al. in ihrer Matlab-Implementierung ursprünglich vorgeschlagen hatten.

ROS-Implementierung Um die Sensordatenverarbeitung zu vereinfachen und die Vorteile der Kommunikation in ROS zu nutzen, wurde der bestehende C++-Code von Ramsundar K G zu einem ROS-Paket umgewandelt. Hierfür mussten *Callback-Funktionen*, *Subscriber* und *Publisher* implementiert werden, die direkt auf die Topics des Bagfiles zugreifen und die Bildverarbeitung auf diesen Daten durchführt.

5.3.1.2 Algorithmen zur Frequenzanalyse

Typische Algorithmen, die zur Frequenzanalyse eingesetzt werden können, sind die Fourier Transformation oder die Peak Detektion. Die Fourier Transformation wurde bereits in Kapitel 2 beschrieben. Die Implementierung basiert auf der *fftw3*-Bibliothek [FFT]. Bei der Peak Detektion wird innerhalb eines Datenfensters die Anzahl der Maxima, deren Werte über dem Mittelwert der Peaks liegen, bestimmt.

5.3.2 Detektion von Gesichts- und Körperarealen basierend auf Neuronalen Netzen

Mit Hilfe von Neuronalen Netzen wird die Region of Interest geliefert, in der die eigentliche Detektion von Herz- und Atemrate bzw. Körpertemperatur stattfindet.

5.3.2.1 Neuronale Netze

Die beiden Neuronalen Netze RetinaNet und YOLOv3 sollen nicht nur mit RGB-Kamera-Bildern bzw. Thermokamera-Bildern trainiert werden, sondern auch mit fusionierten Thermo-RGB-Kamera-Bildern. Bei der Fusion kombiniert man die Vorteile beider Sensoren miteinander. Beispielsweise wird die RGB-Kamera bei Dunkelheit keine hilfreichen Bilder liefern, die Thermokamera aber schon. Das RGB-Bild hingegen liefert noch mehr Strukturen im Gesicht und signifikantere Punkte als die Thermokamera.

RetinaNet In dieser Arbeit wird die RetinaNet Implementierung des Git Repository's von [Fiz] verwendet, welche *keras* und *tensorflow* (v2.4. kommt zum Einsatz) verwenden. Für den funktionalen GPU Support, müssen mehrere Software Pakete installiert werden. Dies beinhaltet NVIDIA Treiber (driver 465.19.01), *CUDA* 11.3, *cuDNN* und *OpenCV* 4.5.254.

YOLOv3 Die YOLOv3 Version im Git repository von AlexeyAB [Ale] wird eingesetzt. Für einen reibungslosen Ablauf müssen verschiedene Software-Requirements erfüllt sein (Details siehe RetinaNet).

ROS-Implementierung Es wurde ein ROS-Interface entwickelt, damit RetinaNet und YOLOv3 direkt auf den Kamera-Topics arbeiten können.

5.3.3 Kontaktlose Atemdetektion

Folgende unterschiedliche Ansätze zur Atemdetektion per RGB-Kamera, Thermokamera, Structured-Light-Kamera, ToF-Kamera und Radar-Sensor sollen getestet und verglichen werden.

5.3.3.1 RGB-Kamera

Die Frequenz der Brustkorbbewegung soll in den RGB-Kamerabildern detektiert werden. Hierfür wurden drei Verfahren getestet: der Dense-Optical-Flow-Algorithmus von Brox et al. [Bro+11], der Lucas-Kanade-Algorithmus von Li et al. [LYT14] und Temporale Differenz (Differenzbilder) wie von [Jor+17] erwähnt. Da in ersten Tests mit keiner dieser Methoden genaue Ergebnisse mit der RGB-Kamera erreicht wurden, soll die Atemdetektion nicht mit dieser Kamera durchgeführt werden. Damit ist auch eine angedachte Fusion mit einem lichtunabhängigen Sensor hinfällig. Als Alternative kommen die Thermo- und die Structured-Light- bzw. Time of Flight-Kamera infrage.

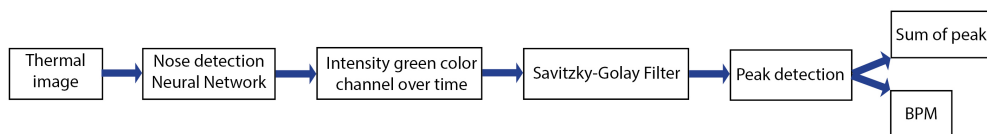


Abb. 5.18: Ablauf des Algorithmus zur Atemdetektion mit einer Thermokamera.

5.3.3.2 Thermokamera

Zunächst wird die Region of Interest der Nase im Thermokamerabild durch das Neuronale Netzwerk detektiert. Die Intensität des grünen Farbkanals des Bildausschnittes wird verwendet, um die Atemrate zu bestimmen. Vor der Peak Detektion kommt ein Savitzky-Golay-Filter mit der Fenstergröße $m = 4$ und dem Grad des Polynoms $n = 2$ für Atemraten, die kleiner gleich 19 BPM oder größer 50 BPM sind, zum Einsatz. Die Grenzwerte wurden empirisch ermittelt. Zwischen diesen Grenzen findet keine Glättung statt. Der Ablauf des Algorithmus ist in Abbildung 5.18 zu sehen.

5.3.3.3 Structured-Light-Kamera

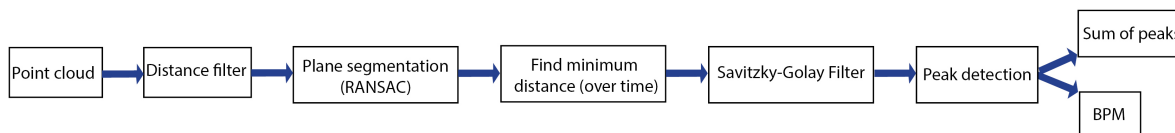


Abb. 5.19: Ablauf des Algorithmus zur Atemdetektion mit einer Structured-Light-Kamera [GNK20].

Die aufgenommenen Daten der Brustkorbbewegung des SimBaby-Simulators werden folgendermaßen weiterverarbeitet (siehe Abbildung 5.19). Um die Tischebene zu entfernen, wird ein Distanzfilter angewandt, der alle Punkte, die mehr als 35 cm von der Kamera entfernt liegen, löscht. Auf die restlichen Punkte wird eine RANSAC-Segmentierung angewandt, die dem Ebenen-Fitting dient [PCLc]. Die Point Cloud Library (PCL) liefert die benötigten Funktionen. Um die Punktwolke des Datentyps *PointCloud2*, die der Sensor liefert, mit der PCL verarbeiten zu können, muss die Punktwolke zunächst umgewandelt werden. Die resultierende Punktwolke nach der Segmentierung ist in Abbildung 5.20 zu sehen.

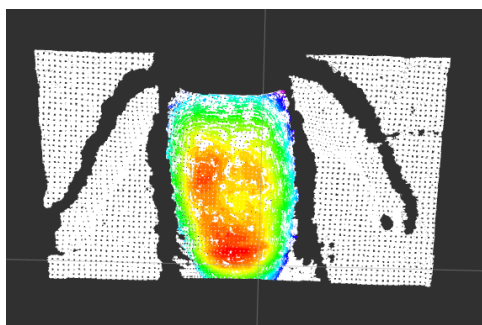


Abb. 5.20: Segmentierung der Punktwolke. Die bunte Ebene zeigt den Brust- und Bauchbereich

Um die Bewegung zu detektieren und hieraus die Atemrate zu berechnen, soll die Distanz des Punktes mit dem kleinsten Abstand über die Zeit betrachtet werden, d.h. dass der minimale Wert in der segmentierten Punktwolke gefunden werden muss. Dieser Wert wird über die Zeit betrachtet. Um nun die Atemrate zu bestimmen, soll eine Peak-Detektion zum Einsatz kommen (eine Fouriertransformation ist nicht notwendig) [cla], die auf einer Matlab-Implementierung basiert. Um die Artefakte zu entfernen wurden verschiedene Methoden getestet:

- **Hilbert-Transformation** [hao], um eine Hüllkurve zu berechnen. Hierbei wird zunächst die Fourier-Transformierte berechnet, eine Phasenverschiebung durchgeführt und die Inverse berechnet. Diese Methode liefert keine ausreichende Glättung.
- Kriterium, dass ein Peak nicht größer sein darf, als die Differenz zwischen Mittelwert und Minimum. Dies funktioniert nicht, da der Mittelwert auch die ungewollten Peaks in die Berechnung einbezieht.
- **Savitzky-Golay Filter basierend auf Gram-Polynomen** [Gor90] [arn]. Dieses Filter liefert ein gutes Glättungsverfahren und wurde daher ausgewählt. Die Parametrierung ist in der untenstehenden Tabelle zu sehen.

Tab. 5.1: Parameter für Savitzky-Golay Filter [GNK20]

BPM	Fenstergröße	Grad des Polynoms
≤ 18	11	10
18 - 50	3	2
50 - 62	3	1
> 62	-	-

5.3.3.4 ToF-Kamera

Der Signalverarbeitungs-Ablauf sieht folgendermaßen aus (siehe Abbildung 5.21): Die ToF-Kamera liefert eine 3D-Punktwolke. Im ersten Schritt entfernt ein Distanzfilter alle Punkte, die mehr als 20 cm von der Kamera entfernt sind, denn diese entsprechen der Tisch-Ebene. Danach wird eine Ebenen-Segmentierung (plane segmentation) der PCL, die auf dem Least Median of Squares (LMEDS) basiert, durchgeführt [PCLd] [PCLb]. Der eingesetzte Modell-Typ heißt *SACMODEL_PLANE*. Es wird ein Distanz-Schwellwert von 2 cm verwendet, was bedeutet, dass Punkte, die einen maximalen Abstand von 2 cm zueinander haben, derselben Ebene angehören. Um Ausreißer am Rand der Punktwolke, die durch Reflexionen verursacht werden, zu entfernen, kommt ein Statistical-Outlier-Removal-Filter zum Einsatz [PCLa], mit einem Multiplikator der Standard-Abweichung von $\alpha = 1$ und $k = 30$ Nachbarn. Zur Rauschreduktion wird ein Median-Filter [Ich] mit einem *maximum_allowed_movement* von 3 mm verwendet. Der Parameter *maximum_allowed_movement* beschreibt, um welche Distanz sich ein Tiefen-Pixel (Dexel) bewegen darf, wenn das Filter angewandt wird.

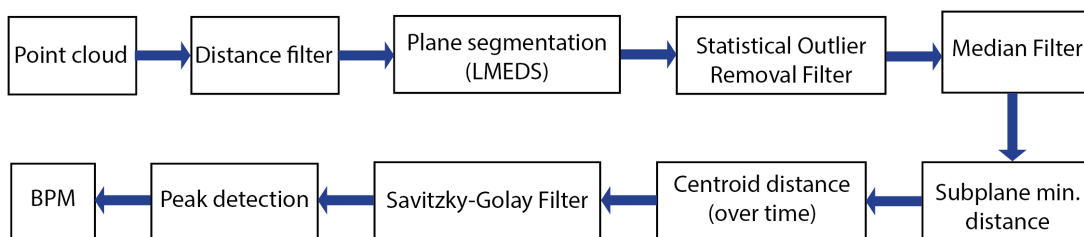


Abb. 5.21: Ablauf der ToF-Atemdetektion. Im Unterschied zur Datenverarbeitung bei der Structured-Light-Kamera wird bei der Ebenendetektion anstelle vom RANSAC der LMEDS Algorithmus angewandt. Hinzu kommt das Statistical-Outlier-Removal-Filter, das Medianfilter und die PCA mit Bestimmung des Schwerpunkts über die Zeit. [Gle+21]

Im nächsten Schritt wird der Punkt mit der minimalen Distanz zur Kamera bestimmt und eine Subebene, die aus allen Punkten besteht, die maximal 4 mm (in die z-Richtung) vom Punkt der minimalen Distanz liegen, berechnet. Der Schwerpunkt dieser Subebene wird über die Zeit geplottet und entspricht der Distanzänderung des Thorax, verursacht durch die Atmung. Ähnlich wie bei der Signalverarbeitung mit der Structured-Light-Kamera werden Pakete von 33 Schwerpunkts-Werten über die Zeit gesammelt. Für die ersten drei Pakete (nach dem Synchronisationssignal) wird mit Hilfe einer Peak Detection [Gon] die Anzahl der Atemzüge pro Minute in breaths per minute (BPM) bestimmt. Die Atemrate in BPM lässt sich durch Division der Anzahl der Peaks durch die Zeit in Sekunden mal 60 s berechnen. Der Durchschnitt der Atemraten der drei Datenpakete wird zur Parametrierung des Savitzky-Golay Filters [Gor90] [arn], der auf Gram-Polynomen beruht, genutzt. Die Fenstergröße und der Grad des Polynoms variieren je nach Größe des ersten BPM-Werts (siehe Tabelle 5.2). Die Stärke der Filterung ist frequenzabhängig. Da bei kleineren Atemraten mehr Rauschen vorliegt, wird dann eine stärkere Filterung benötigt. Die idealen Filterparameter wurden empirisch ermittelt.

Tab. 5.2: Parameter für den Savitzky–Golay Filter [Gle+21].

BPM	Fenstergröße	Grad des Polynoms
≤ 19	11	10
19–32	3	2
>32	-	-

In Abbildung 5.22 ist ein Abschnitt des gefilterten (Magenta) und des ungefilterten (Blau) Atemsignals bei 20 BPM im *deep* Modus des Baby-Thorax-Simulators zu sehen. Danach wird wie zuvor eine Peak Detection auf die gefilterten Werte angewandt, um die Anzahl der Peaks innerhalb eines Datenpakets über die Zeit zu ermitteln.

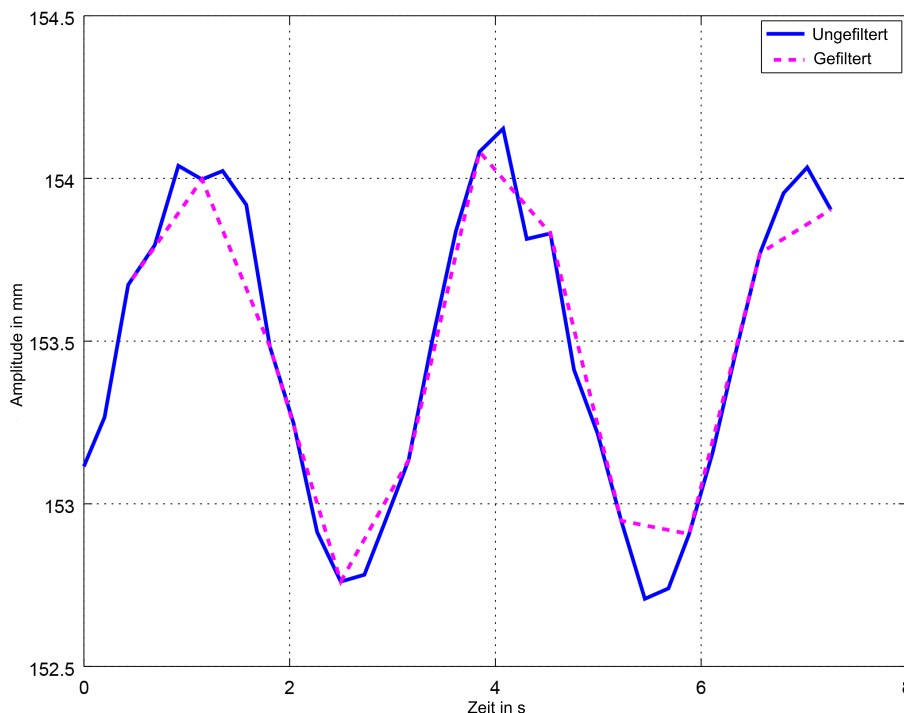


Abb. 5.22: Ausschnitt des ungefilterten (blau) und gefilterten (magenta) Atemsignals bei 20 BPM im *deep* Modus. Das angewandte Savitzky–Golay Filter hat eine Fenstergröße 3 und der Grad des Polynoms liegt bei 2. Durch das Filter wird eine Glättung des Signals bewirkt, sodass die Peaks eindeutig herausgestellt werden. [Gle+21]

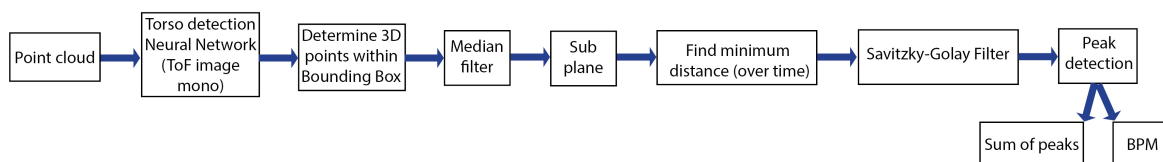


Abb. 5.23: Ablauf der Atemdetektion mit der ToF-Kamera und integriertem Neuronales Netz.

Integration des Neuronales Netzes Die Time-of-Flight-Kamera kommt in Kombination mit dem Neuronales Netz, das mit RGB- und ToF-Mono-Bilddaten trainiert wurde, zum Einsatz. Bei Vorarbeiten am Frühgeborenen-Simulator Paul fielen Ungenauigkeiten auf, die auf eine Fehldetektion des Thorax hinwiesen. Stattdessen wurde fälschlicherweise die Windel als Thorax detektiert, wodurch die Atemratenwerte Fehler aufwiesen. Um dem zu begegnen, wird die ROI "torso" im ToF-Mono-Bild über das Neuronales Netz eingegrenzt und über die intern vorliegende Fusion kann auf die zugehörigen 3D-ToF-Punkte rückgeschlossen werden. Nur Punkte, die innerhalb der Bounding Box liegen, werden weiterverarbeitet. Der Ablauf der Atemdetektion basierend auf der ToF-Kamera mit integriertem neuronalen Netz ist in Abbildung 5.23 zu sehen. Das Savitzky-Golay Filter musste für die durch die Studie gewonnenen realen Daten neu parametrisiert werden. Geeignete Werte wurden empirisch ermittelt (siehe Tabelle 5.3).

Tab. 5.3: Parameter für den Savitzky–Golay Filter reale Daten.

BPM	Fenstergröße	Grad des Polynoms	Fenstergröße mov. aver. Window
≤ 10	5	10	16
> 10 < 80	1	1	16
> 80	2	1	4

5.3.3.5 Radar-Sensor

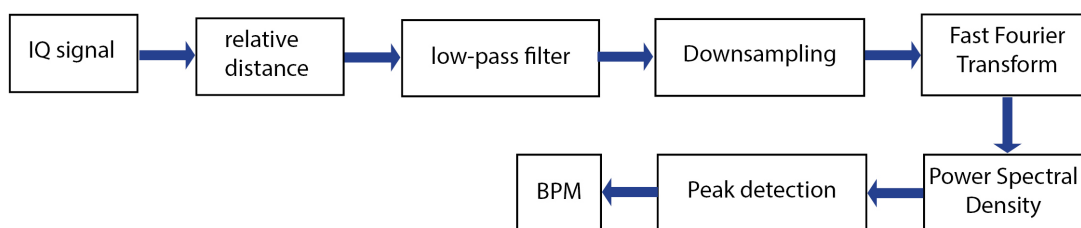


Abb. 5.24: Signalverarbeitung für die Atemraten-Detektion mit dem Radar-Sensor.

In Abbildung 5.24 ist der Ablauf der Signalverarbeitung zur Atemdetektion mit dem Radar-Sensor zu sehen. Der Radar-Sensor liefert ein IQ-Signal, das folgendermaßen mathematisch dargestellt werden kann:

$$(5.8) \quad I = r \cdot \cos(\sigma) = \Re(Z)$$

$$(5.9) \quad Q = r \cdot \sin(\sigma) = \Im(Z)$$

Durch Anwendung der Arkustangens-Methode zur Demodulation kann die relative Phasenverschiebung $\Delta\sigma$ der beiden Signale berechnet werden:

$$(5.10) \quad \Delta\sigma = \arctan\left(\frac{\Re(Z)}{\Im(Z)}\right) = \arctan\left(\frac{I}{Q}\right)$$

Im ersten Schritt wird die relative Distanz wie folgt bestimmt:

$$(5.11) \quad \Delta x = \frac{\Delta\sigma}{2 \cdot \pi} \cdot \frac{\lambda}{2}$$

mit Δx als relative Distanz, $\Delta\sigma$ als relative Phasenverschiebung und λ als Wellenlänge. Die Distanzwerte werden mit einem Tiefpass mit einer Grenzfrequenz von 7 Hz gefiltert. Danach wird ein Downsampling zu 20 Hz durchgeführt. Eine FFT wird angewandt, um das Frequenzspektrum des Signals zu berechnen. Die FFT verwendet die letzten 5 s, was 100 Werten bei 20 Hz entspricht. Der Datensatz wird per Zero Padding auf 2000 Werte erweitert, d.h. die FFT wird mit 2000 Werten berechnet, wobei 1900 auf Null gesetzt sind. Dadurch wird eine spektrale Auflösung von 0,01 Hz erhalten. Die FFT liefert 1000 komplexe Werte, aus denen die Power

Spectral Density (PSD) berechnet wird $((\Re)^2 + (\Im)^2)$. Um die Frequenz mit dem höchsten Peak zu erhalten, wird eine Peak-Detektion, die auf einem quadratischen Savitzky-Golay [Gor90] mit Fenstergröße 1 basiert, angewandt. Der Savitzky-Golay berechnet die erste und zweite Ableitung, die verwendet werden kann, um herauszufinden, ob ein lokales Maximum erreicht wurde. Der größte Peak im Spektrum entspricht der Atemfrequenz in Hz. Alle Peaks, die unterhalb 10 % des höchsten Peaks und unter 0,08 Hz liegen, werden ignoriert. Im letzten Schritt wird die Atemrate von Hz in BPM umgerechnet.

5.3.3.6 Synchronisierung ToF-Kamera und Radar-Sensor

Die ToF-Kamera und der Radar-Sensor werden durch dasselbe launch-File gestartet und erhalten dadurch automatisch ROS-Zeitstempel in Unix-time. Die Sensorik, der Baby-Thorax-Simulator und die Datensatz-Aufnahme werden direkt aus einem bash-Skript heraus gestartet. Durch das bash-Skript wird auch ein Synchronisationssignal simuliert, ein Sinussignal bei 180 BPM, das bevor die Simulation der Atmung startet, eine Sekunde lang andauert. Dieses Signal wird für die Synchronisierung mit dem Laser-Mikrometer benötigt. Zwischen dem Synchronisationssignal und dem Beginn der eigentlichen Simulation findet eine kurze "Schlafphase" statt, sodass das Synchronisationssignal keinen Einfluss auf die Messdaten hat.

Die Messwerte, die das Laser-Mikrometer bei einer Abtastrate von 2500 Hz liefert (Zeitintervall ab Startzeit und Distanz in mm), können als csv-File gespeichert werden. Um das Laser-Mikrometer mit den anderen Sensoren zu synchronisieren, muss zunächst die Startzeit der Messung in Unix-time mit Hilfe eines bash-Skripts umgewandelt werden und alle einzelnen Zeitstempel werden durch Addition der Zeitintervalle aus dem csv-File berechnet. Die Daten werden durch einen ROS-Knoten eingelesen, welcher die Daten des csv-Files zu einer ROS-Topic umwandelt, die gemeinsam mit den anderen Daten durch den Publisher gesendet wird.

Wenn das Distanzsignal aus den ToF-Kamera- und Radardaten berechnet wird, erhält jeder Distanzwert einen korrekten Datensatz-Zeitstempel. Dann können die synchronisierten Laser-Mikrometer-, ToF- und Radar-Distanzsignale mit Hilfe des ROS-Tools *rqt_plot* dargestellt werden.

5.3.4 Kontaktlose Herzratendetektion

Zur Herzratendetektion sollen folgende Ansätze mit einer RGB-Kamera und einem Radar-Sensor getestet werden.

5.3.4.1 RGB

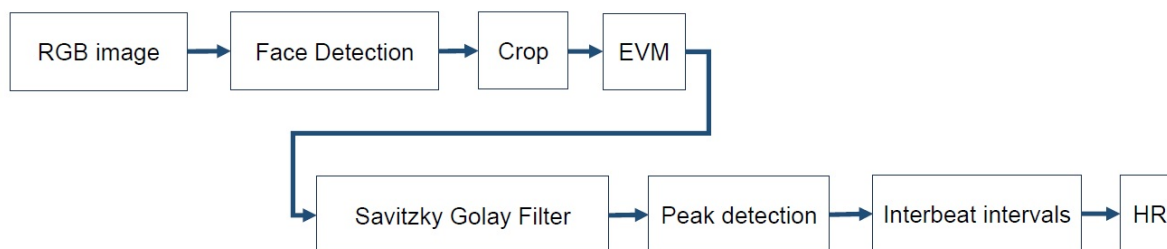


Abb. 5.25: Ablauf der Herzrattendetektion mit der RGB-Kamera und integriertem Neuronalem Netz.

Zunächst wird das Gesicht des Kindes mit dem Neuronalem Netz, das mit RGB-Daten trainiert wurde, detektiert. Der Eulerian Video Magnification Algorithmus verstärkt die Farbänderungen, die durch den Herzschlag im Gesicht in der Region of Interest verursacht werden. Die Parameter des EVM sind in Tabelle 5.4 zu sehen. Die Werte wurden empirisch ermittelt. Das Intensitätssignal über die Zeit des grünen Farbkanals des verstärkten Bildes wird zunächst mit einem Savitzky-Golay-Filter bearbeitet. Der grüne Farbkanal wird verwendet, da dieser die meiste Information über den Herzschlag liefert. Die Filterung wird benötigt, um das durch den EVM verstärkte Rauschen zu reduzieren. Die Parametrierung des Filters ist in Tabelle 5.5 zu sehen. Über das gefilterte Signal läuft eine Peak Detection und die Zeitstempel der Peaks werden gespeichert, um letztendlich die Interbeat Intervals berechnen zu können. Über den reziproken Wert des Interbeat Intervals wird die Herzrate bestimmt. Zur Glättung des Herzratensignals wird ein sliding moving window der Größe 9 verwendet. Zur Ermittlung der idealen Fenstergröße wurde der RMSE über mehrere aufgezeichnete Datensätze berechnet. Die Fenstergröße, die zum niedrigsten RMSE führt, wurde gewählt. Im Laufe der Entwicklung wurde die vom Neuronalem Netz erzeugte Bounding Box sukzessive verschmälert (bis zu einem Sechstel der ursprünglichen Breite), sodass möglichst ausschließlich Hautareale (ab Kinn) enthalten sind. Die Verwendung des gesamten Gesichts ist wegen der zu tragenden Maske während der Phototherapie schlecht geeignet, da es dann verhältnismäßig wenig Hautanteile gibt. Es wurde jeweils der RMSE und der Median der prozentualen Abweichung berechnet, um den besten ROI Bereich zu finden. Eine Bildvorverarbeitung mit Optimierung von Kontrast, Helligkeit und Sättigung wurde getestet, bringt aber keine Verbesserung der Genauigkeit.

Tab. 5.4: Parameter für den Eulerian Video Magnification Algorithmus Herzrate.

Parameter Name	Parameter
alpha	400
lambda_c	1000
cutoff_freq_low	1,33
cutoff_freq_high	3,00
chrom_attenuation	1
exaggeration_factor	7,0
delta	0
lambda	0
levels	5
input fps	9,0

Tab. 5.5: Parameter für den Savitzky–Golay Filter Herzrate.

BPM	Fenstergröße	Grad des Polynoms	Fenstergröße mov. aver. Window
≤ 60	5	10	9
> 60	1	1	9

5.3.4.2 Radar

Es wurden drei verschiedene Algorithmen für die Herzraterdetektion mit Radar entwickelt [Her21]. Diese sollen nun vorgestellt werden. Alle drei Algorithmen haben gemeinsam, dass das Radar-Distanzsignal zunächst mit einem Tiefpass gefiltert wird und die Werte anschließend per Downsampling auf 20 Hz reduziert werden. Danach wird bei Peak2peak und FFT-per-band ein Bandpassfilter verwendet, um unphysiologische Signale auszufiltern.

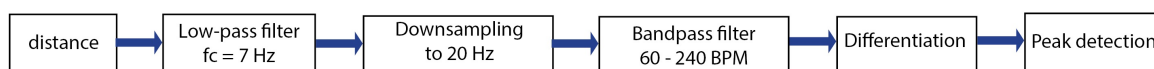


Abb. 5.26: Flowchart des Peak2peak-Algorithmus für die Herzraten-Detektion mit dem Radar-Sensor.

Peak2peak In Abbildung 5.26 ist der Ablauf des Peak2peak-Algorithmus für die Herzraten-Detektion mit dem Radar-Sensor zu sehen. Per Differenzierung werden die Maxima bestimmt. Vom Interbeat Interval, das man durch die Peakdetektion erhält, kann leicht auf die Herzrate in BPM geschlossen werden.

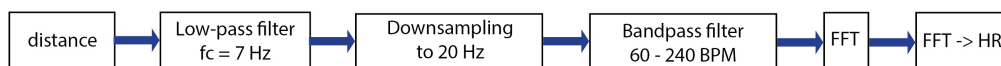


Abb. 5.27: Flowchart des FFT-per-Band-Algorithmus für die Herzraten-Detektion mit dem Radar-Sensor.

FFT-per-band Abbildung 5.27 zeigt den Ablauf des FFT-per-Band-Algorithmus. Er bestimmt die Herzrate nach dem Bandpassfilter über eine FFT. Es wird mit dem Spektrum, also der quadrierten Größe der Fouriertransformierten, gearbeitet. Der höchste Peak des Spektrums ist die relevante Frequenz (Herzrate).



Abb. 5.28: Flowchart des one-FFT-for-all-Algorithmus für die Herzraten-Detektion mit dem Radar-Sensor.

one-FFT-for-all In Abbildung 5.28 ist der Ablauf des one-FFT-for-all-Algorithmus zu sehen. Dieser Algorithmus verzichtet auf den Bandpass und berechnet direkt aus der FFT die Herzrate. Die fünf höchsten Peaks werden basierend auf der Frequenz gruppiert. 5-80 BPM für die Atemrate, 60-240 BPM für die Herzrate. Relevant sind die zwei Peaks mit der höchsten Amplitude. Es besteht eine Überlappung zwischen 60 und 80 BPM. Hierfür wird eine Konfliktstrategie angewandt, bei der die Anzahl der Kandidaten betrachtet wird. Bestehen ein Atemraten (RR) und zwei Herzraten (HR) Kandidaten, so verursacht der erste HR Kandidat den Konflikt und muss durch den zweiten HR Kandidaten ersetzt werden. Bestehen ein HR und zwei RR Kandidaten, so wird der erste RR Kandidat durch den zweiten ersetzt. Wenn für beide Gruppen nur ein Kandidat besteht, muss keinerlei Konfliktstrategie angewandt werden. Bestehen zwei RR und zwei HR Kandidaten, sind die zweiten Kandidaten für den Konflikt verantwortlich. Der Kandidat mit der

niedrigeren Frequenz entspricht dann der RR und der andere der HR. Wenn es keine Konflikte der zweiten Kandidaten gibt, wird die höchste Amplitude gewählt.

5.3.5 Kontaktlose Temperaturdetektion

5.3.5.1 Kalman-Filter

Für die Temperaturdetektion bestand zunächst die Idee, einen Kalmanfilter zu verwenden, um aus den Thermokameradaten die Körperkerntemperatur des Kindes zu berechnen, da einige Umgebungsparameter etc. berücksichtigt werden können und erwartet wurde, dass damit genaueste Ergebnisse zu erzielen wären.

Für die Kalman-Filter-Implementierung wird eine große Menge an Parametern benötigt (vgl. [Lax+18]). Da zur Berechnung mancher Parameter Werte für den Wärmetransport innerhalb der verschiedenen Gewebe- und Hauttypen des Neonaten benötigt werden, und diese der medizinischen Forschung nicht vorliegen, ist die Idee nicht umsetzbar.

5.3.5.2 Differenz periphere und Körperkerntemperatur

Als Alternative besteht die Möglichkeit, die Körperkerntemperatur aus der (altersabhängigen) Differenz zur mit der Thermokamera erfassten peripheren Temperatur zu berechnen (siehe Abbildung 5.29). Hierfür wird eine Look-Up-Table verwendet. Dafür wird zu Beginn der Aufnahme das Alter des Kindes eingegeben, damit der passende Differenzwert zugeordnet wird (siehe Tabelle 5.6). Eine automatische Korrektur der Abweichung der Temperaturwerte, die mit der Thermokamera gewonnen werden, von den Werten des Temperatursensors als Referenz wurde implementiert. Es wird die Differenz zwischen Referenz und Thermokamera berechnet und auf jeden Temperaturwert der Thermokamera addiert. Als von der Thermokamera bestimmte periphere Temperatur wird der Mittelwert des maximalen Temperaturwertes und der mittleren Temperatur des Gesichtsbereiches verwendet (dies wurde empirisch ermittelt). Der Gesichtsbereich wird über das Neuronale Netz geliefert.

Tab. 5.6: Look-Up-Table Temperaturdetektion nach [Lyo+97] und [Jor10]

Früh- (f) / Reifgeboren (r)	Alter (Tage)	Differenz periphere zu Körpertemperatur (K)
f	1	0,4
f	2	0,8
f	3	0,9
f	4	1,0
f	5	1,0
r	beliebig	1,5

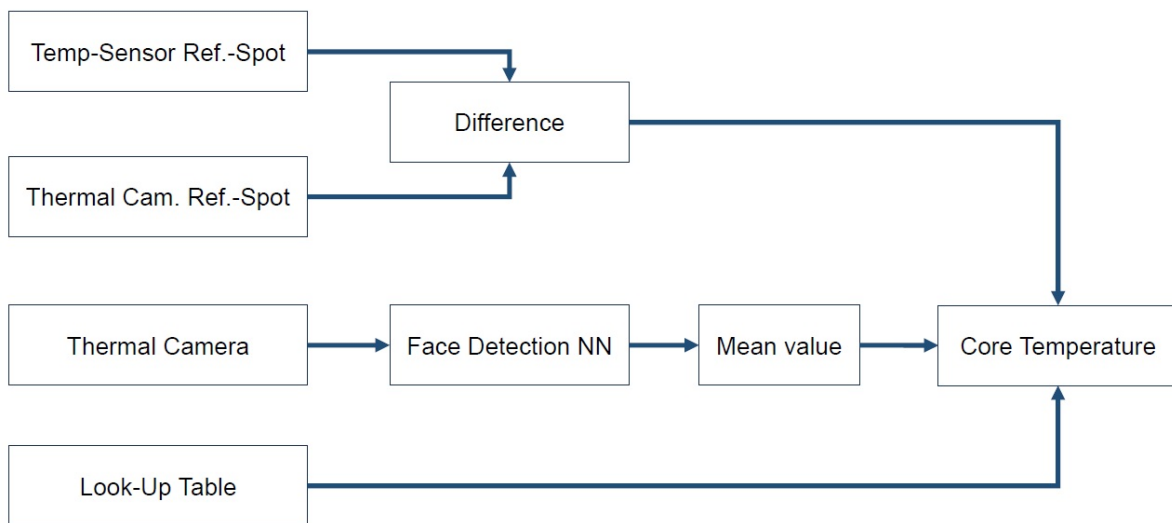


Abb. 5.29: Flowchart Temperaturdetektion über die Differenz der peripheren und der Körperkerntemperatur.

6 Validierung: Experimente und Ergebnisse

Die einzelnen Elemente des in Kapitel 5 vorgestellten Entwurfs wurden implementiert und anhand von Simulatordaten und während der Machbarkeitsstudie an Neonaten gewonnenen Daten getestet. Die durchgeführten Tests und jeweiligen Ergebnisse werden in diesem Kapitel vorgestellt.

6.1 Machbarkeitsstudie

Während der Laufzeit der Studie konnten von 18 Probanden unterschiedlicher ethnischer Herkunft Aufnahmen während der Phototherapie gemacht werden. Zum Schutz der Augen tragen die Neonaten eine Binde über der Augenpartie.

6.1.1 Datengewinnung

Tab. 6.1: Liste der Probanden

Proband	Gestationsalter	Alter bei Aufnahme	Geschlecht	Gewicht
01	reif, 36 + 2	5 Tage	weiblich	2150 g
02	reif	6 Tage	männlich	3260 g
03	reif	7 Tage	männlich	3615 g
04	reif, 37 + 0	keine Angabe	männlich	2760 g
05	reif	3 Tage	männlich	3600 g
06	34 + 0	2 Tage	männlich	1745 g
07	reif	5 Tage	weiblich	3650 g
08	reif	4 Tage	weiblich	2330 g
09	reif	13 Tage	männlich	3300 g
10	reif	2 Tage	weiblich	2750 g
11	reif	3 Tage	weiblich	2880 g
12	reif	3 Tage	männlich	2900 g
13	34 + 5	3 Tage	männlich	2120 g
14	reif	5 Tage	männlich	3330 g
15	reif	6 Tage	männlich	2410 g
16	reif	10 Tage	weiblich	3610 g
17	reif	5 Tage	männlich	2850 g
18	reif	10 Tage	weiblich	3248 g

6.1.2 Goldstandards der Vitalparameterdetektion

Als Referenzmessungen wurden die Werte von EKG, Pulsoximeter und der rektalen Temperatursonde exportiert. Bei der über die Impedanz-Pneumographie gemessenen Atemrate fielen grobe Ungenauigkeiten auf, insbesondere wenn sich das Kind stark bewegt. Aus diesem Grund wurde für alle Datensätze die Anzahl der Atemzüge zusätzlich manuell gezählt und daraus die vorliegende Atemrate mit Hilfe eines MATLAB-Programms berechnet.

6.2 Atemratendetektion mit Structured-Light-Kamera

6.2.1 Messaufbau

Die Structured-Light-Kamera wurde in einem Abstand von 40 cm zum Tisch positioniert, was den Proportionen im Inkubator entspricht. Der SimBaby-Simulator von Laerdal [Lae] wurde für die Testaufnahmen darunter gelegt.



Abb. 6.1: Aufbau mit Structured-Light-Kamera.

6.2.2 Goldstandard

Als Referenz wurde für jeden Datensatz für 15 Minuten die Anzahl der Atemhübe manuell gezählt und mit vier multipliziert, um die Anzahl pro Stunde zu ermitteln. Dabei fiel auf, dass die tatsächliche Atemrate im Referenzbereich um bis zu 15 % unter dem eingestellten BPM-Wert am SimBaby-Simulator lag.

6.2.3 Datengewinnung

Es wurden 40 Datensätze mit der Simulationspuppe SimBaby der Fa. Laerdal aufgezeichnet (Punktwolke und tf). Hierbei lag die Puppe entweder auf dem Bauch oder Rücken, und es wurden verschiedene Atemtiefen eingestellt (*deep* mit Atemhub von 3 - 5 mm und *shallow* mit Atemhub von 1 - 3 mm). Außerdem wurden unterschiedliche Atemfrequenzen eingestellt: 3, 20, 25, 30, 40, 45, 50, 55, 60 und 80 BPM. Die Aufnahmen wurden zur Entwicklung und Validierung der Atemdetektion mit der Structured-Light-Kamera eingesetzt.

6.2.4 Messergebnisse

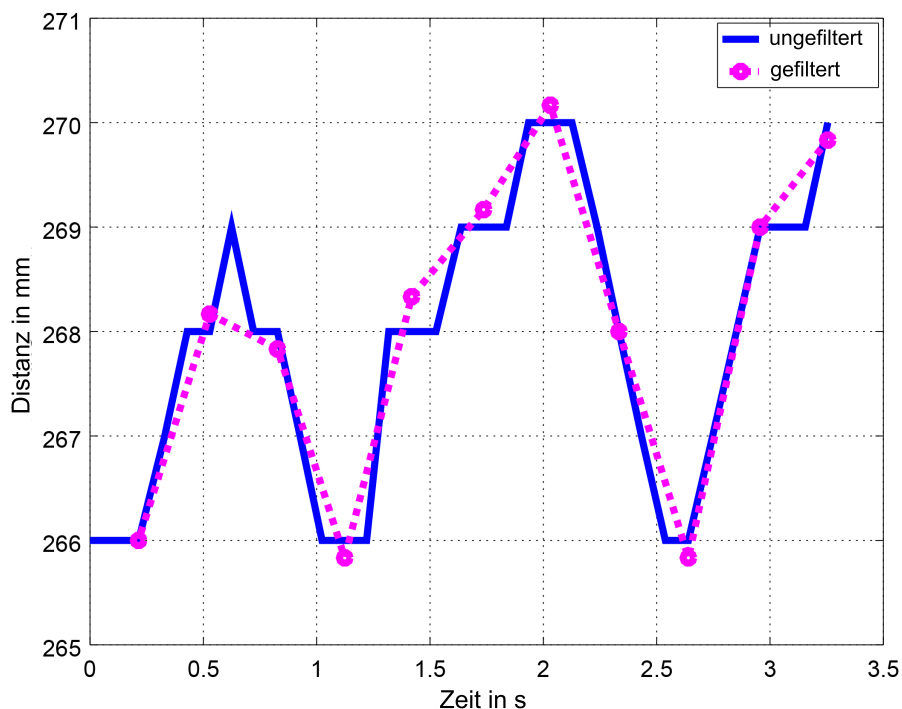


Abb. 6.2: Abschnitt des respiratorischen Signals des SimBabies in Rückenlage in *deep* mode. Die eingestellte Atemrate liegt bei 50 BPM. Daraus resultiert, dass die Fenstergröße des Filters 3 ist und es insgesamt 11 Filterfenster in diesem Abschnitt gibt (vgl. Abschnitt 5.3). Jedes Filterfenster endet an dem markierten magentafarbenen Punkt. Gefiltertes Signal (magenta), ungefiltertes Signal (blau). [GNK20]

Abbildung 6.2 zeigt einen Abschnitt des respiratorischen Signals des SimBabies in Rückenlage bei einer Atemrate von ca. 50 BPM in *deep* mode. Das blaue Signal ist das ungefilterte Signal, das magentafarbene Signal das gefilterte.

Getestet wurde mit 36 Datensätzen à 1h. Der zur automatischen Berechnung der Atemzüge verwendete Algorithmus lieferte ungefähr alle 3 s eine Messung. Abbildung 6.3 und Abbildung 6.4 zeigen die Ergebnisse für das SimBaby in Bauchlage bzw. Rückenlage im *deep* bzw. *shallow* Modus. Der Box-Whisker-Plot ist ein statistisches Werkzeug, das den Median, Minimum, Maximum und oberes und unteres Quartil von jedem Datensatz darstellt. Die kleinen roten Punkte zeigen den tatsächlichen Referenzwert. An den Randfrequenzen steigen die Abweichungen an. Je kleiner die Frequenz, desto größer der Effekt von Artefakten, somit ist die Genauigkeit niedriger. Wenn die Frequenz hoch ist, können manche Peaks nicht mehr detektiert werden. Atemraten von 3 BPM oder 80 BPM liegen nicht im Normbereich für ein Neonat, aber wurden getestet, um zu zeigen, dass ein Anstieg oder Abfall der Atemrate detektiert werden kann. Zusammengefasst kann gesagt werden, dass die Genauigkeit im *deep* Modus bei Bauchlage am größten ist, da die sich bewegende Ebene größer ist und somit stabilere Messungen liefert. Auch die Fenstergröße beeinflusst die Genauigkeit, je nachdem, wo das Signal abgeschnitten wird. Bei Atemraten zwischen 20 und 60 BPM wird eine maximale Differenz von 6 BPM erhalten. [GNK20]

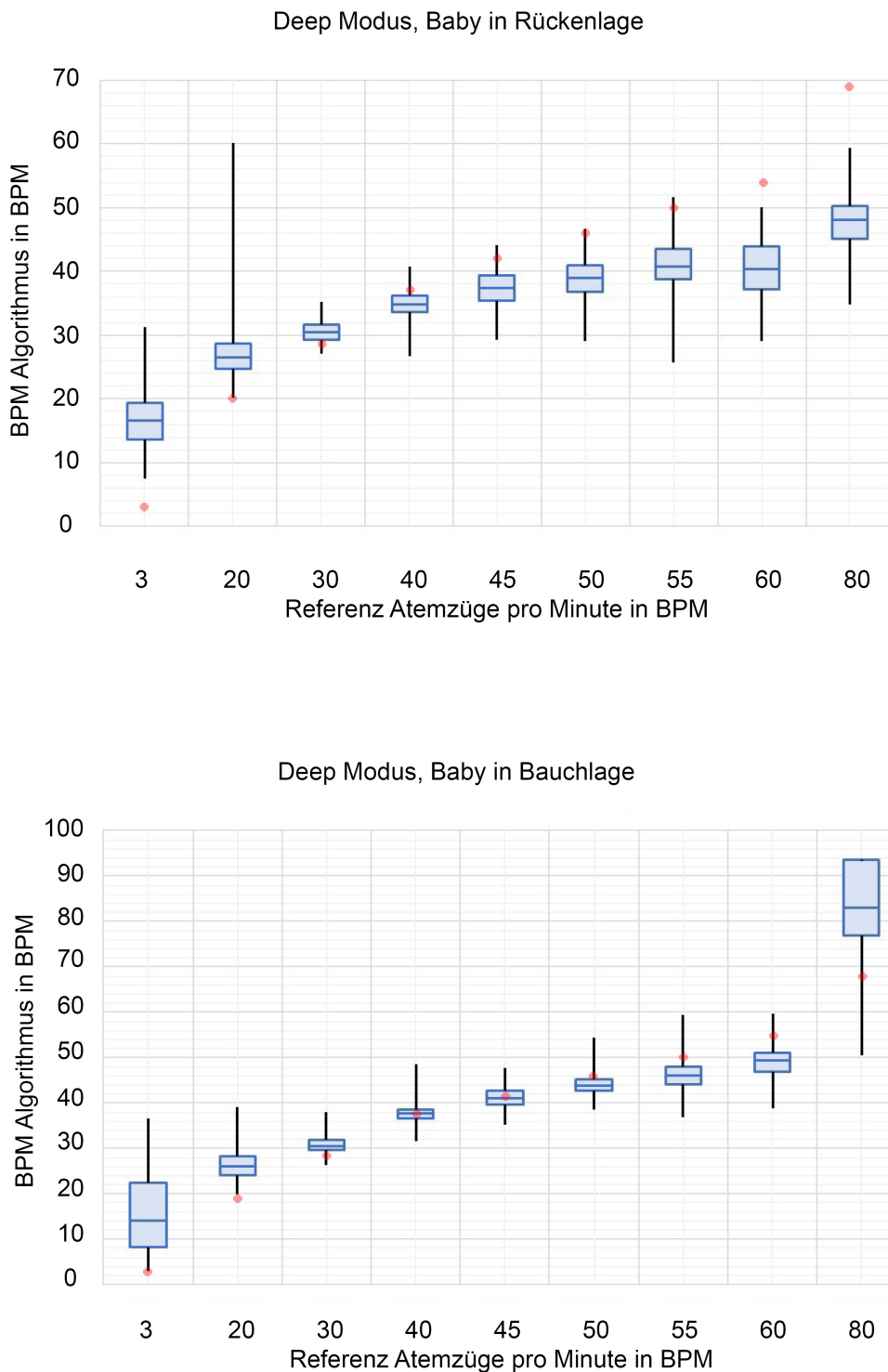


Abb. 6.3: Box-Whisker-Plot im *deep* Modus während das SimBaby auf dem Rücken liegt (oben) und in Bauchlage (unten). Die maximale Differenz zur Referenz im Normbereich liegt bei 14 BPM während das SimBaby auf dem Rücken liegt und bei 6 BPM in Bauchlage. [GNK20]

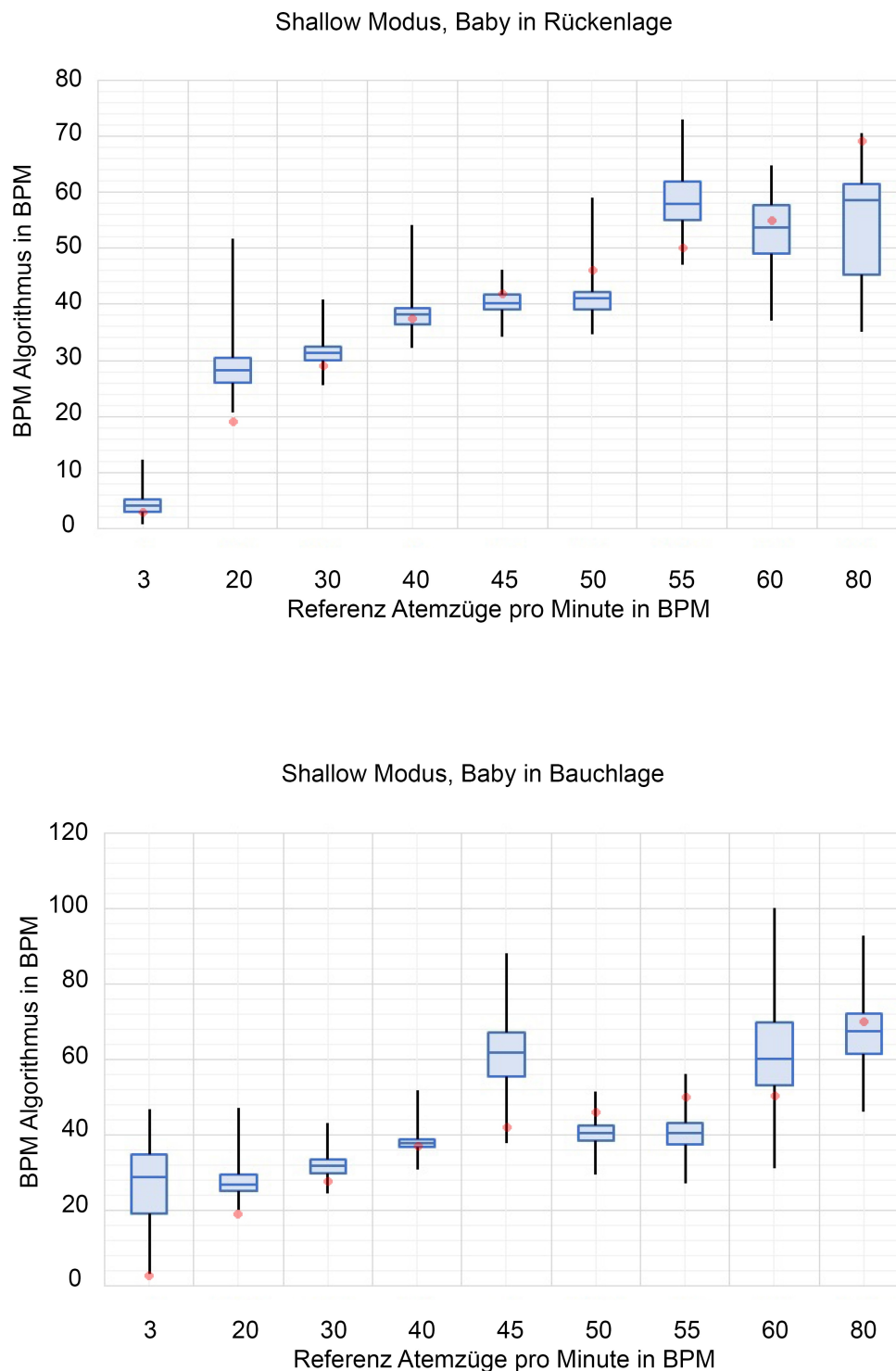


Abb. 6.4: Box-Whisker-Plot im *shallow* Modus während das SimBaby auf dem Rücken liegt (oben) und in Bauchlage (unten). Die maximale Differenz zur Referenz im Normbereich liegt bei 9 BPM in Rückenlage und 20 BPM in Bauchlage. [GNK20]

6.3 Synchrone Auswertung der Atemrate mit ToF-Kamera und Radar-Sensor

6.3.1 Messaufbau

Die Messungen wurden am selbst entwickelten Baby-Thorax-Simulator und dem Frühgeborenen-Simulator "Paul" durchgeführt. Der Messaufbau wurde bereits in Abschnitt 5.1 vorgestellt.

6.3.2 Goldstandard

Als Referenzmessung für den Baby-Thorax-Simulator wurde ein Laser-Mikrometer eingesetzt, bei dem Frühgeborenen-Simulator "Paul" gilt die voreingestellte Frequenz als Referenz. Die Eigenschaften und Genauigkeit sind in Abschnitt 5.1 beschrieben. Um sicherzustellen, dass das Laser-Mikrometer korrekte BPM-Werte liefert, wurde eine manuelle Zählung der Atemhübe zur Verifizierung durchgeführt.

6.3.3 Baby-Thorax-Simulator

6.3.3.1 Datengewinnung

Es wurden ToF-, Radar- und Laser-Mikrometer-Daten im Abstand von 33,5 cm bzw. 23,5 cm zwischen ToF-Kamera und dem Tisch aufgenommen. Diese Abstände entsprechen den Verhältnissen in zwei möglichen Inkubator-Optionen. Folgende Atemfrequenzen wurden einmal bei tiefer (*deep*) und einmal bei flacher (*normal*) Atmung aufgenommen: 5, 10, 20, 30, 40, 45, 50, 55, 60 und 80 BPM. Bei der tiefen Atmung liegt ein etwa doppelt so hoher Hub wie bei der flachen Atmung vor. Für den finalen Testaufbau wurden zusätzlich schwarze Oberflächen des Simulators mit weißem Papier abgedeckt, um Reflektionen zu vermeiden.

Tabelle 6.2 zeigt die gemessenen Atemhübe (Spitze-Spitze) durch das Laser-Mikrometer im *deep* und *normal* Modus. Der Atemhub ist abhängig von der eingestellten Atemfrequenz. Zudem ist der Atemhub proportional zur eingestellten Spannung des Baby-Simulator [Gle+21].

Tab. 6.2: Gemessene Atemhübe in mm mit dem Laser-Mikrometer im *deep* und *normal* Modus [Gle+21].

BPM	Atemhub (mm) <i>deep</i> Mode	Atemhub (mm) <i>normal</i> Mode
5	0,305	0,136
10	0,644	0,273
20	1,344	0,589
30	1,965	0,928
40	2,591	1,285
45	2,797	1,394
50	3,047	1,66
55	3,285	1,72
60	3,472	1,917
80	4,068	2,481

6.3.3.2 Ergebnisse der synchronen Auswertung Baby-Thorax-Simulator

Es wurde bereits erwähnt, welche Daten mit dem Radar-Sensor, der ToF-Kamera und dem Laser-Mikrometer aufgenommen wurden. In diesem Abschnitt sollen die Ergebnisse präsentiert werden. Im ersten Schritt wurden die Algorithmen mit allen aufgezeichneten Datensätzen getestet und die ermittelten BPM-Werte der ToF- und der Radar-Methode verglichen. Im zweiten Schritt wurden die Distanzsignale aller Messsysteme synchron evaluiert.

BPM Vergleich Die Box-Whisker-Plots der BPM-Werte im *deep* Modus in Abbildung 6.5 und Abbildung 6.6 und im *normal* Modus in Abbildung 6.7 und Abbildung 6.8 zeigen den Median, Minimum, Maximum, oberes und unteres Quartil und Ausreißer (mit Kreisen markiert). Die roten Kreuze markieren die Positionen der Referenz-BPM-Werte, die durch das Laser-Mikrometer gemessen wurden.

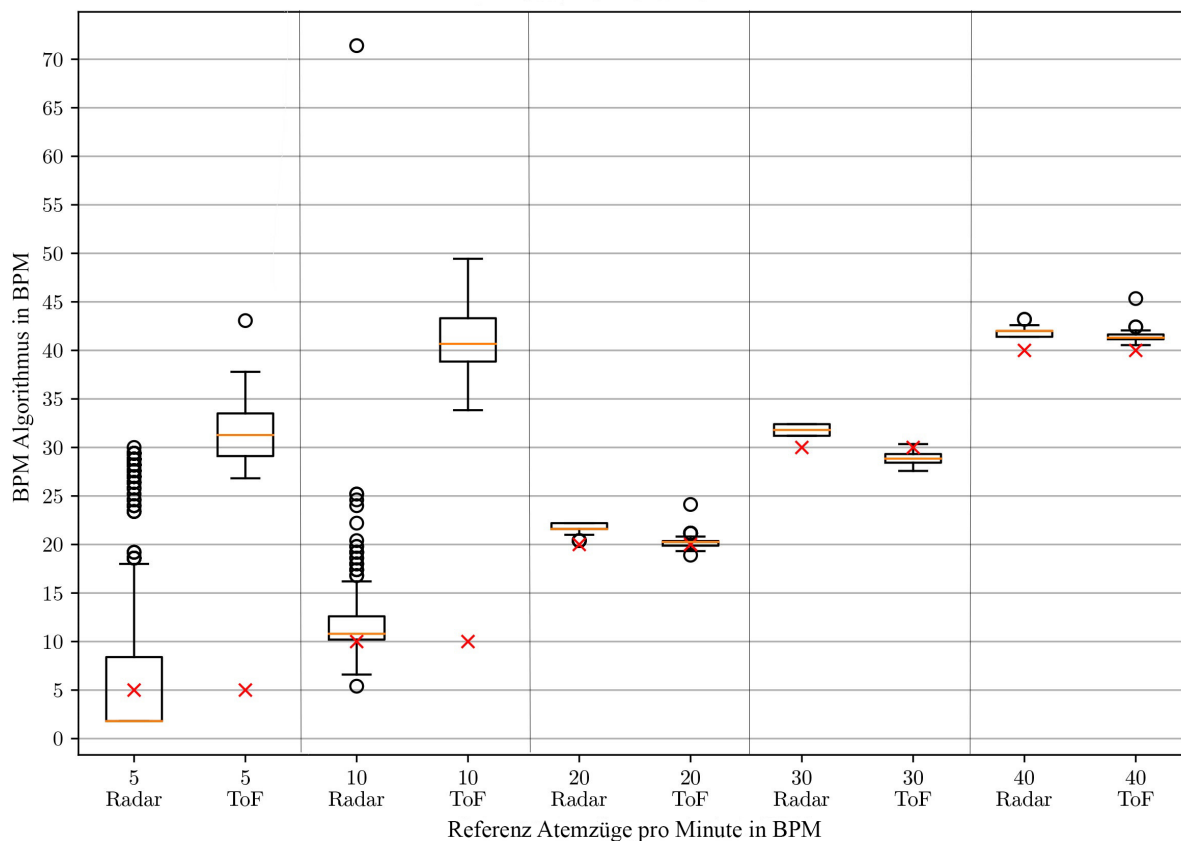


Abb. 6.5: Box-Plot, bei dem die BPM-Werte des ToF- und des Radar-Algorithmus für Atemraten zwischen 5 BPM und 40 BPM im *deep* Modus verglichen werden [Gle+21].

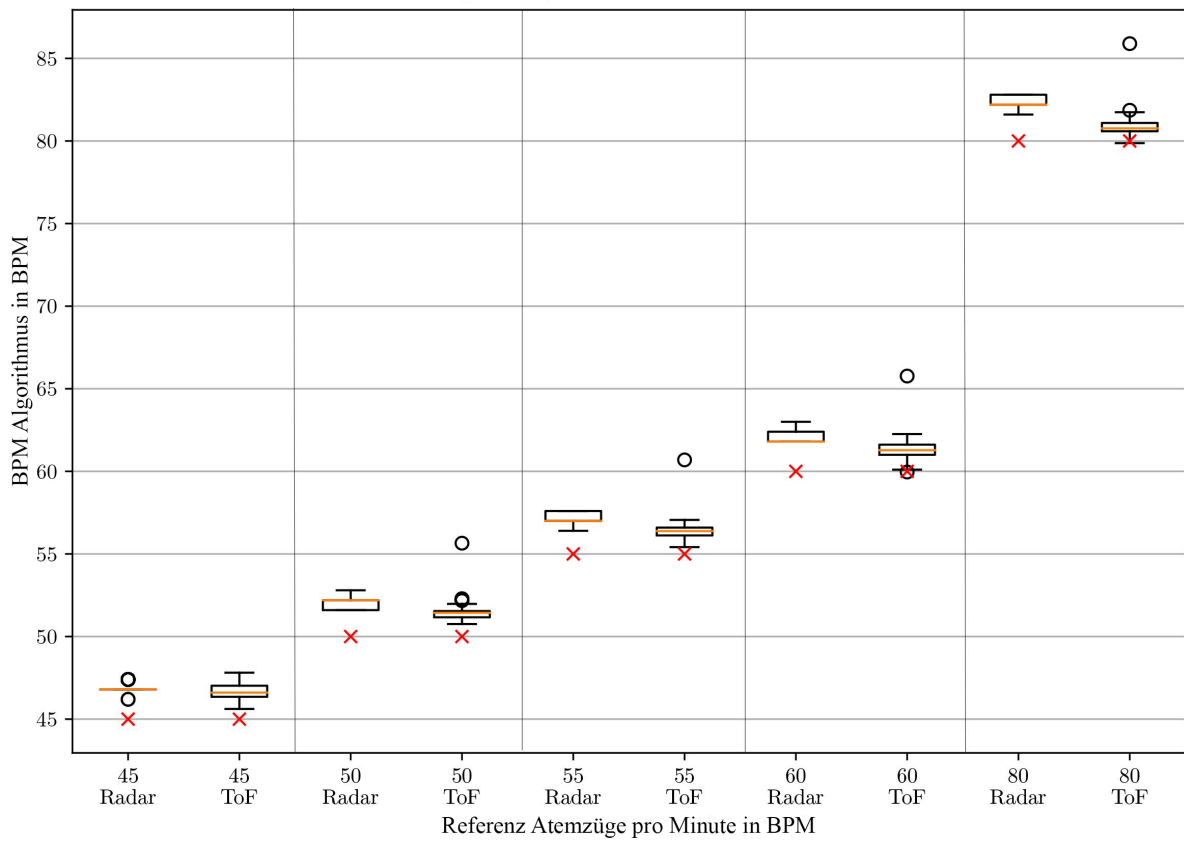


Abb. 6.6: Box-Plot, bei dem die BPM-Werte des ToF- und des Radar-Algorithmus für Atemraten zwischen 45 BPM und 80 BPM im *deep* Modus verglichen werden [Gle+21].

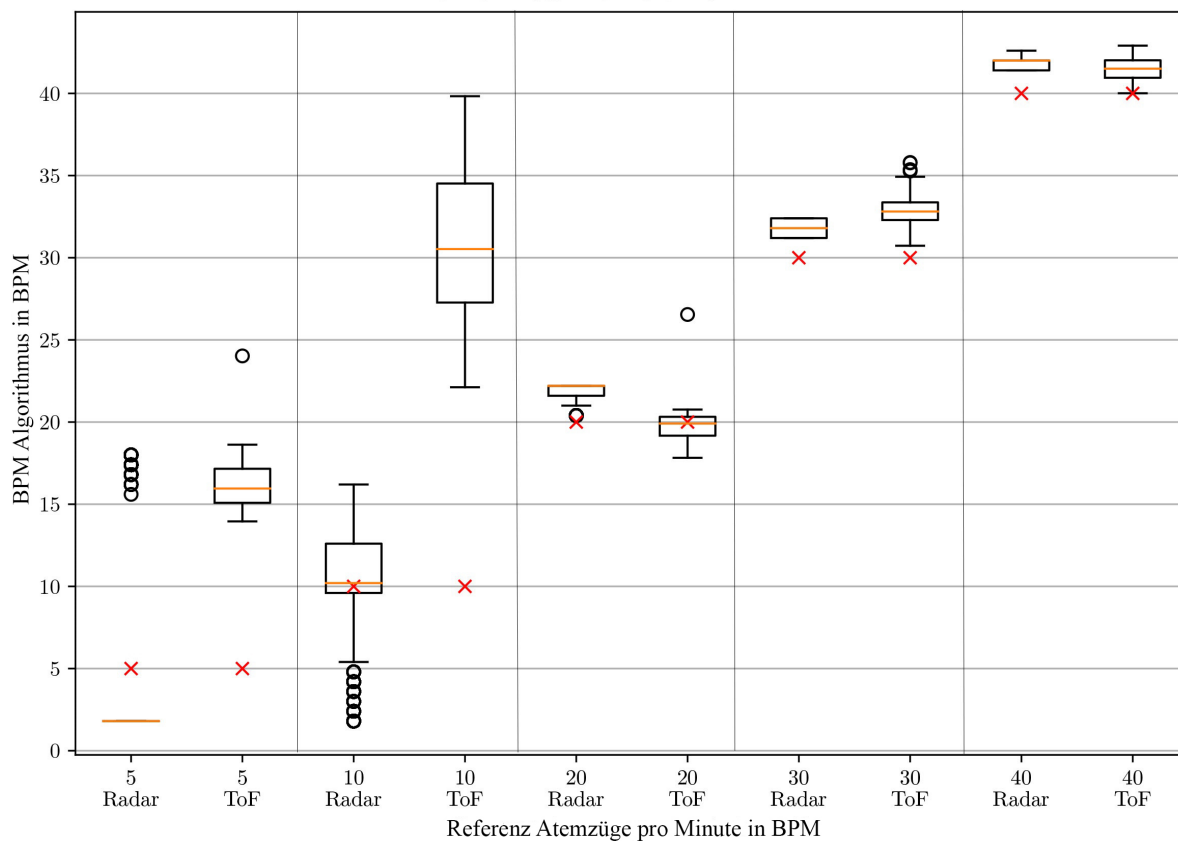


Abb. 6.7: Box-Plot, bei dem die BPM-Werte des ToF- und des Radar-Algorithmus für Atemraten zwischen 5 BPM und 40 BPM im *normal* Modus verglichen werden [Gle+21].

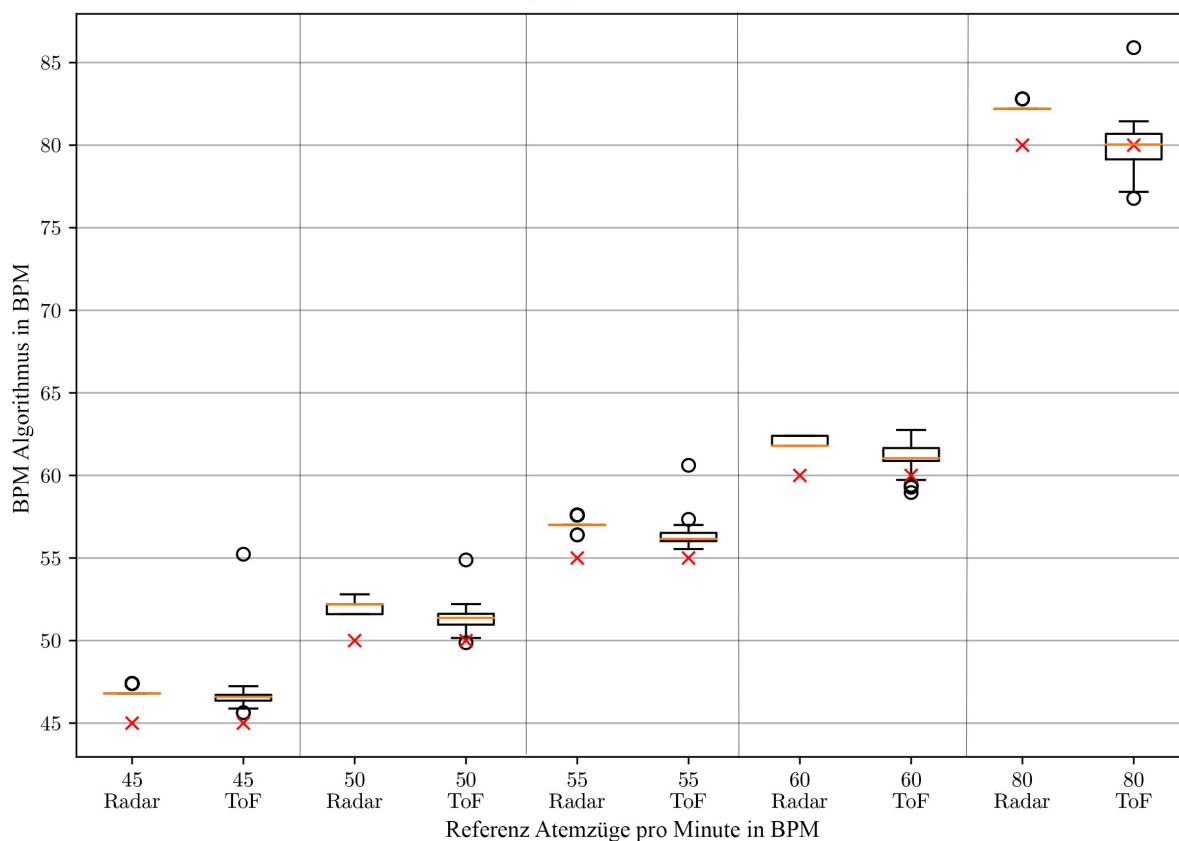


Abb. 6.8: Box-Plot, bei dem die BPM-Werte des ToF- und des Radar-Algorithmus für Atemraten zwischen 45 BPM und 80 BPM im *normal* Modus verglichen werden [Gle+21].

Wie zu erkennen ist, liegen für die ToF-Kamera bei niedrigen Atemraten (5 und 10 BPM) große Unterschiede zur Referenz vor. Im *deep* Modus wird ein Unterschied von 27 bzw. 31 BPM zu den eingestellten Atemraten erhalten. Im *normal* Modus liegt ein Unterschied von 11 bzw. 21 BPM vor. Das Ergebnis hängt vom ersten BPM-Wert ab, welcher die Parametrierung des Savitzky-Golay-Filters bestimmt. Bei 5 und 10 BPM im *deep* Modus und 10 im *normal* Modus sind die Startwerte größer 32 BPM, was bedeutet, dass kein Filter angewandt wird. Bei 5 BPM im *normal* Modus liegt der Startwert bei 24,02 BPM, was bedeutet, dass der Savitzky-Golay-Filter mit Fenstergröße 3 und Grad des Polynoms von 2 angewandt wird. Dies entspricht einer stärkeren Filterung. Wegen des starken Rauschens des Distanzsignals im niedrigen Frequenzbereich erkennt die Peak-Detektion zu viele Peaks. Ein weiterer Aspekt, welcher den Startwert beeinflusst, ist die Zeit, über die die Peak-Detektion angewandt wird. Beim Testen auf zwei verschiedenen Computern fiel auf, dass die Zeit für dieselbe Anzahl an Datenpunkten unterschiedlich war. Dies liegt an den ROS-Prozessen, die im Hintergrund laufen und zu Zeitvariationen führen. Selbst wenn auf nur einem Computer gearbeitet wird, kann die Zeit, um ein Datenpaket von 33 Datenpunkten zu verarbeiten, leicht variieren.

Zwischen 20 und 80 BPM sind die Ergebnisse mit dem ToF-Ansatz etwas besser als die Radar-Ergebnisse. Für kleine Frequenzen liefert der Radar-Ansatz bessere Ergebnisse und eine maximale Abweichung von 3 BPM im Vergleich zur Referenz, aber mit großen Ausreißern. Der Vergleich wurde zu dem Median der Werte vorgenommen.

Distanzsignal Vergleich In diesem Kapitel werden die Distanzsignale verglichen, die durch das Laser-Mikrometer und den ToF- bzw. Radar-Algorithmus erhalten werden. In Abbildung 6.9 ist ein Abschnitt des Atemsignals bei 10 BPM im *deep* Modus zu sehen. Zu Beginn ist das Synchronisationssignal zu erkennen. Die Signale sind ungefiltert, was im Kontext der ToF-Kamera bedeutet, dass der Savitzky-Golay-Filter noch nicht angewandt wurde. Der erste Abschnitt nach dem Synchronisationssignal wird zur Berechnung der ersten BPM für die Parametrierung des Savitzky-Golay-Filters verwendet. Diese basiert auf einer Peak-Detektion. Da das Signal stark verrauscht ist, wird der detektierte BPM-Wert hoch sein (vergleiche BPM Box-Plots). Das Radar-Signal ist ebenfalls verrauscht, aber mit dem FFT-Ansatz ist es leichter, die zugrundeliegende Frequenz zu detektieren. Um das dargestellte Radar-Signal zu erhalten, wurde nachträglich ein Ellipse-Fitting zweimal auf die Daten angewandt. Die IQ-Werte des Radar-Sensors werden auf eine Ellipse positioniert, um zur Kalibrierung die Transformation zum IQ Einheitskreis berechnen zu können. Die Implementierung basiert auf einer MATLAB Implementierung von Gal [Gal].

In Abbildung 6.10 ist das Atemsignal bei 45 BPM im *deep* Modus geplottet. Das ToF-Signal ähnelt dem Referenz Laser-Mikrometer-Signal und ist in Phase. Die Radar-Amplitude wirkt asymmetrisch, was aber keinen Effekt auf die FFT und die Berechnung der BPM hat. In diesem Fall wurde das Ellipse-Fitting auch nachträglich angewandt. Die Größe der Radar-Amplitude ist davon abhängig, ob Ellipse-Fitting angewandt wurde oder nicht.

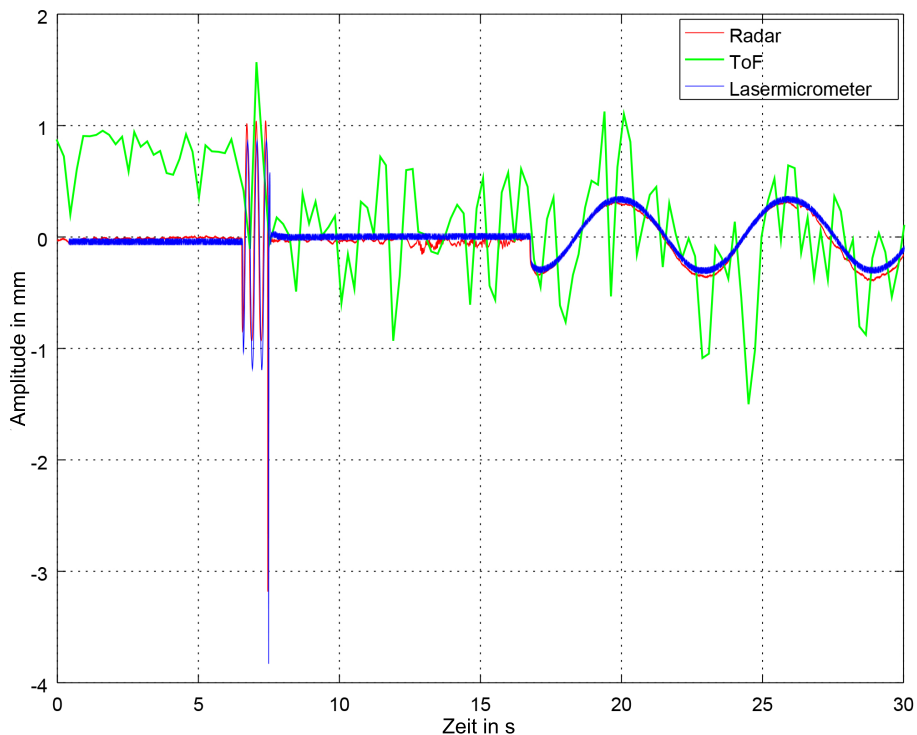


Abb. 6.9: Ungefilterter Abschnitt des Atemsignals bei 10 BPM im *deep* Modus, gemessen mit der ToF-Kamera, dem Radar-Sensor und dem Laser-Mikrometer. Im Kontext der ToF-Kamera bedeutet "ungefiltert", dass der Savitzky-Golay-Filter noch nicht angewandt wurde. Zu Beginn ist das Synchronisationssignal zu erkennen. Das ToF-Kamerasignal ist stark verrauscht und hat große Amplitudensprünge. Das Radar- und Laser-Mikrometer-Signal liegen nahe aneinander. [Gle+21]

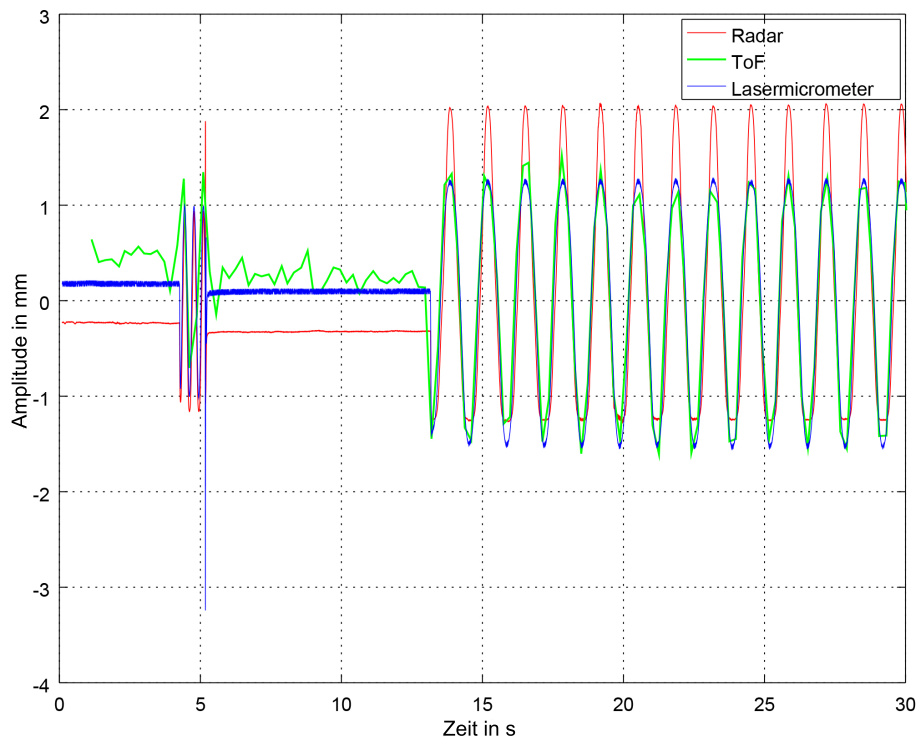


Abb. 6.10: Noch ungefilterter Abschnitt des Atemsignals bei 45 BPM im *deep* Modus, gemessen mit der ToF-Kamera, dem Radar-Sensor und dem Laser-Mikrometer. Zu Beginn ist das Synchronisationssignal zu erkennen. Zu Beginn des ToF-Signals ist das Rauschen klar sichtbar. Später ist der Hub des Simulator-Thorax höher als das Rauschen. Die Amplitude ähnelt der des Laser-Mikrometers. Die Amplitude des Radar-Sensors ist leicht verschoben und ungleichmäßig. Dies kann an den Effekten des Ellipsen-Fittings liegen. [Gle+21]

Es wurde mit einem Atemsignal bei 45 BPM im *deep* Modus getestet, wie sich die Amplitude abhängig vom Ellipse-Fitting, einem angewandten Mittelwert, gleitendem Durchschnitt und einem gefensterterem Durchschnitt verhält. Diese wurden jeweils auf das IQ-Signal des Radar-Sensors angewandt. Das Ergebnis ist in Abbildung 6.11 zu sehen. Das Ziel war, die größtmögliche Amplitude zu erreichen. Da der gleitende Durchschnitt und der gefensterter Durchschnitt negative Effekte bei Anwendung auf niedrige Atemfrequenzen hatte und die FFT schlechtere Ergebnisse lieferte, wurde Ellipse-Fitting verwendet. Bisher ist es nicht möglich, eine stabile und passende Ellipse in Echtzeit zu generieren. Aus diesem Grund wurde eine Ellipse aus den gesamten Radar-Daten eines aufgezeichneten Datensatzes berechnet und diese Ellipse wurde nachträglich auf die Radar-Daten angewandt.

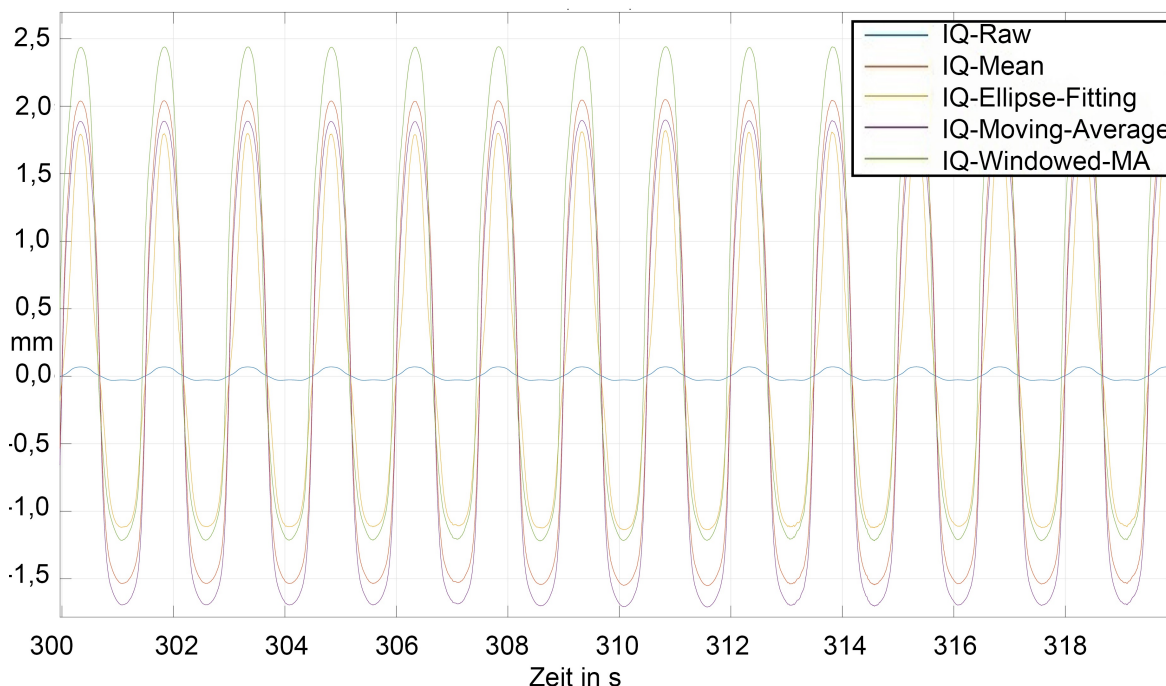


Abb. 6.11: Variierende Radaramplitude, abhängig davon, ob Ellipse-Fitting (gelb), der Mittelwert (rot), gleitender Durchschnitt (lila) oder ein gefensterter Durchschnitt (grün) verwendet wird. In diesem Beispiel wird eine Atemrate von 40 BPM im *deep* Modus verwendet. [Gle+21]

6.3.4 Simulationspuppe Paul

6.3.4.1 Datengewinnung

Für die Aufnahmen mit der ToF-Kamera und dem Radar-Sensor wurde der Frühchen-Simulator Paul genutzt, weil keine Genehmigung der Ethikkommission für die Radar-Aufnahmen am Kind vorlag.

ToF-Kamera Es wurden Datensätze von jeweils 15 Minuten Dauer bei Atemraten von 5, 10, 20, 40, 45, 60, 80 und 100 BPM aufgezeichnet. Zusätzlich wurde das Steigen und Fallen der Rate simuliert, indem die vorher genannten Werte jeweils für eine Minute eingestellt wurden und dann

auf die nächste Rate erhöht wurde. Nachdem 100 BPM erreicht waren, wurde die abfallende Richtung simuliert. Alle Aufnahmen wurden als normale Atmung und mit Respiratory Distress Syndrome (RDS) aufgenommen. Als Referenz wurde zu Beginn mit dem Laser-Mikrometer die Atemrate bestimmt, dabei fielen bei hohen Raten kleine Abweichungen (max. 3,33 % Abweichung bei normaler Atmung, 11,7 % Abweichung bei RDS) zwischen dem eingestellten und dem tatsächlichen Wert auf.

Radar-Sensor Es wurden Atemraten bei 5, 10, 20, 40, 45, 60, 80 und 100 BPM für jeweils 15 Minuten aufgenommen. Wie bei der ToF-Kamera wurden ändernde Raten simuliert und dies jeweils als normale Atmung und mit RDS.

6.3.4.2 Ergebnisse Simulationspuppe Paul

Auswertung ToF-Daten zur Bestimmung der Atemrate In den folgenden Box-Plots sind die Ergebnisse der Atemdetektion der ToF-Kamera an der Simulationspuppe Paul im Vergleich zur voreingestellten Atemrate dargestellt. Es fiel auf, dass die eingestellte Atemrate nicht immer der tatsächlichen Atemrate (rot) entsprach. Wie zu erkennen ist, liefert die ToF-Kamera für beide Atemmodi keine den klinischen Anforderungen entsprechenden Abweichungen kleiner ± 2 BPM.

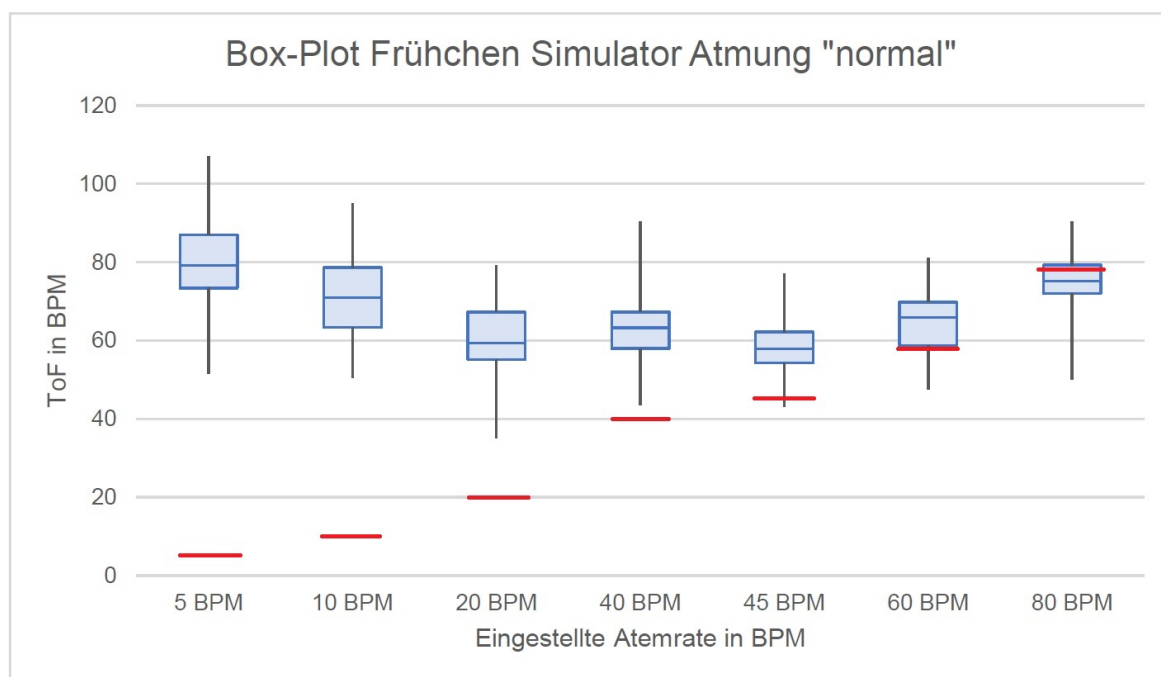


Abb. 6.12: Box-Plot zum Vergleich der Atemraten von der Time-of-Flight-Kamera (blau) und der eingestellten Atemrate (normal) des Simulators nach Frequenzen ausgewertet.

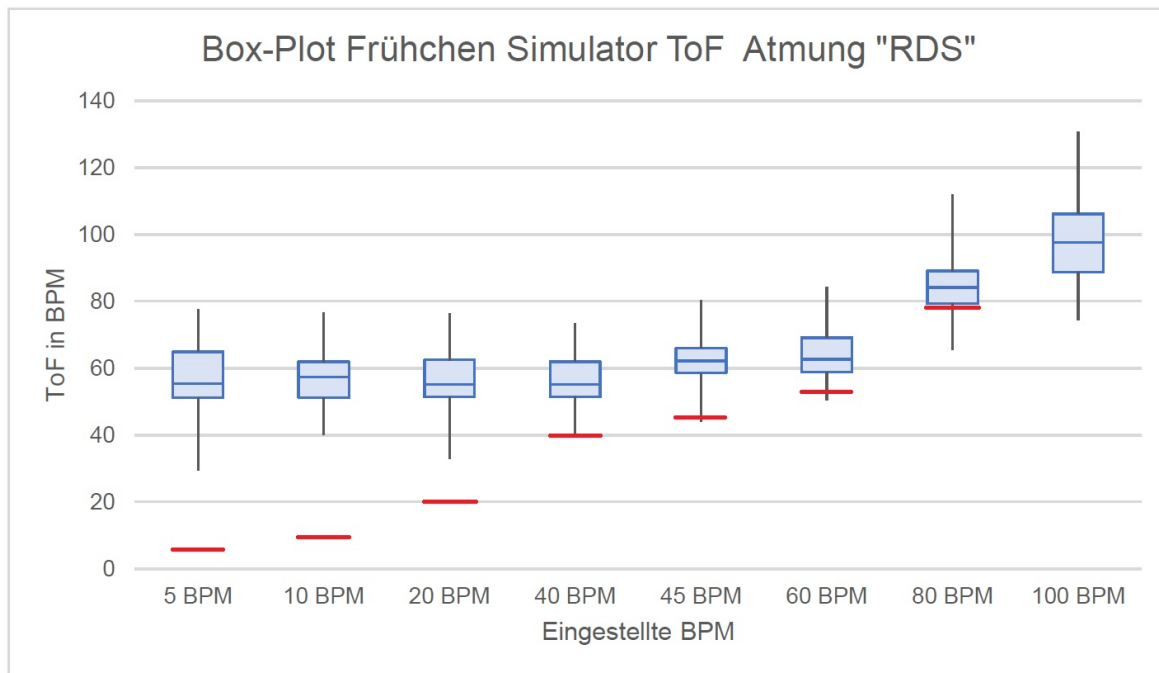


Abb. 6.13: Box-Plot zum Vergleich der Atemraten von der Time-of-Flight-Kamera (blau) und der eingestellten Atemrate RDS des Simulators nach Frequenzen ausgewertet.

Es wurde erörtert, wodurch die Ungenauigkeiten entstehen und herausgefunden, dass anstelle des Thorax die Windel des Simulators detektiert wird, weil die Windel den geringsten Abstand zur Kamera hat. Für die Detektion des Torsos in den Bilddaten der Studien-Probanden wurde ein Neuronales Netz trainiert, das den Thorax von der Windel unterscheiden kann. Bilddaten von Paul wurden nicht in das Neuronale Netz eingespeist.

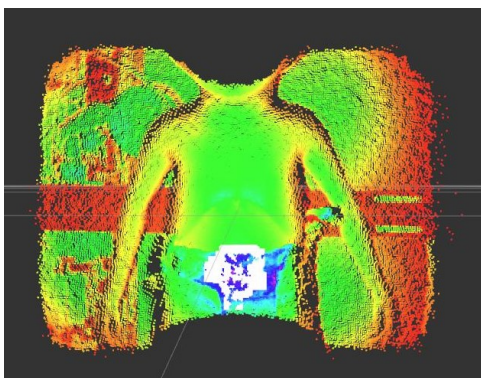


Abb. 6.14: Es wird anstelle des Thorax die Windel des Simulators detektiert, da die Windel den geringsten Abstand zur Kamera hat.

Genauigkeit der manuellen Zählung Die Genauigkeit der manuellen Zählung aus dem RGB-Bild mithilfe des Matlab Codes wurde mit den Ergebnissen des Lasermikrometers bei 10, 40 und 80 BPM verglichen. Bei eingestellten 10 BPM ist die manuelle Zählung 0,93 % höher als das Lasermikrometer, bei 40 BPM 3,61 %. Bei am Simulator voreingestellten 80 BPM misst das Lasermikrometer 5 % weniger, die manuelle Zählung 1 % mehr.

Auswertung Radar-Daten zur Bestimmung der Atemrate Die Ergebnisse mit dem OneFFT4all Algorithmus sind schlecht. Es fällt auf, dass häufig der Wert 1,8 BPM auftritt. Dies könnte durch einen Gleichanteil entstehen, wenn z.B. die Aufnahme gestartet wird, der Simulator sich aber nicht in der "Nullposition" befindet. Es wurde ein Code geschrieben, um den Gleichanteil rechnerisch zu entfernen (Mittelwert über Fenster und diesen Wert von allen Distanzwerten abziehen). Damit lässt sich der Gleichanteil nicht komplett eliminieren, die Ergebnisse mit OneFFT4all sind aber leicht besser als zuvor. Grundsätzlich sind die Messergebnisse am Simulatorbaby schlechter als an dem Baby-Thorax-Simulator.

Laut Hersteller des Simulatorbabies sind elektronische Bauteile verbaut, die den Radar-Sensor durch elektromagnetische Effekte stören können. Dies kann die obigen Ergebnisse erklären, sie sind also zur Beurteilung der Methode nur eingeschränkt geeignet.

6.4 Detektion von Körperarealen mit Neuronalen Netzen

Das Training der Neuronalen Netze erfolgte in mehreren Schritten. Zunächst wurden nur die Daten von vier Probanden gelabelt und zum Training verwendet. Der Trainingsverlauf und die erreichten Average Precisions (AP) sind in den folgenden Graphen ?? bis ?? jeweils in der oberen Zeile und Tabelle 6.5 ersichtlich.

6.4.1 Training der Neuronalen Netze

Die ersten Trainings-Datensätze beinhalten 1400 gelabelte Bilder plus 1400 augmentierte Bilder pro Modalität von den Probanden 06 bis 09. Diese Daten bilden die Grundlage für den Lernprozess des Neuronalen Netzes. Sie werden wiederholt durch das Modell ausgewertet, um eine optimale Anpassung des Netzes zu erreichen. Der zugehörige Validierungs-Datensatz enthält 300 Bilder von jeder Modalität von den Probanden 06 bis 09, die nicht im Trainings-Datensatz vorhanden sind. Die Validierungsdaten werden während des Trainings verwendet, um erste Tests bezüglich Genauigkeit durchzuführen und gleichzeitig neue Daten hinzuzufügen, die das Neuronale Netz bis dahin nicht kannte. Tabelle 6.3 und Tabelle 6.4 zeigen die Anzahl der Instanzen pro Klasse pro Modalität. In jedem verwendeten Bild sind alle erkennbaren Klassen gelabelt. Die Klasse "nose" ist wegen der Augenbinde, die während der Phototherapie getragen wird, häufig nicht sichtbar.

Tab. 6.3: Anzahl von Instanzen pro Klasse pro Modalität im Trainings-Datensatz (Probanden 06 bis 09)

Modality	head	nose	torso	intervention
RGB	8680	5292	7983	1085
Thermo	8487	2712	8520	762
Fusion	8285	100	7462	753

Jedes Netzwerk wurde für 64 Epochen trainiert und die Epoche mit der besten Genauigkeit wurde bestimmt. Die Average Precisions des Validierungsdatensatzes in Tabelle 6.5 zeigen

Tab. 6.4: Anzahl von Instanzen pro Klasse pro Modalität im Validierungs-Datensatz (Probanden 06 bis 09)

Modality	head	nose	torso	intervention
RGB	1193	575	1129	160
Thermo	1199	305	1183	123
Fusion	1190	8	1055	80

deutlich, dass im Fusions-Trainingsdatensatz nicht genügend Bilder der Klassen "nose" und "intervention" vorhanden waren, um die Neuronale Netze erfolgreich zu trainieren.

Tab. 6.5: Average Precision pro Klasse pro Modalität pro Netzwerk für Trainingsphase 1 (Validierungs-Datensatz)

*Die Genauigkeitswerte für "nose" sind wegen einer unzulänglichen Anzahl von Klasseninstanzen im Fusions-Datensatz nicht aussagekräftig.

	Modalität	Beste Epoche	Average Precision			
			head	nose	torso	intervention
RetinaNet	RGB	25	1,0	0,9937	0,99	0,94
	Thermo	28	0,9969	0,9864	0,9862	0,8695
	Fusion	38	0,9949	0,0*	0,9934	0,7683
YOLOv3	RGB	64	1,0	0,9885	0,9991	0,9821
	Thermo	61	0,9983	0,9993	0,9963	0,9225
	Fusion	56	0,9949	0,3274*	0,9948	0,8390

Die Aufnahmen von acht weiteren Probanden wurden sukzessive hinzugefügt, zum Schluss wurde jedes Netz mit allen Trainingsdaten trainiert. Pro Proband wurden 1400 Bilder (plus Augmentierung) verwendet. Da Proband 14 nicht genügend Daten beinhaltete, wurden diese mit Daten von Proband 3 kombiniert. Das RGB-Netz wurde zusätzlich mit Time-of-Flight Mono-Bildern angereichert. Die damit erreichten Average Precisions können Tabelle 6.8 entnommen werden. Insbesondere für die Klassen "nose" und "intervention" wurde eine deutliche Verbesserung erreicht.

Tab. 6.6: Anzahl von Instanzen pro Klasse pro Modalität im Trainings-Datensatz weitere Probanden

Modality	head	nose	torso	intervention
RGB	21692	15366	19637	3174
Thermo	16981	5959	15877	2593
Fusion	16755	4283	14712	2007
ToF	5822	1317	8345	612

Werden die Ergebnisse bei dem Training mit weiteren Probanden und nur mit den vier anfänglichen Probanden verglichen, so werden grundsätzlich ähnliche Average Precisions erhalten. Auffallend ist eine deutliche Verbesserung mit dem Fusions-Netzwerk bei der Klasse "nose". Dies liegt an der höheren Anzahl an Instanzen im Trainingsdatensatz (siehe Tabelle 6.6).

Tab. 6.7: Anzahl von Instanzen pro Klasse pro Modalität im Validierungs-Datensatz weitere Probanden

Modality	head	nose	torso	intervention
RGB	2971	1744	2726	470
Thermo	2373	663	2211	342
Fusion	2354	463	2032	281
ToF	843	127	1158	83

Tab. 6.8: Average Precision pro Klasse pro Modalität pro Netzwerk für Trainingsphase 2 (Validierungs-Datensatz)

	Modalität	Beste Epoche	Average Precision			
			head	nose	torso	intervention
RetinaNet	RGB	43	0,9977	0,9904	0,9957	0,9004
	Thermo	34	0,9983	0,9564	0,9914	0,9100
	Fusion	29	0,9968	0,9537	0,9762	0,8143
	ToF	26	0,9320	0,9425	0,9285	0,7719
YOLOv3	RGB	64	0,9983	0,9918	0,9958	0,9640
	Thermo	61	1,0	0,9940	0,9962	0,9560
	Fusion	60	0,9983	0,9935	0,9947	0,9514
	ToF	59	0,9327	0,9526	0,9289	0,8990

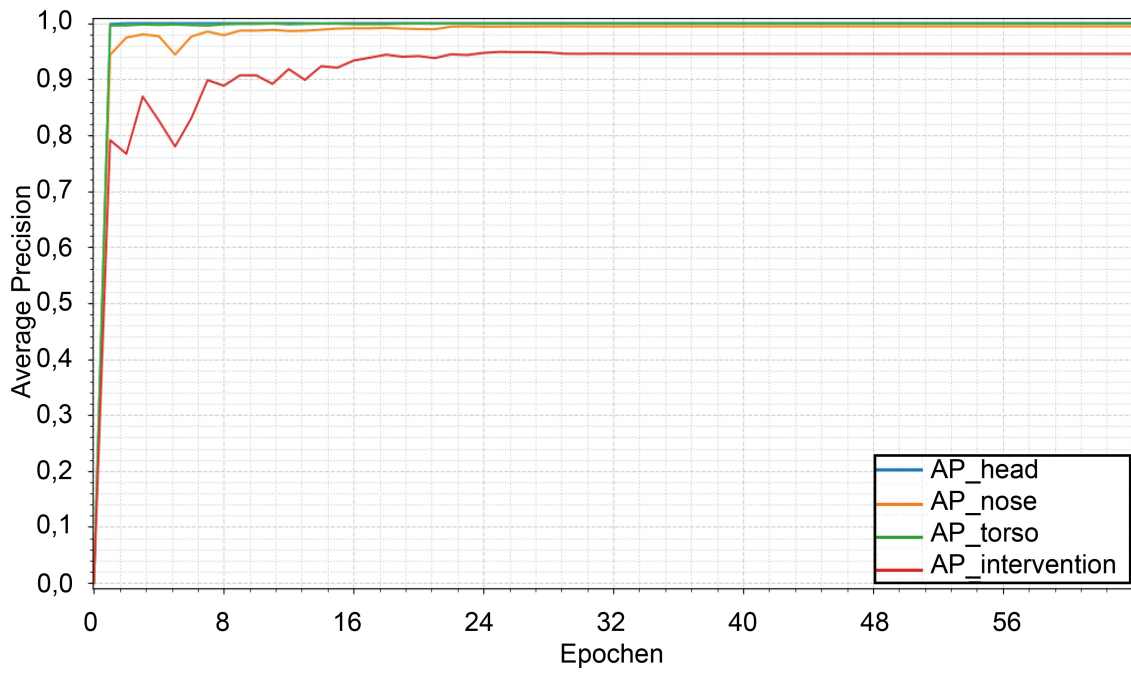


Abb. 6.15: Trainingsverlauf Phase 1 RGB-Daten für RetinaNet, AP per Epoche

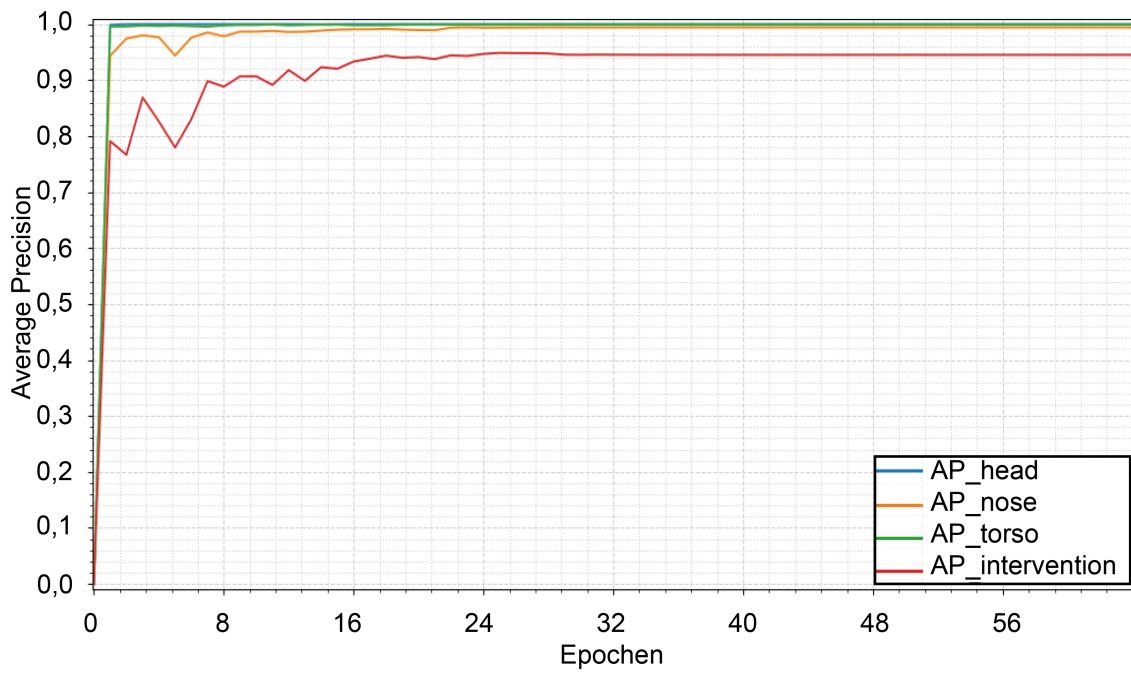


Abb. 6.16: Trainingsverlauf Phase 1 RGB-Daten für YOLOv3, AP per Epoche

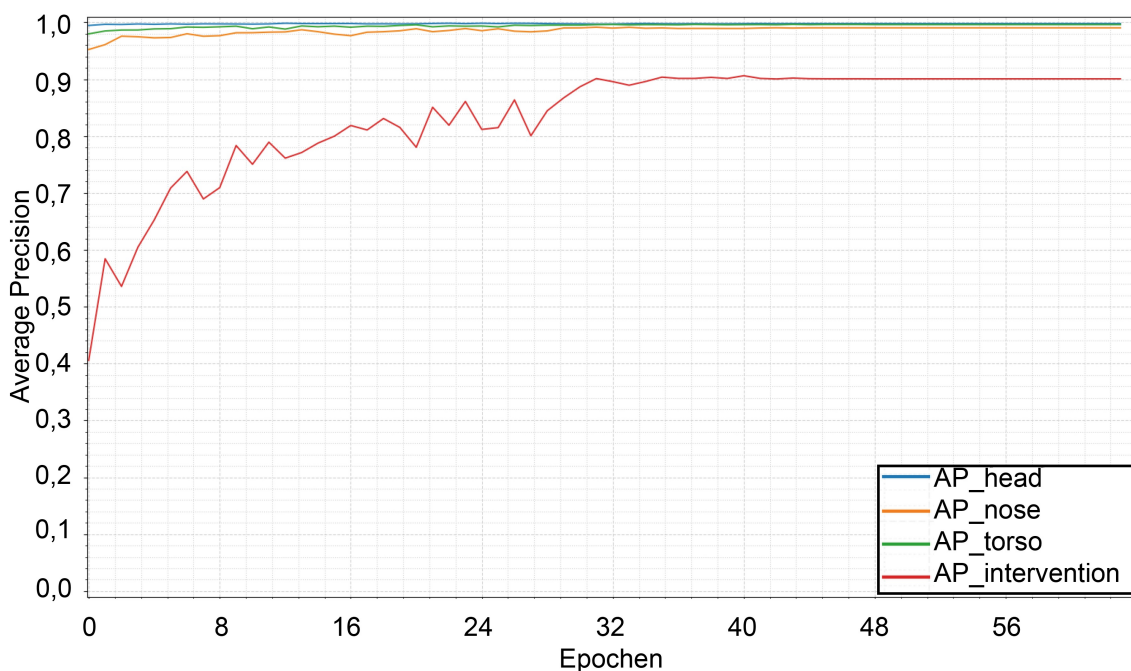


Abb. 6.17: Trainingsverlauf Phase 2 RGB-Daten für RetinaNet, AP per Epoche

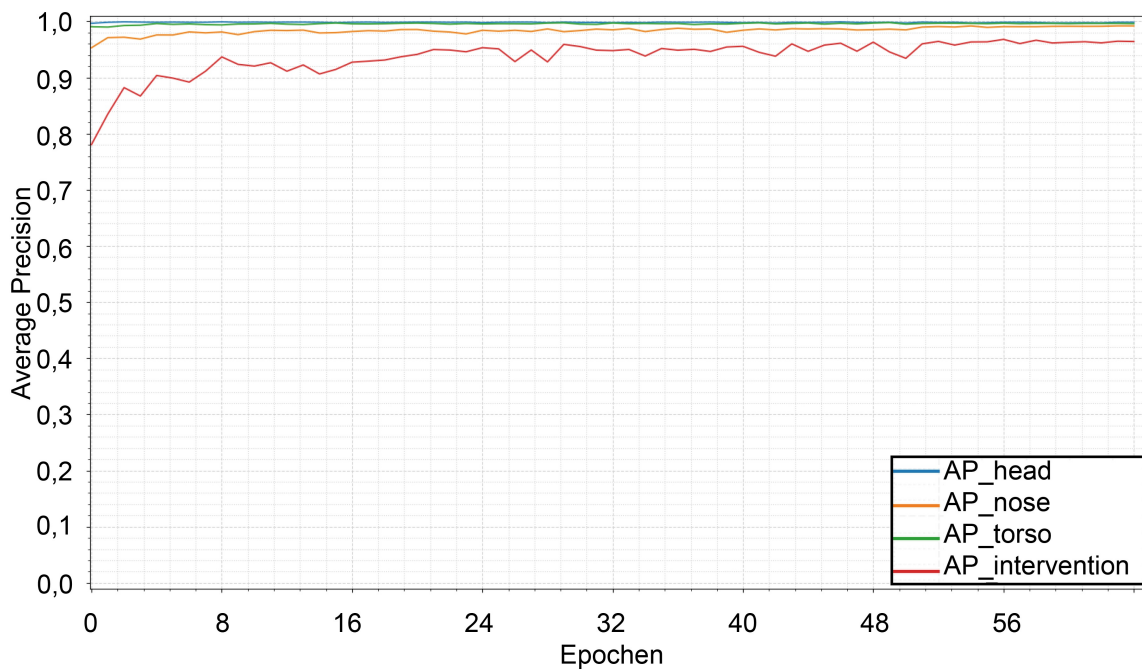


Abb. 6.18: Trainingsverlauf Phase 2 RGB-Daten für YOLOv3, AP per Epoche

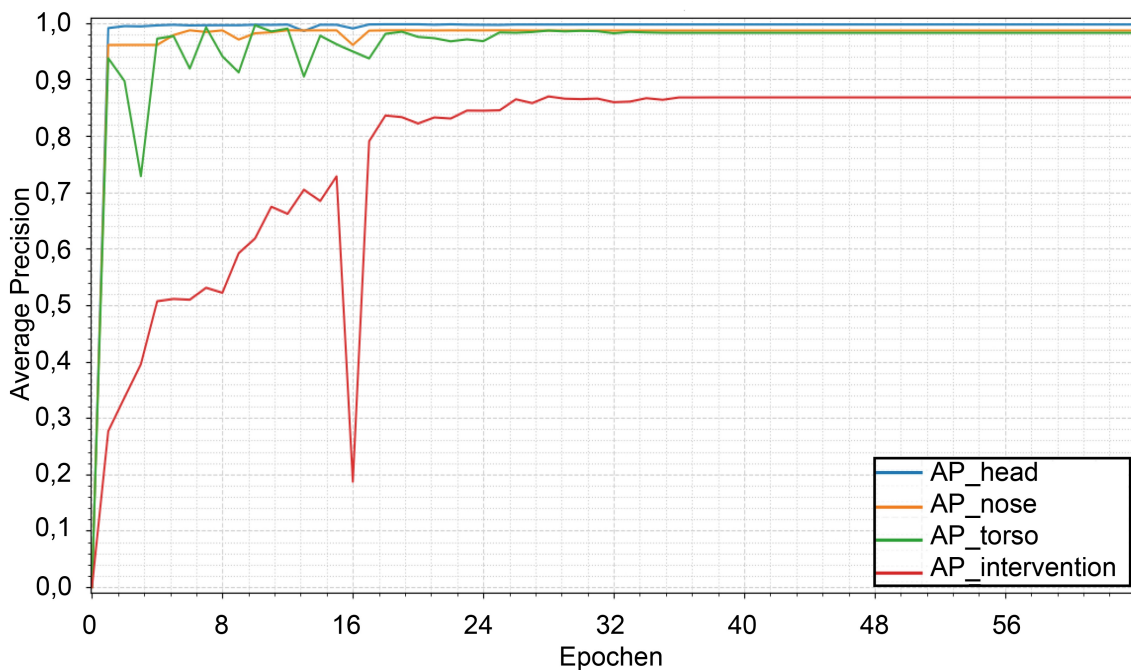


Abb. 6.19: Trainingsverlauf Phase 1 Thermo-Daten für RetinaNet, AP per Epoche

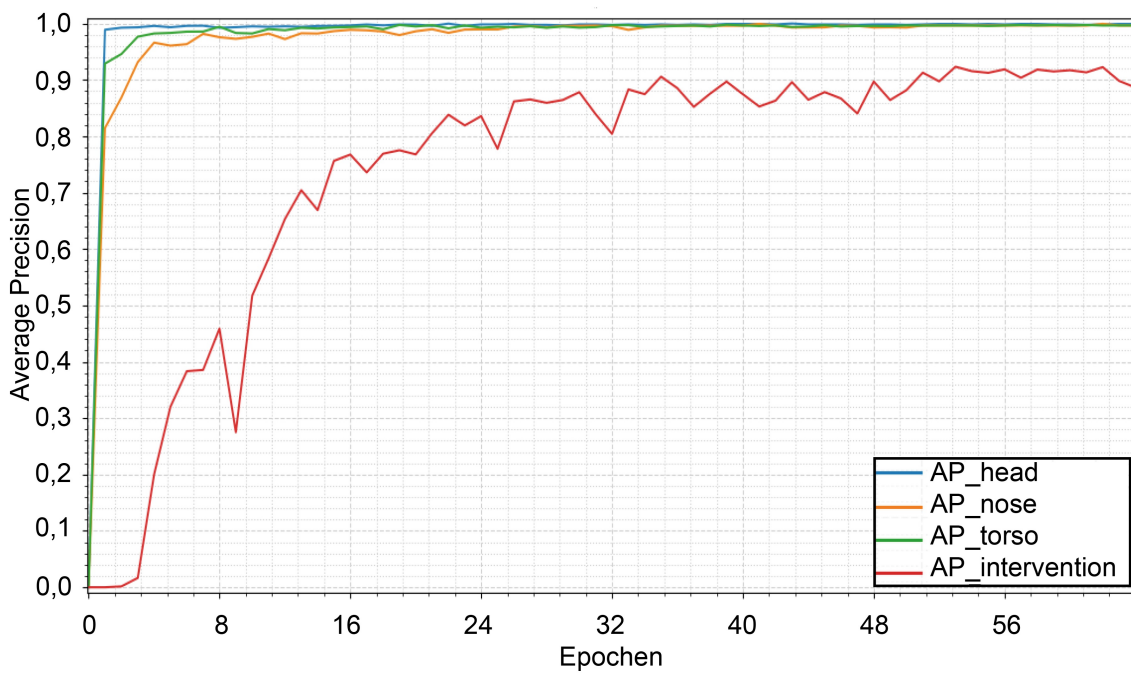


Abb. 6.20: Trainingsverlauf Phase 1 Thermo-Daten für YOLOv3, AP per Epoche

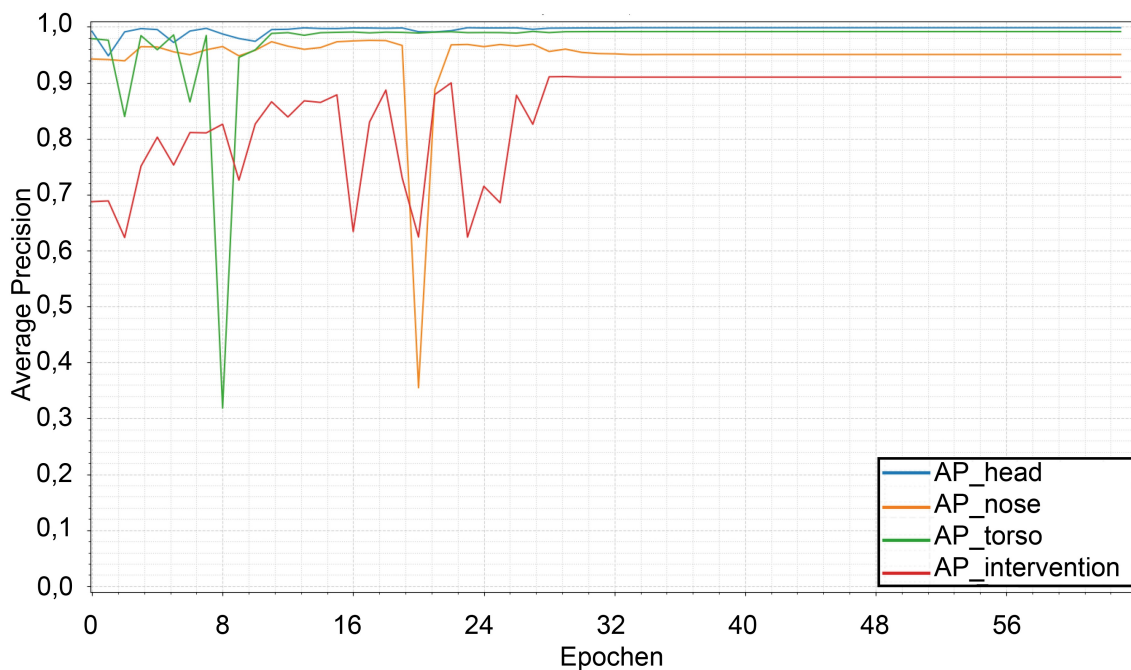


Abb. 6.21: Trainingsverlauf Phase 2 Thermo-Daten für RetinaNet, AP per Epoche

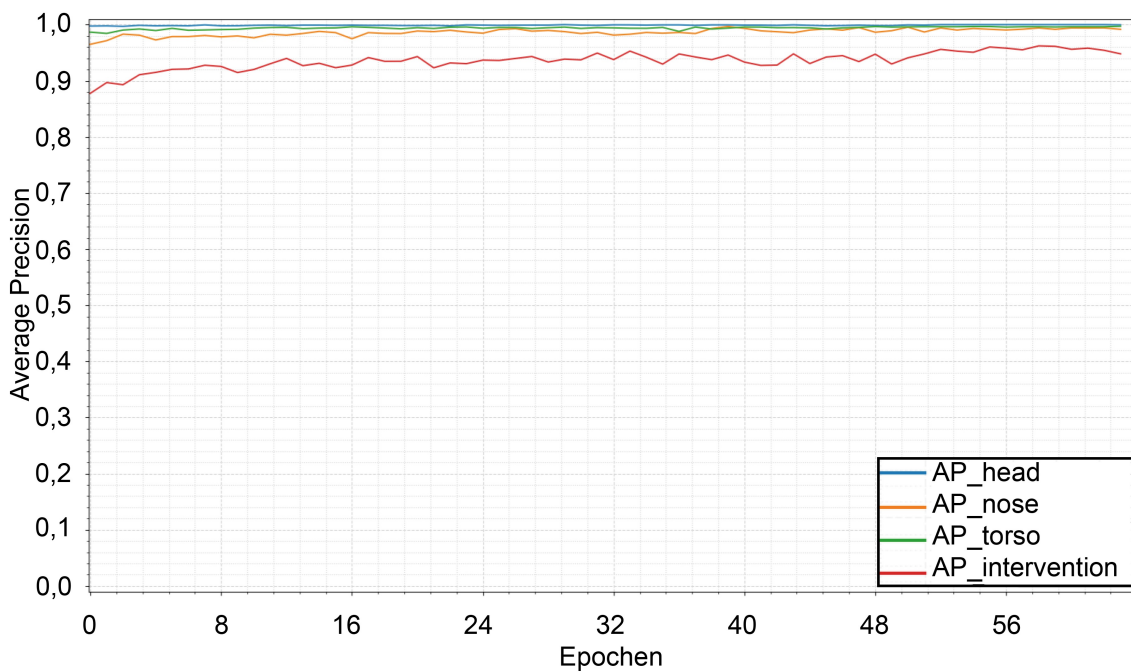


Abb. 6.22: Trainingsverlauf Phase 2 Thermo-Daten für YOLOv3, AP per Epoche

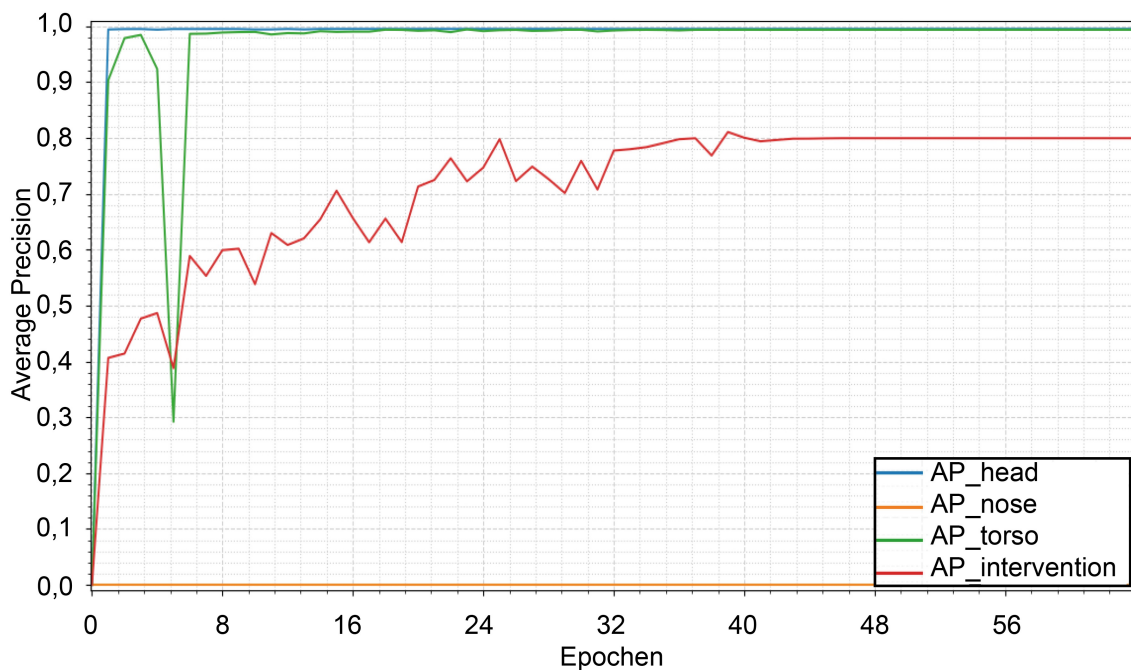


Abb. 6.23: Trainingsverlauf Phase 1 Fusions-Daten für RetinaNet, AP per Epoche

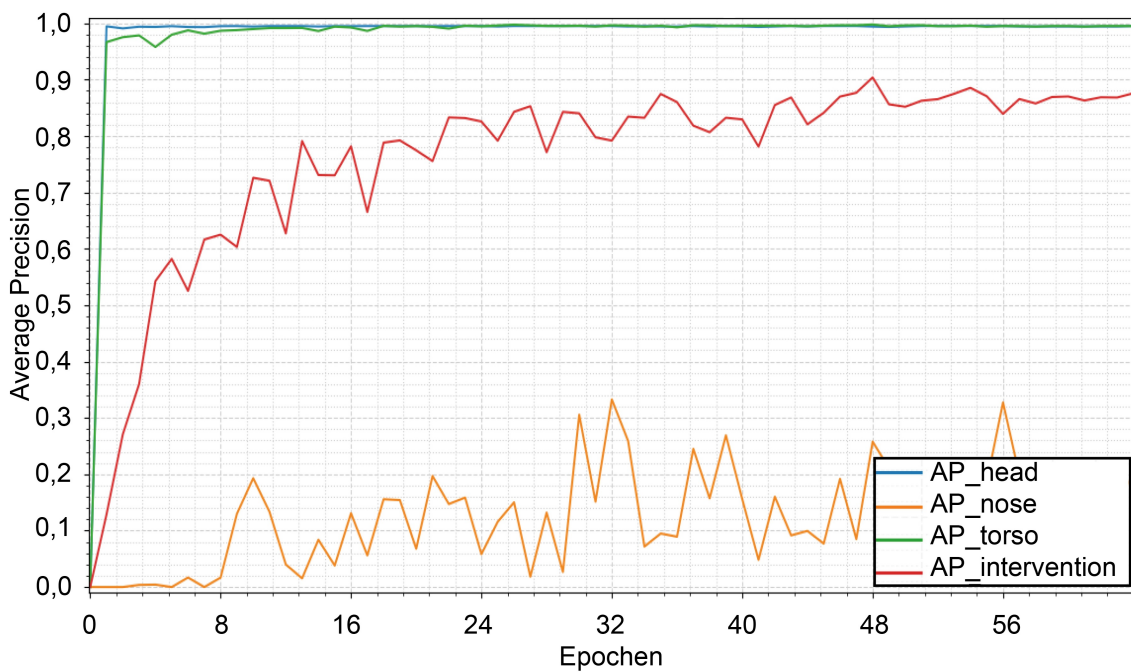


Abb. 6.24: Trainingsverlauf Phase 1 Fusions-Daten für YOLOv3, AP per Epoche

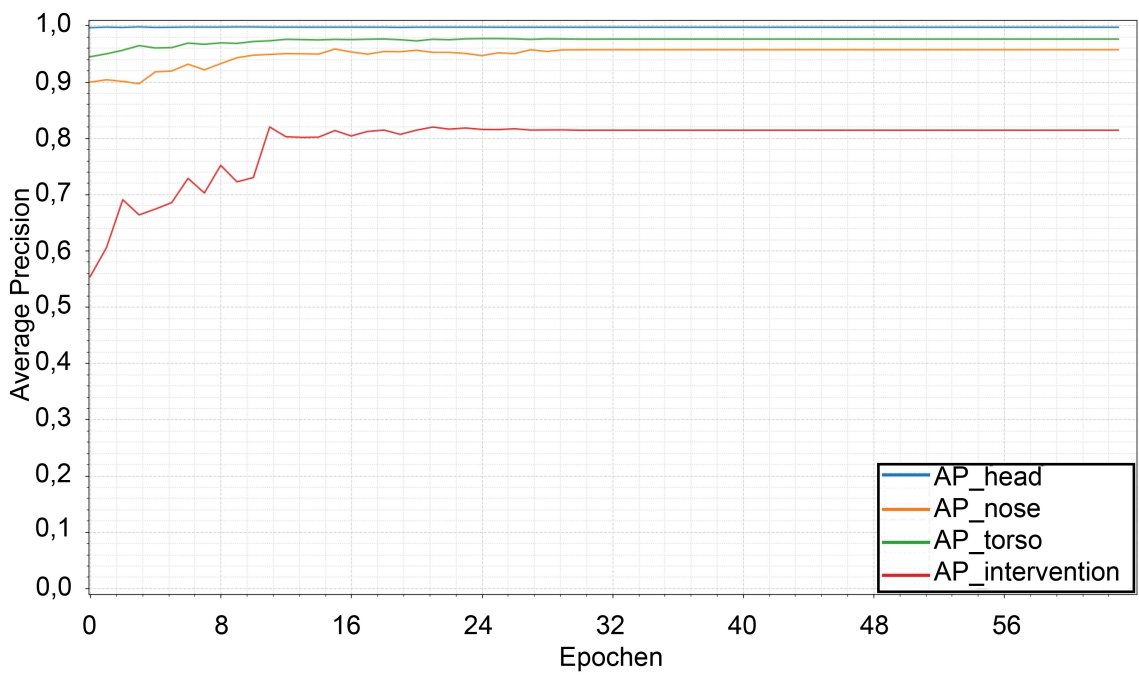


Abb. 6.25: Trainingsverlauf Phase 2 Fusions-Daten für RetinaNet, AP per Epoche

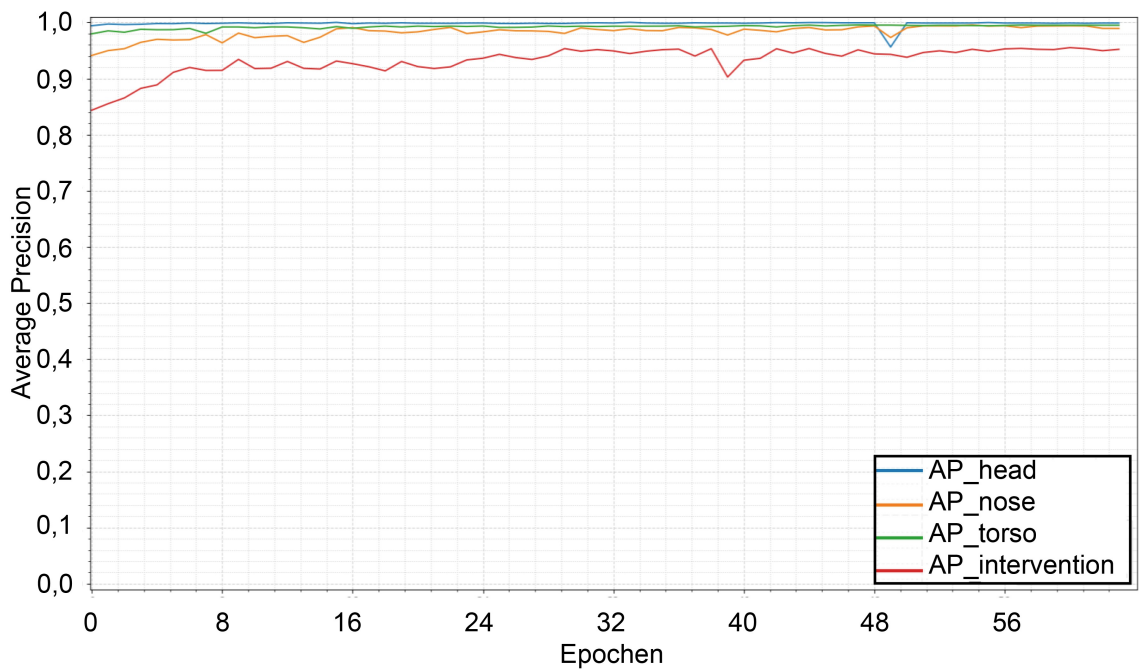


Abb. 6.26: Trainingsverlauf Phase 2 Fusions-Daten für YOLOv3, AP per Epoche

6.4.2 Neuronale Netze Trainingsverläufe

6.4.2.1 RetinaNet

RGB Bei den RGB-Daten wird ein steiler Anstieg in der ersten Trainingsphase gefunden. Für "torso", "nose" und "head" sind bereits ausreichend Daten vorhanden, um APs von über 0.99 zu erreichen. Von "intervention" liegen etwas weniger Daten vor, daher wird nur ein AP von 0.94 erreicht. Dieser Wert wird auch durch das weitere Training nicht besser. Offenbar sind im Trainingsdatensatz zu wenige geeignete Instanzen der Klasse "intervention" enthalten.

Thermo Bei den Thermo-Daten weist das RetinaNet Oszillationen für die Klassen "torso" und "intervention" auf, sowie einen Einbruch in einer Epoche. Oszillierendes Verhalten entsteht durch die Anpassung der Lernrate während des Trainings-Prozesses. Je weiter das Optimum entfernt ist, desto größer ist die Schrittlänge. Der Einbruch kann durch eine falsche Anpassung der Gewichte am Anfang verursacht werden, zusätzlich sind in dem Batch wenige Instanzen der Klasse "intervention" enthalten und können somit weniger in die Gewichtung eingehen. Die Genauigkeiten sind etwas schlechter als für die RGB-Daten. "intervention" erreicht nur eine AP von 0.87 nach der ersten Trainingsphase und von 0.91 mit allen Bildern.

Fusion "head" und "torso" erreichen beim Training mit den Daten von nur vier Probanden früh ein stabiles Plateau, mit APs über 0.99. Von "nose" und "intervention" waren deutlich zu wenige Instanzen im ersten Trainingsdatensatz enthalten. Durch zusätzliche Daten der weiteren acht Probanden kommt auch die Klasse "nose" auf eine AP von 0.95 und die Klasse "intervention" auf 0.81, wobei für "intervention" nur zwei Drittel der Anzahl der RGB-Bilder vorhanden war.

ToF Ein Teil der ToF-Daten wurde für zusätzliches Training des RGB-Netzes verwendet. Die Erkennungsgenauigkeit von ToF-Bildern ist entsprechend der niedrigen Anzahl der Trainingsinstanzen schlechter als bei den anderen Modalitäten.

6.4.2.2 YOLOv3

RGB Für die RGB-Daten ist ein steiler Anstieg der APs für "head" und "torso" ersichtlich, ähnlich schnell wie bei RetinaNet. Bei "nose" und "intervention" fallen mehr Oszillationen auf. Das YOLOv3 Netz benötigt deutlich mehr Epochen bis ein stabiles Plateau erreicht wird. Dort sind die Genauigkeiten aber ähnlich wie bei RetinaNet. Bei "intervention" wird sogar eine AP von 0.98 erreicht. Es folgt keine weitere Verbesserung durch zusätzliche Daten in der zweiten Trainingsphase.

Thermo Bereits in der ersten Trainingsphase wird nach einem steilen Anstieg bei "head", "torso", sowie "nose" schnell eine hohe Genauigkeit größer 0.99 erreicht. Bei "intervention" ist der Graph stark oszillierend und die AP liegt nur bei 0.92. Diese lässt sich durch weitere Daten bis auf 0.96 steigern.

Fusion Für "head" und "torso" steigt der Graph schnell steil an, bis zu einem AP größer 0.99. "nose" und "intervention" sind stark oszillierend. Die beste AP für "nose" liegt bei 0.33 nach

der ersten Trainingsphase. Durch weiteres Training kann auch dort eine AP von 0.99 erreicht werden.

ToF Die erreichten Genauigkeiten sind ähnlich wie beim RetinaNet, bei der Klasse "intervention" etwas besser.

6.4.3 Ergebnisse Testdatensätze

Mit dem Test-Datensatz wird ausgewertet, wie effektiv die Netzwerke für generalisierte Anforderungen trainiert wurden. Zur Auswertung der Testdatensätze werden die Gewichte der jeweilig besten Epoche verwendet. Hierbei ist wichtig, dass die Verteilung des Test-Datensatzes dem des Training-Datensatzes entspricht. Bei der vorhandenen Datenmenge war das nicht zu gewährleisten. Testdatensatz 1 enthält 65 Bilder von Proband 10. Für die Klasse "nose" waren unverhältnismäßig wenige Testbilder verfügbar. Die übrigen ToF-Daten von Proband 10 flossen in das Training der Phase 2 ein. Im Testdatensatz 2 kamen zufällig ausgewählte Bilder von Proband 17 zum Einsatz, mit jeweils 140 Bildern pro Modalität, die bisher noch nicht vom Netzwerk gesehen wurden. Für die Klasse "intervention" waren allerdings ausschließlich RGB-Bilder enthalten.

Tab. 6.9: Anzahl von Instanzen pro Klasse pro Modalität im Testdatensatz 1

Modality	head	nose	torso	intervention
RGB	62	12	58	25
Thermo	62	2	54	15
Fusion	64	2	51	22

Tab. 6.10: Anzahl von Instanzen pro Klasse pro Modalität im Testdatensatz 2

Modality	head	nose	torso	intervention
RGB	132	67	128	128
Thermo	140	10	139	0
Fusion	140	30	134	0
ToF	86	41	140	0

6.4.3.1 Ergebnisse mit Testdatensätzen 1 und 2 nach Trainingsphase 1

Aus der untenstehenden Tabelle 6.11 ist klar ersichtlich, dass die Trainingsdaten zu wenige Instanzen der Klasse "nose" und "intervention" enthielten. Mit den Testdatensätzen waren die Average Precisions erheblich niedriger als die dokumentierten Validierungsgenauigkeiten, was ein klares Zeichen für Overfitting ist. Auch für "torso" hätten aufgrund der hohen Komplexität der Bilder im Trainingsdatensatz eine höhere Anzahl an Instanzen enthalten sein müssen, um eine gute Generalisierung zu erreichen.

Tab. 6.11: Average Precision pro Klasse pro Modalität pro Netzwerk (Test-Datensatz 1 nach Trainingsphase 1)

	Modalität	Epoche	Average Precision			
			head	nose	torso	intervention
RetinaNet	RGB	25	0,9255	0,0000	0,7841	0,5884
	Thermo	28	0,9227	0,0000	0,7816	0,4856
	Fusion	38	0,9958	0,0000	0,6863	0,6574
YOLOv3	RGB	64	0,8563	0,0000	0,8301	0,5360
	Thermo	61	0,8924	0,0000	0,8611	0,5706
	Fusion	56	0,9455	0,0000	0,7864	0,5199

Tab. 6.12: Average Precision pro Klasse pro Modalität pro Netzwerk (Test-Datensatz 2 nach Trainingsphase 1)

	Modalität	Epoche	Average Precision			
			head	nose	torso	intervention
RetinaNet	RGB	25	0,7942	0,0000	0,0002	0,4776
	Thermo	28	1,0000	0,0000	0,0000	0,0000
	Fusion	38	0,8869	0,0000	0,0198	0,0000
YOLOv3	RGB	64	0,6955	0,0000	0,0262	0,1594
	Thermo	61	0,9940	0,0000	0,0001	0,0000
	Fusion	56	0,9678	0,0000	0,0377	0,0000

6.4.3.2 Ergebnisse mit Testdatensätzen 1 und 2 nach Trainingsphase 2

Tab. 6.13: Average Precision pro Klasse pro Modalität pro Netzwerk (Test-Datensatz 1 nach Trainingsphase 2)

	Modalität	Epoche	Average Precision			
			head	nose	torso	intervention
RetinaNet	RGB	43	0,9177	0,1071	0,8244	0,4333
	Thermo	34	0,9576	0,0000	0,8660	0,5490
	Fusion	29	0,9401	0,0000	0,4716	0,4063
YOLOv3	RGB	64	0,9333	0,1742	0,8587	0,5231
	Thermo	61	0,9220	1,0000	0,8986	0,6119
	Fusion	60	0,9172	0,0000	0,7576	0,5334

Tab. 6.14: Average Precision pro Klasse pro Modalität pro Netzwerk (Test-Datensatz 2 nach Trainingsphase 2)

	Modalität	Epoche	Average Precision			
			head	nose	torso	intervention
RetinaNet	RGB	43	0,9473	0,5836	0,4635	0,4538
	Thermo	34	0,9999	0,0000	0,1552	0,0000
	Fusion	29	0,9997	0,2088	0,6075	0,0000
	ToF	26	0,8839	0,0000	0,6622	0,0000
YOLOv3	RGB	64	0,9113	0,7658	0,4955	0,3638
	Thermo	61	1,0000	0,0000	0,0005	0,0000
	Fusion	60	0,9857	0,2720	0,0322	0,0000
	ToF	59	0,9936	0,9855	0,7018	0,0000

Durch Trainingsphase 2 verbesserten sich die APs insbesondere bei Testdatensatz 2 für alle Klassen. Besonders für die wichtige Klasse "head" konnten hohe Genauigkeiten erzielt werden. Die unterschiedlichen Testergebnisse mit den beiden Testdatensätzen zeigen, dass keine ausreichende Generalisierung erreicht wird. Dafür wäre ein Training mit zusätzlichen unterschiedlichen Probanden nötig.

Abbildung 6.27 zeigt beispielhaft die visuelle Auswertung von Fusions-Bildern mit RetinaNet und YOLOv3. Die erreichten Genauigkeiten sind der Bildunterschrift zu entnehmen.

Für die Herzraten- und Temperaturdetektion ist die robuste Detektion der Klasse "head" am wichtigsten. Da die Genauigkeiten bei beiden Neuronalen Netzen ähnlich gut sind, aber YOLOv3 eine deutlich geringere Rechenzeit benötigt, wurde entschieden, im Gesamtsystem YOLOv3 einzusetzen.

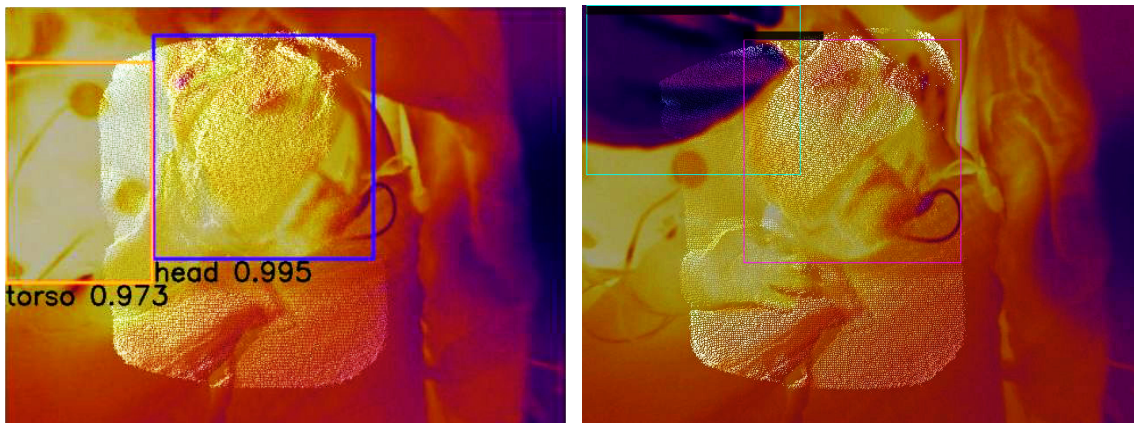


Abb. 6.27: Visuelle Auswertung von Beispiels Fusions-Bildern mit RetinaNet (links) und YOLOv3 (rechts). Mit RetinaNet werden confidence scores von 0.995 für den Kopf und 0.973 für den Torso erhalten. Für YOLOv3 ergeben sich confidence scores von 0.95 für den Kopf und 0.94 für die Intervention. [Gle+23]

6.5 Atemdetektion mit Time-of-Flight-Kamera und Neuronalem Netz

Die austrianierten Neuronalen Netze liefern den Torso-Bereich für die Atemdetektion mit der Time-of-Flight-Kamera.

In Abbildung 6.28 sind für Proband 12, 15, 16 und 17 Box-Plots der prozentualen Abweichungen der Atemdetektion mit der ToF-Kamera von der manuell gezählten Referenz aufgeteilt nach aufgezeichnetem Datensatz zu sehen. Hierbei muss auch die Aktivität der Probanden während der Aufnahmen beachtet werden. Die Aktivitäten werden wie in folgender Tabelle klassifiziert:

Tab. 6.15: Klassifizierung Aktivität

Aktivität Kürzel	Beschreibung
0	Schlaf
1	wach / leichte Bewegung
2	mittlere Bewegung / weinen
3	starke Bewegung / schreien
i	Intervention
l	leerer Inkubator
x	Video steht

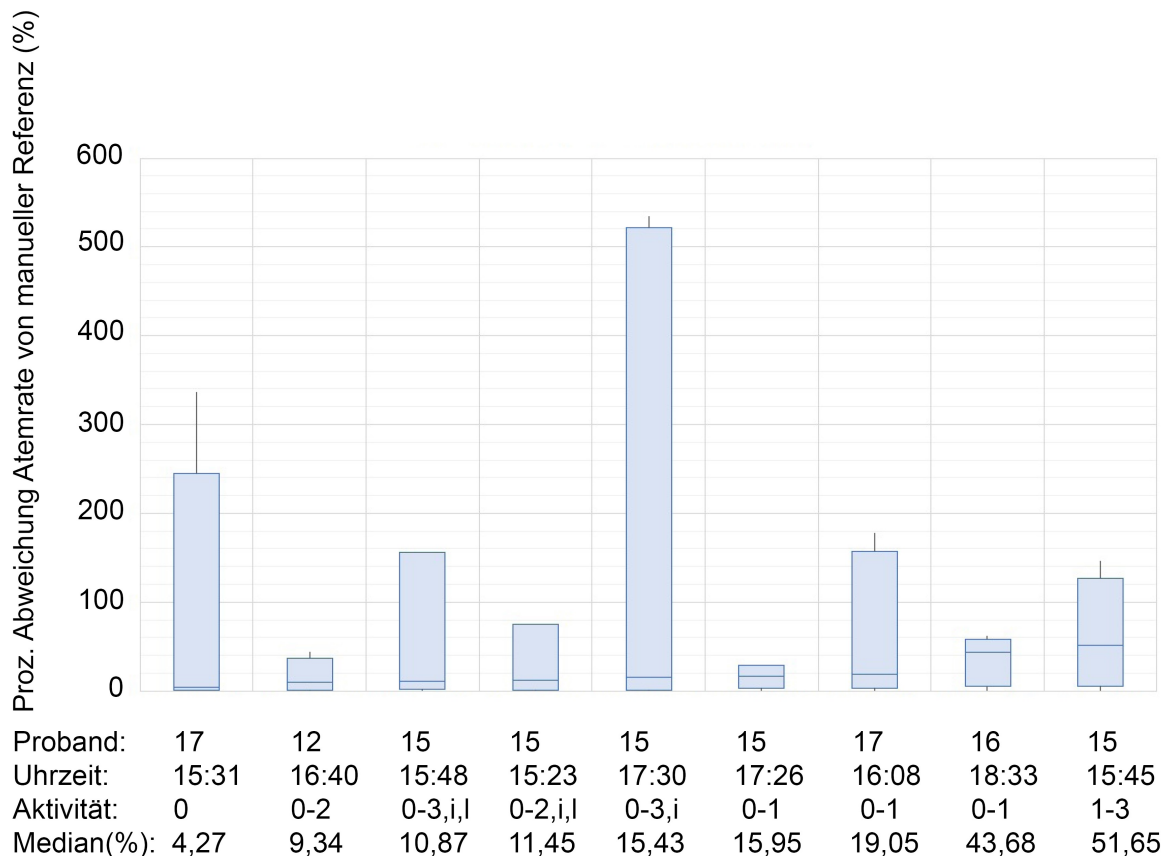


Abb. 6.28: Box-Plot der prozentualen Abweichungen der Atemdetektion mit der ToF-Kamera sortiert nach aufsteigendem Median.

Bei Proband 17 15:31 ist der kleinste Median der prozentualen Abweichung zu finden. Das Kind schläft während der Aufnahme. Die Aufnahme ist unter Phototherapie-Blaulicht entstanden. Die Bilder von Proband 17 sind dem Neuronalen Netz unbekannt, daher eventuell schlechte Detektion des Torsos.

Proband 12 16:40 schläft die meiste Zeit des aufgezeichneten Datensatzes mit Bauch frontal zur Kamera. Schläuche liegen über dem Bauch. Das Kind bewegt sich kurz, dabei verdreht sich der Torso und die erhobene Hand ist im Bild.

Proband 15 15:48 bewegt sich stark. Es findet Intervention durch das Pflegepersonal statt. Zeitweise wird das Kind aus dem Inkubator genommen.

Bei Proband 15 15:23 wurde in der Hälfte der Zeit der leere Inkubator aufgenommen. Anschließend liegt der Thorax des Kindes frontal zur Kamera und das Kind atmet gleichmäßig. Es findet eine ausführliche Intervention statt: Das Kind wird gewickelt und neu verkabelt.

Bei Proband 15 17:30 liegen die Atembewegungen meistens in einer Ebene. Kurze starke Bewegung und Verdrehung tritt auf, Schreien und anschließende Intervention (Hand auf Thorax). Proband 15 17:26 liegt relativ ruhig, der Torso befindet sich frontal vor der Kamera. Die Atembewegung ist gleichmäßig.

Proband 17 16:08 schläft während der Aufnahme und bewegt sich leicht. Die Bilder von Proband 17 sind dem Neuronalen Netz unbekannt, daher eventuell schlechte Detektion des Torsos.

Proband 16 18:33 liegt seitlich leicht verdreht, dadurch ist die Brustkorbbewegung keine reine Bewegung in die Höhe, sondern eine auch laterale Wellenbewegung.

Proband 15 15:45 schläft zunächst, macht dann viele starke Bewegungen, wirft den Kopf hin und her.

6.6 Atemdetektion mit Thermokamera und Neuronalem Netz

Für die Atemdetektion mit der Thermokamera liefern die Neuronalen Netze die Bounding Box des Nasenbereichs.

Abbildung 6.29 und Abbildung 6.30 zeigen die Ergebnisse der Atemdetektion mit der Thermokamera. Ausgewertet wurden nur die Zeiten, in denen keine Bewegung oder Intervention vorliegt, da bei Bewegung die Referenzmessung ungenau ist. Es sind deutliche Abweichungen zur Referenz zu erkennen. Für Proband 9 liegt der Median der prozentualen Abweichung bei 32,55 % und bei 36,31 % für Proband 15.

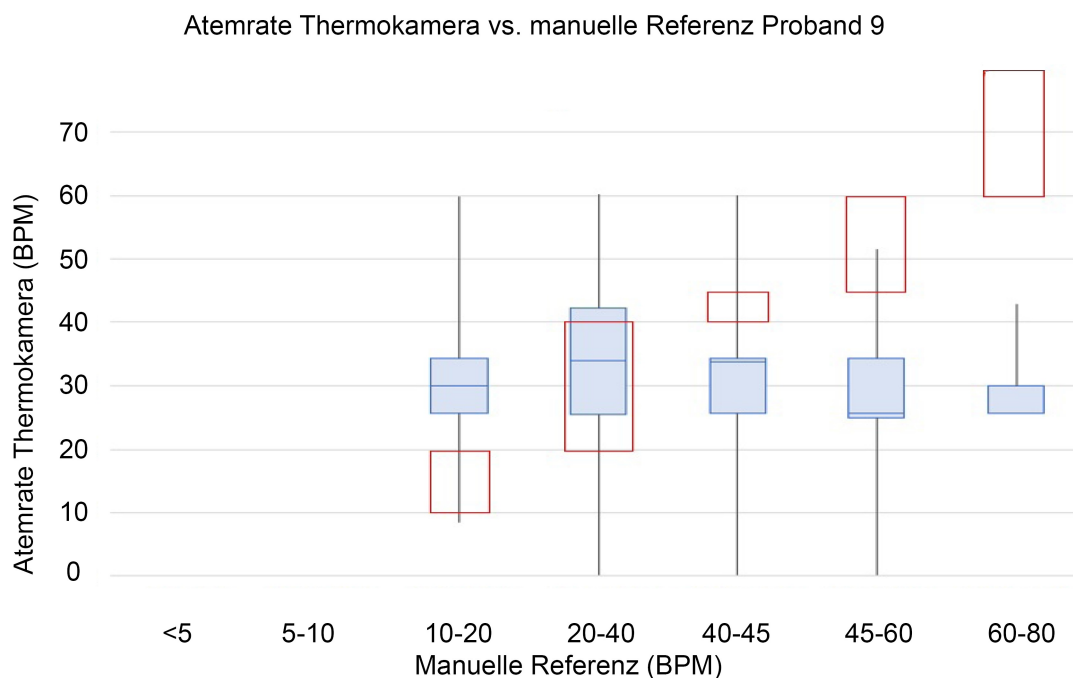


Abb. 6.29: Box-Plot der Atemdetektion mit der Thermokamera von Proband 9 aufgeteilt nach Frequenzbereichen. (Rot: Frequenzbereich der manuellen Referenz, blau: Ergebnisse der Thermokamera)

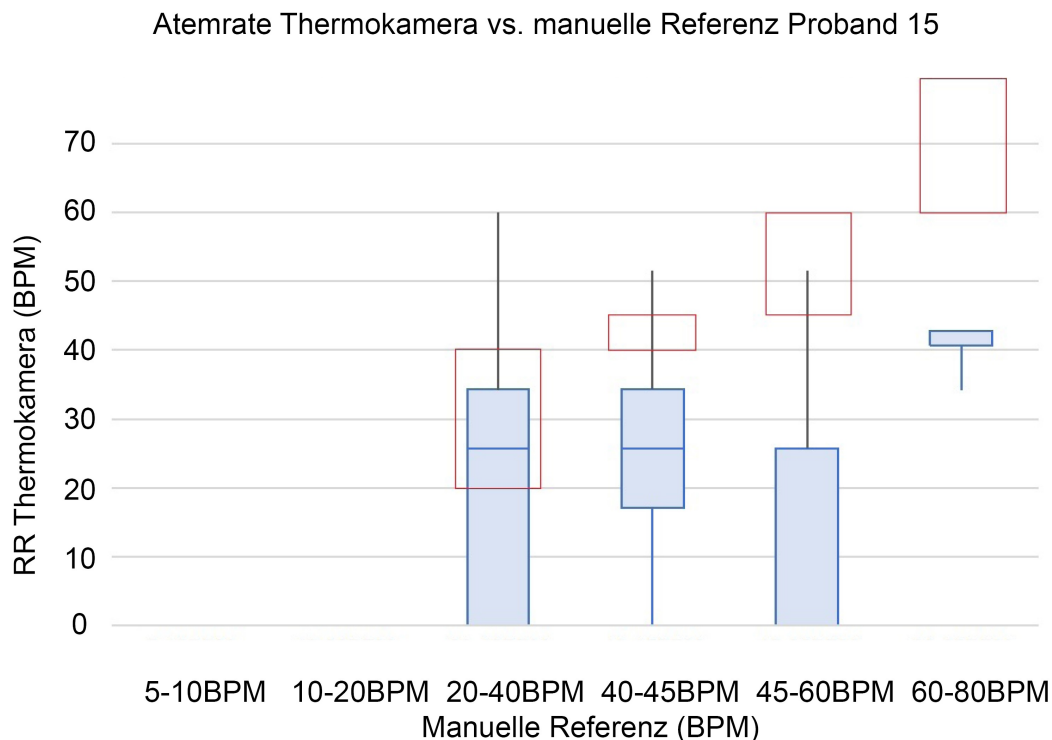


Abb. 6.30: Box-Plot der Atemdetektion mit der Thermokamera von Proband 15 aufgeteilt nach Frequenzbereichen. (Rot: Frequenzbereich der manuellen Referenz, blau: Ergebnisse der Thermokamera)

6.7 Herzratendetektion mit RGB-Kamera und Neuronalem Netz

Für die Herzratendetektion werden die sichtbaren Gesichtsareale verwendet. Dafür liefern die Neuronalen Netze die Bounding Box des Kopfes. Die besten Ergebnisse werden erzielt, wenn nur ein Sechstel der Bounding Box (ab dem Kinn) genutzt wird.

Es wurde mit aufgezeichneten Datensätzen verschiedener Lichtverhältnisse (bei Tageslicht, Dunkelheit, bzw. Blaulicht während der Phototherapie) getestet und der Einfluss von Bewegung und Interventionen erfasst. Dafür wurde jeweils der RMSE und der Median der prozentualen Abweichung berechnet. Mit Blaulicht (Proband 17, 15:31) wird ein RMSE von 14,07 BPM erreicht und ein Median der prozentualen Abweichung von 9,49 %. Der Verlauf der Herzrate im Vergleich zur Referenz des EKGs ist in Abbildung 6.32 zu sehen. Bei Tageslicht (Proband 17, 16:08) liegt der RMSE bei 26,35 BPM und einem Median der prozentualen Abweichung von 16,49 %. Im Vergleich dieser beiden aufgezeichneten Datensätze ist erkennbar, dass die Phototherapie keinen schädlichen Einfluss auf die Qualität der Messung hat. Ähnliche Ergebnisse werden erreicht, wenn der Arm des Kindes vor dem Gesicht liegt (Proband 16, 18:33) mit einem RMSE von 29,60 BPM und einem Median der prozentualen Abweichung von 18,46 %. Hier wurde offensichtlich die Herzrate nicht im Gesicht, sondern am Arm gemessen. Deutlich schlechtere Ergebnisse sind bei Bewegung und Intervention (Proband 15, 17:30) zu finden mit einem RMSE von 53,07 BPM und einem Median der prozentualen Abweichung von 23,36 %.

Bei Dunkelheit (Proband 16, 17:29) steigt der Fehler der gemessenen Werte abermals an, auf einen RMSE von 61,32 BPM und einem Median der prozentualen Abweichung von 37,34%. Zum Vergleich: Wenn die gesamte Bounding Box der Klasse "head" herangezogen wird, sind die Ergebnisse nochmal deutlich schlechter mit RMSE von 55,67 BPM und einem Median der prozentualen Abweichung von 45,57% (Proband 17, 15:31).

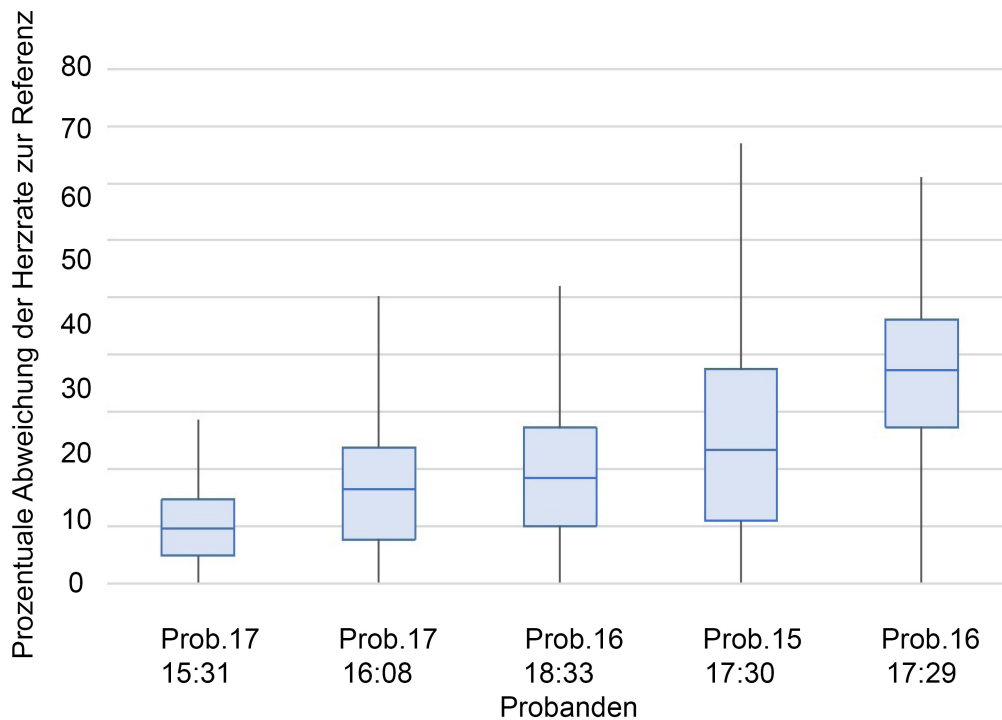


Abb. 6.31: Box-Plot der prozentualen Abweichungen der Herzratendetektion mit der RGB-Kamera. Es wurde mit aufgezeichneten Datensätzen verschiedener Situationen getestet.

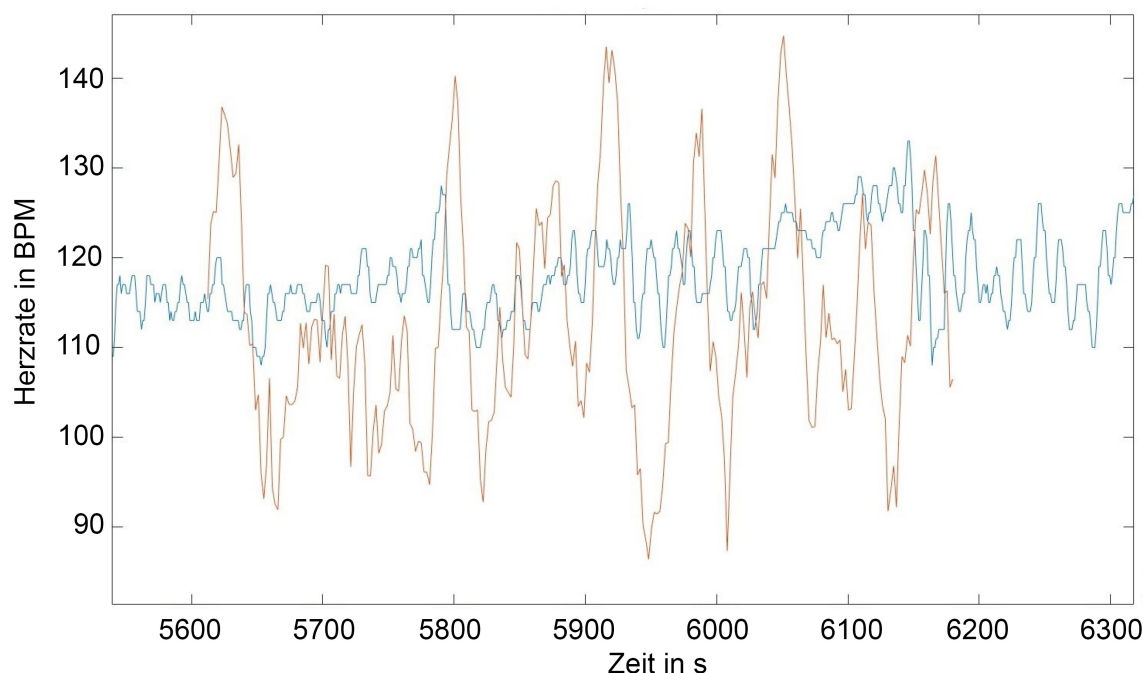


Abb. 6.32: Verlauf der Herzrate mit RGB-Kamera (rot) vs. Referenz vom EKG (blau) von Proband 17, 15:31.

6.8 Herzratendetektion mit Radar-Sensor

Die drei vorgestellten Algorithmen Peak2Peak, FFT per Band und One FFT for all wurden als Software Simulation und am Baby-Thorax-Simulator getestet.

6.8.1 Software Simulation

Es wurde der reine Herzschlag bei folgenden Raten simuliert: 60, 80, 100, 120, 140, 160, 180, 200, 220 und 240 BPM. Zusätzlich wurde zwischen zwei Wellenformen unterschieden: Sinuswelle und Pulswelle. Die Überlagerung mit der Atmung wurde ebenfalls getestet. Hierbei wurden alle Kombinationen der Herz- und Atemrate simuliert, bei welcher die Atemrate niedriger als die Herzrate ist. Die Kombinationen (80, 80), (80, 60) und (60, 60) wurden vernachlässigt. Wenn tiefe und normale Atmung, Sinus- oder Rechtecksignal der Atmung, sowie die Sinus- oder Pulswelle miteinbezogen werden, ergaben sich 536 Kombinationen. Die Dauer der einzelnen Testfälle betrug jeweils 300 s.

Ergebnisse der Software Simulation Über alle Testfälle hinweg wurde der Root Mean Square Error (RMSE) für die drei Algorithmen berechnet. Wie in Tabelle 6.16 zu sehen, werden die niedrigsten Werte für den FFT per Band Algorithmus erhalten. Somit ist dieser für die Herzratendetektion am besten geeignet.

Tab. 6.16: RMSE Software Test in BPM

Algorithmus	min	mean	median	max
FFT per band	2,50	28,98	4,61	163,20
Peak2Peak	5,12	57,47	19,80	626,88
one FFT for all	0,00	109,94	100,18	240,00

6.8.2 Hardware Simulation

Nach den Tests mit der Software Simulation wurden auch Tests am Hardware-Simulator getätigt.

6.8.2.1 Messaufbau

Die Messungen wurden am selbst entwickelten Baby-Thorax-Simulator durchgeführt. Der Messaufbau wurde bereits in Abschnitt 5.1 vorgestellt.

6.8.2.2 Goldstandard

Als Referenzmessung wurde ein Laser-Mikrometer eingesetzt. Die Eigenschaften und Genauigkeit sind in Abschnitt 5.1 beschrieben. Um sicherzustellen, dass das Laser-Mikrometer korrekte BPM-Werte liefert, wurde eine manuelle Zählung der Atemhübe zur Verifizierung durchgeführt.

6.8.2.3 Datengewinnung

Genau wie mit der reinen Software Simulation wurden mit dem Baby-Thorax-Simulator alle Kombinationen von Herz- und Atemrate, allerdings nur mit der Sinuswellenform getestet. Daraus ergaben sich 134 Testfälle. Wie zuvor lag die Dauer jedes einzelnen Testfalles bei 300 s.

6.8.2.4 Ergebnisse am Baby-Thorax-Simulator

Auch hier wurden über alle Testfälle der RMSE für die drei Algorithmen berechnet und unterschieden, ob ein Ellipsen-Fitting vorlag oder nicht. Die niedrigsten Werte wurden wie zuvor für den FFT per Band Algorithmus erhalten. Im Vergleich zum Softwaretest sind die RMSE Werte beim Test am Baby-Thorax-Simulator etwas höher.

Tab. 6.17: RMSE ohne und mit Ellipsen Fitting in BPM

RMSE ohne Ellipsen Fitting in BPM.				
Algorithmus	min	mean	median	max
FFT per band	2,21	41,18	22,71	183,30
Peak2Peak	2,90	136,66	133,90	472,49
one FFT for all	0,00	69,71	57,89	240,00
RMSE mit Ellipsen Fitting in BPM				
Algorithmus	min	mean	median	max
FFT per band	2,26	42,71	24,70	183,35
Peak2Peak	3,34	135,69	125,87	417,61
one FFT for all	0,00	66,89	59,08	239,94

6.9 Temperaturdetektion mit Thermokamera und Neuronalem Netz

Das Neuronale Netz liefert die Bounding Box "head". Die in der Studie teilnehmenden Probanden tragen als Schutz während der Phototherapie eine mehr oder weniger große Binde über den Augen. In der Folge ist der detektierte Gesichtsbereich unterschiedlich groß.

Der Temperatursensor im Inkubator zur Kalibrierung der Thermokamera stand erst ab Proband 13 zur Verfügung.

Die Ergebnisse sind beispielhaft in den untenstehenden Box-Plots (Abbildung 6.33) zu sehen. Für die dargestellten Box-Plots wurden jeweils 5000 Werte verwendet. In der ersten Sekunde der Messung treten falsch-niedrige Werte auf, die von der Auswertung ausgeschlossen wurden. Die geringste Abweichung liegt bei Proband 13 (Bag 1) mit einem Median von 0,0197 K und RMSE von 0.1208 K. Das Kind trug nur eine schmale Augenmaske und es fand während der Messung keine Intervention statt. In der Standardsituation des Monitorings eines Kindes mit unverdecktem Kopf im geschlossenen Inkubator werden immer ähnlich gute Ergebnisse erreicht. Proband 17 (Bag 2) mit einem Median von -0,0275 K und RMSE 0.1568 K trug zwar eine breite Augenbinde, hat aber einen ausgeprägten Hinterkopf, sodass in der Detektion etwa jeweils zur Hälfte die Binde bzw. der Kopf sichtbar ist. Bei Proband 16 (Bag 1) ist der Fehlermedian etwas größer mit -0,2090 K und einem RMSE von 0.2117 K. Das Kind schläft mit seitlich gelegtem Kopf, dabei wird wenig Hautareal von der Thermokamera erfasst. Bei Proband 16 (Bag 2) findet Intervention statt, der Kopf kann verdeckt werden. Der Fehlermedian beträgt 0,7215 K und der RMSE 0.8689 K. Proband 15 (Bag 4) trägt eine breite Augenmaske, die das Gesicht zu zwei Dritteln verdeckt, wodurch der Fehlermedian von 3,0651 K bzw. einem RMSE von 3.0498 K entsteht. In einer Phase des aufgezeichneten Datensatzes, in der sich das Kind stark bewegt und dabei sein Gesicht Richtung Kamera dreht, wodurch mehr Hautareal sichtbar wird, sinkt die Abweichung von der Referenz.

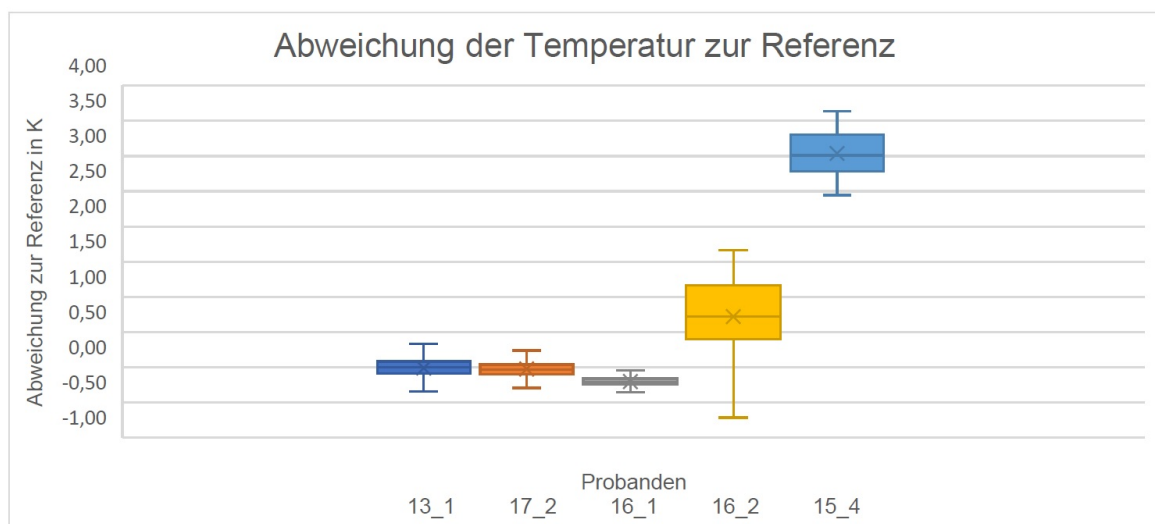


Abb. 6.33: Beispielhafte Box-Plots der Temperaturdetektion unter verschiedenen Bedingungen.

7 Diskussion

An dieser Stelle sollen die Ergebnisse diskutiert und mit dem theoretischen Ansatz und den beschriebenen Anforderungen sowie dem Stand der Technik verglichen werden.

7.1 Atemratendetektion mit Structured-Light-Kamera

Grundsätzlich sind die Ergebnisse der Tests mit der Structured-Light-Kamera vielversprechend. Da die Kamera aber größere Distanzen benötigt, als im Inkubator gegeben sind, wurde ein Einsatz verworfen. Das Verfahren und die Algorithmik sind Basis der Atemratendetektion mit der Time-of-Flight-Kamera.

7.2 Atemratendetektion mit ToF-Kamera und Radar-Sensor

7.2.1 Einordnung der Ergebnisse

Obwohl mehrere Nachteile der Sensoren aus der Literatur bekannt waren, hatten diese Effekte einen größeren Einfluss in der Praxis als erwartet. Dies hat zu einigen unerwarteten Beschränkungen der eingesetzten Algorithmen geführt.

Zum einen wurde nicht erwartet, dass die ToF-Kamera sich deutlich empfindlicher hinsichtlich Rauschen und Reflexionen verhält, als die zuvor eingesetzte Structured-Light-Kamera [GNK20]. Obwohl Vorkehrungen wie das Abdecken von dunklen Oberflächen des Simulators vorgenommen wurden, um Reflexionen durch hohen Kontrast zu vermeiden, sowie der Einsatz von zusätzlichen Filtern, sind die Distanzsignale bei niedrigen Frequenzen stark verrauscht. Im niedrigen Frequenzbereich ist der Hub des Baby-Thorax-Simulators klein (proportional zur eingestellten Spannung des Lautsprechers). Das bedeutet, dass der SNR klein ist. Die Peak-Detektion liefert für diese Signale keine befriedigenden Ergebnisse, da eine zu hohe Anzahl an Peaks erkannt wird, was zu einer inkorrekten Parametrierung des Savitzky-Golay-Filters führt. Da der Savitzky-Golay-Filter feste Grenzen hat, die auch noch bei hohen Frequenzen gelten, können diese nicht weiter verschoben werden, ohne eine zu starke Filterung im hohen Frequenzbereich zu verursachen.

Es wurde erwartet, dass die Ergebnisse bei kleinen Atemraten im *deep* Modus besser sind als im *normal* Modus, da dann größere Hübe vorliegen. Diese Erwartung ließ sich in den Tests nicht bestätigen.

Es wird mindestens alle 5 s ein neuer Atemraten-Wert benötigt (denn so oft liefert ein Monitoring-Gerät auf der neonatologischen Intensivstation einen neuen Wert). Ideal wäre ein neuer Wert alle 3 s. Der Radar-Algorithmus liefert alle 50 ms einen neuen BPM-Wert. Der Ansatz mit der ToF-Kamera berechnet ca. nur alle 7,5 s einen neuen Wert. Dieser Wert ist niedrig im Vergleich

zum Structured-Light-Kamera Ansatz, bei dem alle 3 s ein neuer Wert generiert wird [GNK20]. Dies kann durch die Erweiterung mit dem Median und Statistical Outlier Filter begründet werden, die mehr Verarbeitungszeit benötigen. Außerdem wird nun eine Subebene berechnet, was im anderen Ansatz nicht der Fall war. Da aber über die synchrone Auswertung von Radar-Sensor- und ToF-Kamera-Daten in Kombination geredet wird, werden auf jeden Fall durch den Radar-Sensor häufig genug neue Werte generiert.

Der Einsatz des Ellipsen-Fittings beim Radar-Sensor führte zu einer unerwarteten Änderung der Amplitude des Distanzsignals und erfordert zusätzliche Rechenzeit und sollte daher im Gesamtsystem nicht angewendet werden, um zu gewährleisten, dass alle 5 s ein neuer Wert kommt.

Es wurde gezeigt, dass beide Sensoren ihre Stärken und Schwächen haben, welche sich gut ergänzen, wenn die BPM bestimmt werden. Die Ergebnisse sind eine gute Basis, auf die aufgebaut werden kann, um eine größere Robustheit durch die Kombination beider Sensoren zu erhalten. Es wurde festgestellt, dass bei niedrigen Atemraten von 5 bzw. 10 BPM die mit dem ToF-Algorithmus berechneten BPM-Werte zu hoch sind. In diesem Fall muss der niedrigere Wert des Radar-Sensors berücksichtigt werden, da diese Werte einen kritischen physiologischen Zustand und somit ein Risiko für das Neugeborene darstellen können. Für alle Atemraten im Normbereich (20 bis 60 BPM) detektiert die ToF-Kamera grundsätzlich etwas genauer. Daher wird ein Entscheidungsalgorithmus benötigt, der die Daten des jeweils genaueren Sensors anzeigt.

Eine "early" Sensorfusion würde bei diesem Ansatz nicht zur Verbesserung der Genauigkeit oder Robustheit beitragen, da bei der Synchronisierung kein eigentliches Punktemapping sinnvoll ist, sondern die Ergebnisse von getrennt berechneten Peak Detections verglichen werden.

7.2.2 Vergleich mit dem Stand der Technik

Mit dem Journal Beitrag [Gle+21] wurde gezeigt, dass diese Arbeit die Erste ist, bei der Daten eines mikrowellen-interferometrischen Radar-Sensors und einer ToF-Kamera synchron zur Atemdetektion von Neonaten ausgewertet wird. Wie in Kapitel 3 erwähnt, wertet WikiSpiro die Daten beider Sensoren nicht synchron aus [Ngu+16]. Die ToF-Kamera wird nur zum Tracking der Position des Patienten verwendet, um eine genaue Positionierung des Radar-Sensors zu ermöglichen. Außerdem wurde ihr System nur an Erwachsenen getestet.

Das hier beschriebene System liefert genaue BPM-Werte im Normbereich zwischen 20 und 60 BPM mit einer maximalen Abweichung von 3 BPM für die ToF-Kamera bei einer Atemrate von 30 BPM im *normal* Modus. Im Vergleich dazu erreichen Beltrao et al. mit ihrem 24 GHz CW-Radar-Sensor nur Abweichungen von 5-7 BPM an Frühgeborenen [Bel+22]. Der hier präsentierte kombinierte Ansatz ist beim Test am Simulator genauer. Da die Ethikkommission keinen Test mit einem Radar-Sensor an Studienteilnehmern erlaubte, ist ein direkter Vergleich nicht möglich, es werden aber gute Ergebnisse auch an Neonaten erwartet.

7.3 Detektion von Körperarealen mit Neuronalen Netzen

Die anschließend folgenden Verfahren nutzen eine automatische Eingrenzung der Region of Interest mittels Neuronalen Netzen.

7.3.1 Diskussion der Trainings-Ergebnisse

Bei der Diskussion der Trainings-Ergebnisse wird Bezug auf die Average Precisions in Tabelle 6.8 genommen. Die bereits mit weniger Probanden als geplant erreichte hohe Genauigkeit beider Netze zeigt, dass sie wie erwartet ausgezeichnet geeignet sind, um die Region of Interest für die weitere Auswertung automatisch zu bestimmen.

7.3.1.1 Training mit Fusions-Daten vs. Training mit RGB- oder Thermo-Daten

Tab. 7.1: Differenz zwischen Fusion und RGB / Thermo in Average Precisions pro Klasse für RetinaNet (Validierungs-Datensatz)

	Modalität	Epoche	Average Precision Differenz			
			head	nose	torso	intervention
RetinaNet	Fusion - RGB	-14	-0,0009	-0,0367	-0,0195	-0,0861
	Fusion - Thermo	-5	-0,0015	-0,0027	-0,0152	-0,0957
YOLOv3	Fusion - RGB	-4	0,0	0,0017	-0,0011	-0,0126
	Fusion - Thermo	-1	-0,0017	-0,0005	-0,001	-0,0046

In der Vorbereitungsphase für das Training der verschiedenen Netze zeigte sich ein entscheidender Vorteil des Trainings mit Fusionsdaten, nämlich, dass nur ein Netzwerk trainiert werden muss und damit wesentlich weniger Bilder gelabelt werden müssen als beim separaten Training eines Netzes mit RGB- bzw. Thermo-Daten. Dieses vereinfacht die Generierung von Datensätzen hinsichtlich Zeit und Kosten, es werden 66 % gespart.

Da die Fusions-Bilder mehr Informationen beinhalten, als die RGB- und Thermo-Bilder, dauert das Training maximal 14 Epochen länger, um dieselbe Genauigkeit zu erreichen, aber es sind 25 % weniger Trainingsbilder notwendig, um mit den Testdatensätzen eine höhere Genauigkeit als mit RGB-Bildern zu erzielen.

Die wichtigste Klasse ist "head". Hier gibt es bis auf das RGB-Netz nur geringe Unterschiede der Testdaten-Genauigkeit zu den Genauigkeiten des Validierungs-Datensatzes, was auf gut trainierte Netze hinweist. Es fiel auf, dass der Fusions-Datensatz schon in der ersten Trainingsphase mit den Daten von nur vier Probanden gute Ergebnisse mit dem Testdatensatz 2 bringt, während bei den RGB-Datensätzen die Genauigkeit gegenüber dem Validierungsdatensatz um 20,6 % beim RetinaNet bzw. 30,4 % bei YOLOv3 absank. Dies zeigt, dass das Training des Neuronalen Netzes mit fusionierten Daten die besten Ergebnisse hervorbringt. Nach der zweiten Trainingsphase liefern die RGB-Testdaten aus Testdatensatz 2 überraschenderweise nur Genauigkeiten von

0,9473 (RetinaNet) und 0,9113 (YOLOv3) für die Klasse "head", obwohl hierfür die meisten Trainingsbilder eingespeist wurden. Hier zeigt sich wohl eine Art Overfitting, es wurde mit zu vielen ähnlichen Bildern trainiert.

Es gibt von allen Modalitäten nicht ausreichend viele Instanzen der Klassen "nose" und "intervention" in den Trainings-Datensätzen, um eine genaue Detektion durchzuführen. Auch für "torso" werden bei ähnlicher Anzahl der Trainingsbilder wie bei RGB und Thermo mit den Testdaten schlechtere Average Precisions als bei "head" gemessen. Die Ursache liegt vermutlich in der komplexeren Darstellung des Torsos, der auf Bildern verdreht erscheinen kann, von den Armen des Kindes bedeckt ist und eine stärker variierende Form als der Kopf hat. Es wird Overfitting angenommen, was bedeutet, dass der Trainingsdatensatz zu klein war und das Netz die Features der vorhandenen Bilder zu genau erlernt hat, aber mit kleinsten Abweichungen davon nicht zurecht kommt, also nicht generalisieren kann. Im Unterschied dazu hat der Kopf eine einfachere Form, ist weniger häufig verdeckt und in den Thermokamera-Bildern immer am hellsten und deutlich zu erkennen.

Die Verläufe von "intervention" und "nose" zeigen ein schwingendes Verhalten. Dies kann durch die Verwendung des Stochastic Gradient Descent (SGD) Optimierer und seiner Lernrate erklärt werden. Je weiter weg das Optimum liegt, desto größer wird die Schrittlänge. Hier kann ein einzelner Batch eine Rolle spielen. Sind nicht genügend Instanzen einer Klasse in einem Batch vorhanden, wird das Netz jene Klassen bevorzugen, in der mehr Instanzen vorhanden sind, da dadurch die Kostenfunktion mehr gesenkt werden kann.

Zusammenfassend kann gesagt werden, dass das Fusions-Netzwerk eine ähnlich gute Genauigkeit wie das RGB- bzw. Thermo-Netzwerk liefert. Dies ist der Fall, wenn die Anzahl der Bilder und die Anzahl der Instanzen pro Klasse für jede Modalität und Klasse in den Trainingsdaten gleich groß sind. Wenn die extrinsische Transformation zwischen Thermo- und RGB-Kamera gegeben ist, kann die Position der detektierten Region automatisch für das Thermo- und RGB-Bild aus dem Fusionsbild berechnet werden, unter der Bedingung, dass die Region of Interest (ROI) in der ToF-Punktwolke liegt.

7.3.1.2 Vergleich RetinaNet und YOLOv3

Tab. 7.2: Unterschied der Average Precisions pro Klasse pro Modalität RetinaNet - YOLOv3 (Validierungs-Datensatz)

	Modalität	Epoche	Average Precision			
			head	nose	torso	intervent.
RetinaNet - YOLOv3	RGB	-21	-0,0006	-0,0014	-0,0001	-0,0636
	Thermo	-27	-0,0017	-0,0376	-0,0048	-0,046
	Fusion	-31	-0,0015	-0,0398	-0,0185	-0,1371

Wenn die Unterschiede in den Ergebnissen von RetinaNet und YOLOv3 für den Validierungsdatensatz betrachtet werden, ergibt sich, dass YOLOv3 für alle Modalitäten und Klassen minimal genauer als RetinaNet ist (max Differenz von -0,1371), auch wenn RetinaNet beim Training

weniger Epochen benötigt, um die höchste Genauigkeit zu erreichen. Beide Netze zeigen für "nose", "torso" und "intervention" Overfitting, es waren also zu wenige Instanzen der Klassen im Trainingsdatensatz enthalten. Dies ist der Tatsache zu entnehmen, dass der Testdatensatz schlechter abschneidet als der Validierungsdatensatz. Auch wenn bei RetinaNet die Ergebnisse für die Validierungsdaten grundsätzlich etwas schlechter als bei YOLOv3 sind, sticht bei den Testdatensätzen keines der beiden Netze hervor. Nur bei der Klasse "nose" in den Time-of-Flight-Testdaten detektiert YOLOv3 erheblich genauer.

Wenn bei beiden Netzen die Ergebnisse mit dem Test-Datensatz 2 (Tabelle 6.14) betrachtet werden, fallen niedrigere Average Precisions bei RGB "intervention" auf. Der Grund dafür liegt in der Komplexität mancher Testbilder, die in dieser Art nicht im Trainings-Datensatz vorlagen, wie zum Beispiel das Füttern mit einer Flasche. Für die Klasse "intervention" waren keine Thermo-, ToF- oder Fusionsbilder im zufällig generierten Testdatensatz 2 enthalten, womit die Average Precisions von 0,0000 zu erklären sind. Ähnliches gilt für die Klasse "nose" im Testdatensatz 1 für den Test nach der ersten Trainingsphase. Unverhältnismäßig wenige Instanzen führen hier ebenfalls zu einem schlechteren Ergebnis.

Wie auch schon von Hennemann gezeigt, werden über alle Klassen hinweg bessere Ergebnisse erhalten, wenn mit einer größeren Anzahl unterschiedlicher Probanden trainiert wird [Hen23].

Wie von Redmond et al. [JA18] und Hennemann [Hen23] erwähnt, ist die Rechenzeit pro Bild bei RetinaNet ungefähr dreimal länger als bei YOLOv3. Dies konnte auch in dieser Arbeit beobachtet werden. Es werden ca. 58 ms mit RetinaNet und 14 ms mit YOLOv3 bei Verwendung der beschriebenen Hardware benötigt. Die höhere Geschwindigkeit von YOLOv3 bei ähnlicher Genauigkeit führte zur Entscheidung, YOLOv3 zu favorisieren.

7.3.2 Vergleich mit dem Stand der Technik

Nachfolgend werden die Trainingsergebnisse für jede Klasse mit den besten in der Literatur beschriebenen Ergebnissen anderer Neuronaler Netzwerke verglichen. Green et al. erreichen eine Average Precision von 0,982 für die Kopf-Detektion [Gab+18]. Der Fusions-Ansatz liefert eine Validierungs-Genauigkeit von 0,9968 (RetinaNet) und 0,9983 (YOLOv3). Für den Test-Datensatz 2 nach Trainingsphase 2 wurden APs von 0,9997 und 0,9857 für RetinaNet resp. YOLOv3 erreicht. Dies zeigt, dass ähnlich hohe Genauigkeiten für Instanzen der Klasse "head" auch für Bilder, die das Netz noch nicht gesehen hat, erhalten werden können. Es ist zu erwarten, dass bei erweitertem Training der Neuronalen Netze Genauigkeiten auf dem Niveau der Validierungsergebnisse erreicht werden. Da die Detektionen für die RGB- und Thermo-Bilder direkt berechnet werden können, werden dafür mindestens die selben hohen Genauigkeiten erlangt. Für die Klassen "nose" und "intervention" gibt es keine Genauigkeitswerte in der Literatur. Beppu et al. nennen einen AP von 0,985 für den Torso [Bep+22]. In dieser Arbeit wurden Validierungs-Genauigkeiten von 0,9762 (RetinaNet) und 0,9947 (YOLOv3) mit dem Fusions-Netzwerk erzielt. Bei Betrachtung der komplexen Bilder im Test-Datensatz fallen die APs aufgrund der wenigen Probanden deutlich ab. Für "torso" werden die besten Ergebnisse nach Trainingsphase 2 mit Testdatensatz 1 in den Thermo-Bildern erzielt, mit einer AP von 0,8660 für RetinaNet und 0,8986 für YOLOv3. Es treten deutliche Unterschiede bei Testdatensatz 2

auf, was darauf hinweist, dass keine genügende Generalisierung erreicht wurde und die Netze mit Bildern von weiteren Probanden trainiert werden müssen, um eine verlässliche Detektion zu gewährleisten.

In dieser Arbeit werden erstmals Thermo-RGB-Fusions-Bilder von Neonaten zum Training von Neuronalen Netzen verwendet.

7.4 Atemdetektion mit ToF-Kamera und Neuronalem Netz

Wie oben beschrieben wird der Torso in unbekanntem Bildern durch die Neuronalen Netze beim bisherigen Trainingsstand schlecht detektiert. Für die Tests der Methode zur Atemdetektion mit ToF-Kamera und Neuronalem Netz spielt das jedoch keine Rolle, da an aufgezeichneten Datensätzen getestet wird, die den Neuronalen Netzen bekannt sind, also die Genauigkeit des Validierungsdatensatzes herangezogen werden muss.

7.4.1 Einordnung der Ergebnisse

Aus den Ergebnissen ist ersichtlich, dass die Aktivität des Kindes und Interventionen durch das Pflegepersonal einen starken Einfluss auf die Qualität der Messung haben. Nur beim ruhig schlafenden Kind ist der Median der prozentualen Abweichung mit 4,27 % so niedrig, dass er keine klinische Relevanz hat.

Das Neuronale Netz mit YOLOv3 benötigt pro Bild 14 ms Verarbeitungszeit [Gle+23], aber von der ToF-Kamera wird nur alle 7,5 s ein neuer Wert geliefert [Gle+21]. Dies entspricht ohne Kombination mit dem Radar-Sensor nicht den genannten Anforderungen, dass alle 5 s ein neuer Wert generiert werden muss.

Wie in [Gle+21] erwähnt, können durch die Synchronisierung mit einem Radar-Sensor die Defizite der ToF-Kamera hinsichtlich Genauigkeit und Echtzeitfähigkeit kompensiert werden. Für den Test der Synchronisierung an den Probanden gab es keine Genehmigung der Ethikkommission.

7.4.2 Vergleich mit dem Stand der Technik

Nach dem jetzigen Stand gibt es keine Publikation, die eine ToF-Kamera an Neonaten einsetzt. Aus diesem Grund soll das Verfahren allgemein mit kontaktlosen Verfahren zur Atemdetektion an Neonaten verglichen werden. Pereira et al. geben für ihr Verfahren mittels IR-Kamera eine prozentuale Abweichung von 3,42 % (vgl. Tabelle 3.1) an, erfasst an vier untersuchten Frühgeborenen, allerdings ohne Angabe des Aktivitätslevels. In der vorliegenden Arbeit wird bei schlafenden Probanden eine Abweichung von 4,27 % erreicht. Bei starken Bewegungen liegt die Abweichung allerdings bei bis zu 51,65 %. Dies zeigt, dass das Verfahren bei hoher Aktivität des Kindes nicht einsetzbar ist und nicht der klinischen Forderung nach einer Genauigkeit von ± 2 BPM entspricht. Ein ähnliches Verhalten wird jedoch beim Goldstandard, der Impedanz-Pneumographie auch beobachtet. Sahakin et al. zeigen, dass die weit verbreitete bipolar-Elektrodenkonfiguration (zwei Elektroden zur Impedanz-Pneumographie, eine zusätzliche für das EKG) hinsichtlich Artefaktsensitivität am schlechtesten abschneidet [STW85]. Davis et al. berichten von einem ähnlichen Problem mit Rauschen bei der Impedanz-Pneumographie

mit einem Philips IntelliVue MX800 Monitoring-Gerät. Sie weisen nach, dass bei 60 % der gemessenen Daten Abweichungen von ± 10 Atemzügen / Minute vorliegen [Dav+14], was die vom Hersteller genannte Messgenauigkeit von ± 1 BPM in Frage stellt. Der Signal Quality Index bei Charlton et al. identifiziert nur 34,9 % der realen Impedanz-Pneumographie-Daten als mit hoher Qualität [Cha+21]. Jeyhani et al. weisen darauf hin, dass das Problem von Bewegungsartefakten und Fehlern durch Kabelbewegung bisher nicht gelöst werden konnte [JVV16]. Somit ist die Nutzung von Impedanz-Pneumographie als Goldstandard schlecht geeignet.

7.5 Atemdetektion mit Thermokamera und Neuronalem Netz

7.5.1 Einordnung der Ergebnisse

Wie schon an anderen Stellen in dieser Arbeit erwähnt, tragen die an der Studie teilnehmenden Probanden während der Phototherapie eine Augen-Maske zum Schutz vor UV-Licht. Diese verdeckt häufig die Nase des Kindes, sodass diese nicht erkannt bzw. gelabelt werden kann. Dadurch liegt eine zu geringe Anzahl an Trainingsdaten für die Neuronalen Netze vor. Auch ein gut trainiertes Netz kann im Setting der Phototherapie häufig keine Nase des Kindes erkennen, z.B. wenn das Kind auf der Seite liegt oder die Binde über die Nase rutscht. Eine höhere Anzahl an Probanden würde die Detektionsgenauigkeit der Neuronalen Netze verbessern. Aber selbst wenn die Nase detektiert werden kann, ist der Atemstrom nicht unbedingt sichtbar oder liegt außerhalb der Bounding Box der Klasse "nose". Dies ist in den Ergebnissen von Proband 9 und 15 sichtbar. Um die Detektion zu verbessern, wäre es notwendig, eine Region of Interest "breath" zu definieren und in den Trainingsbildern für die Neuronalen Netze zu labeln.

Die Methode ist im derzeitigen Entwicklungsstand nicht einsetzbar.

Es wurde Impedanzpneumographie als Referenz genutzt zu den Zeiten, bei denen keine Bewegung und Intervention vorliegen. Nichtsdestotrotz ist die Impedanz-Pneumographie kein zuverlässiges Verfahren, wie in 7.4.2 bei der Atemdetektion mit der ToF-Kamera beschrieben.

7.5.2 Vergleich mit dem Stand der Technik

Pereira et al. erreichen mit ihrer kontaktlosen Atemdetektion mit der Thermokamera einen relativen Fehler von 3,42%. Bei ihrer Methode ist eine manuelle Auswahl der ROI und des kleineren Messbereichs (Region of Measurement) nötig. Als Referenz wird die Impedanz-Pneumographie genutzt. [Per+17] In dieser Arbeit wird die ROI automatisch detektiert. Allerdings kann der Atemstrom außerhalb liegen. Es wird ein Median der prozentualen Abweichung von mindestens 32,55% erreicht. Die klinischen Anforderungen an die Genauigkeit sind mit diesem Verfahren somit nicht erfüllt.

7.6 Herzratendetektion mit RGB-Kamera und Neuronalem Netz

7.6.1 Einordnung der Ergebnisse

Aus den gemessenen Daten wird ersichtlich, dass für eine möglichst genaue Messung der Herzrate mit RGB-Kamera eine gute Detektion der Hautareale durch das Neuronale Netz erforderlich ist. Die Ergebnisse für die gesamte Bounding Box "head" sind deutlich schlechter, als wenn ein Bildbereich mit mehr Hautareal-Anteilen verwendet wird. Erstaunlicherweise ist bei einem aufgezeichneten Datensatz, bei dem der nackte Arm des Kindes über dem Gesicht liegt, keine große Abweichung zum aufgezeichneten Datensatz, bei dem das Gesicht voll sichtbar ist, erkennbar. Es wird daraus geschlossen, dass eine Verbesserung der Ergebnisse erreicht werden könnte, wenn ein Neuronales Netz eingesetzt wird, das bestimmte Hautareale segmentiert.

Bei starker Bewegung und Intervention, z.B. wenn das Pflegepersonal Hände über das Gesicht des Kindes legt, können beliebige Abweichungen der erfassten Herzrate zur Referenz auftreten. Blaulicht hat keinen Einfluss auf die Genauigkeit, aber bei Dunkelheit fällt die Qualität der Messung stark ab.

Insgesamt zeigen die Ergebnisse starke Abweichungen zur Referenz, sind also ohne weitere Verbesserungsmaßnahmen nicht zuverlässig. Es ist nicht zu erwarten, dass eine Fusion mit anderen Sensoren zu einer Verbesserung der Messergebnisse führt, da der EVM allein zur Verstärkung der Farbänderungen eingesetzt wird. Eine Synchronisierung mit einem auf den Thorax gerichteten Radar-Sensor könnte die Detektion robuster machen.

7.6.2 Vergleich mit dem Stand der Technik

[Dav+14] weisen in ihrer Arbeit einen RMSE von 3,45 bis 4,27 BPM für die Herzratendetektion mit RGB-Kamera auf. Ihre Ergebnisse sind im Unterschied zu dieser Arbeit zustande gekommen, indem bei ähnlichem Verfahren (EVM, FFT) aufgrund der bekannten Probleme (Bewegung des Kindes, Lichtänderungen, zu geringe Hautdetektion) ausschließlich "gute" Bilder in Ruhephasen gewählt wurden, also 50 % des erhobenen Datensatzes nicht betrachtet wurde und kleine ROIs manuell ausgewählt wurden. Daher sind die Ergebnisse nur eingeschränkt vergleichbar mit dem hier präsentierten Verfahren. Es konnte nachgewiesen werden, dass die Genauigkeit steigt, wenn nur Bilder ohne Bewegung, Intervention und schlechtem Licht betrachtet werden und die ROI aus möglichst viel Hautareal besteht. In der praktischen Anwendung können Störfaktoren allerdings nicht ausgeblendet werden. Die automatische Detektion der ROI ist zwingend notwendig. Chen et al. erreichen mit ihrem ähnlichen Ansatz basierend auf dem EVM und einer FFT einen durchschnittlichen RMSE von 15,1829 [Che+20]. Obwohl das Gesicht segmentiert wird, werden keine besseren Ergebnisse erreicht. Das beste Ergebnis in dieser Arbeit mit einem RMSE von 14,07 BPM liegt jedoch darunter. Dennoch sind die klinischen Anforderungen mit einer maximalen Abweichung von ± 5 BPM damit nicht erfüllt.

7.7 Herzratendetektion mit Radar-Sensor

7.7.1 Einordnung der Ergebnisse

Wie zu erwarten, sind die Ergebnisse am Baby-Thorax-Simulator schlechter als mit der Software Simulation. Durch die verwendete Hardware treten Störeinflüsse auf. Darüber hinaus ist der Baby-Thorax-Simulator nicht in der Lage, die identische Wellenform wiederzugeben, die softwaremäßig simuliert wurde. Die getesteten Algorithmen, die alle auf einer FFT basieren, liefern nicht die benötigte Genauigkeit. Geeignere Algorithmen sind zu entwickeln.

7.7.2 Vergleich mit dem Stand der Technik

Zum Vergleich herangezogen wird die Arbeit von Lee et al., die mit einem IR-UWB-Radar-Sensor an 34 Neonaten Daten gesammelt haben. Dokumentiert wurde ein mittlerer Fehler von $-0,23$ BPM. [Lee+21] Bei dem Software Test mit dem Algorithmus FFT per band wird ein Median des RMSE von $4,61$ BPM erreicht. Dieser liegt unter der klinisch geforderten Genauigkeit von ± 5 BPM. Gemessen mit dem Mikrowellen interferometrischen Radar-Sensor wird am Baby-Thorax-Simulator ein RMSE Median von $22,71$ BPM erhalten, wenn die Berechnung ohne Ellipsen Fitting durchgeführt wird. Die in der Literatur angegebene Aussage, dass der Mikrowellen interferometrische Radar-Sensor genauere Ergebnisse bringt als UWB, konnte in der vorliegenden Untersuchung nicht gezeigt werden. Da unterschiedliche Referenzverfahren verwendet wurden, EKG vs. Laser-Mikrometer, sind die Ergebnisse nur eingeschränkt vergleichbar. Es wird erwartet, dass das hier vorgestellte Verfahren in der realen Anwendung zu ungenaue Messergebnisse liefern würde.

7.8 Temperaturdetektion

7.8.1 Einordnung der Ergebnisse

Während der Phototherapie tragen die Probanden eine Augenbinde. Je nachdem von welchem Typ und Größe diese ist, kann sie zu Fehlmessungen führen. Die große Binde lässt wenig Hautfläche frei, was zu erniedrigten Messwerten führen kann, weil die Oberfläche der Binde kühler ist als die Haut. Aus medizinischer Sicht kann es bei der großen Binde (siehe Abbildung 7.1) allerdings auch zu einem Wärmestau kommen, welcher die Erhöhung der Oberflächentemperatur des Kopfes verursachen kann. Dies führt zu erhöhten Messwerten mit der Thermokamera. Dieser Fehler kann auch durch ein Neuronales Netz, das Hautareale segmentiert, nicht verhindert werden. Wenn jedoch der Kopf des Kindes unbedeckt ist, kann die Thermokamera genaue Werte generieren, solange keine Intervention stattfindet. Bei Intervention kann es durch die geöffneten Klappen zu Temperaturschwankungen im Inkubator kommen, was aber durch den Referenz-Temperatursensor in der Berechnung berücksichtigt wird. Nur wenn der Kopf des Kindes zu stark verdeckt wird, ist eine korrekte Temperaturdetektion nicht mehr möglich. Diese Situation wird zu diesem Zeitpunkt aber als nicht klinisch relevant beurteilt, da das Kind gleichzeitig vom Pflegepersonal betreut wird.

Die hohe Genauigkeit zeigt, dass die verwendete Look-Up table für die Berechnung geeignet ist. Der Referenzsensor weist eine Genauigkeit von $0,1$ K auf und kompensiert damit den Fehler

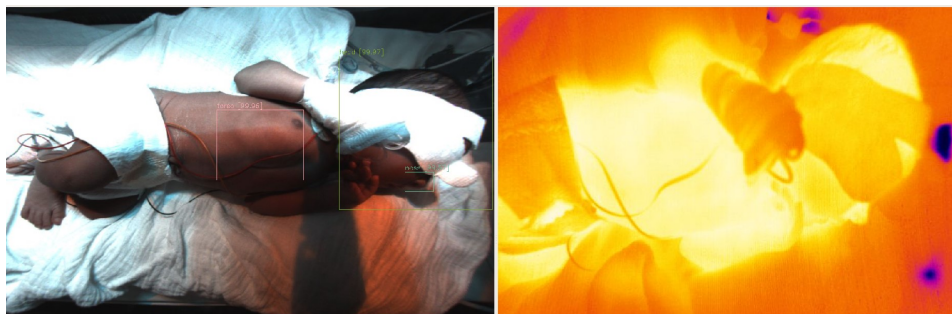


Abb. 7.1: Neonat mit breiter Binde im RGB-(li.) und Thermokamerabild (re.).

der Thermokamera von bis zu ± 2 K.

Die kontaktlose Temperaturmessung während Intervention könnte durch Verwendung der Methode mit dem Kalman-Filter verbessert werden, da dort mehr Umgebungsparameter wie z.B. Umgebungstemperatur und Luftfeuchtigkeit berücksichtigt werden. An eine Umsetzung ist allerdings erst zu denken, wenn die medizinische Forschung die benötigten Gewebekennzahlen der Neonaten bereitstellt.

7.8.2 Vergleich mit dem Stand der Technik

Bislang gibt es keine Veröffentlichungen, bei der ein kontaktloses Verfahren zur Ermittlung der Körperinnentemperatur von Neonaten vorgestellt wird. Daher liegen keine Vergleichswerte vor. Der kontaktlose Ansatz von Laxminarayan et al., der nur für Erwachsene ausgelegt ist und mit definierten Aktivitäten der Probanden arbeitet, liefert einen RMSE von 0,33 K [Lax+18]. Der in dieser Arbeit geringste RMSE von 0,12 K liegt noch darunter.

7.9 Fazit

Als bestes System erwies sich in der Validierungsphase eine Kombination aus folgenden Sensoren und Algorithmen:

ToF-Kamera zur Messung der Atemrate: Bei Messungen der Atemrate mit der ToF-Kamera liegt die geringste Abweichung bei 1,28 BPM bei einer Atemrate von 30 BPM, wenn das Kind ruhig schläft. Bei mittlerer und hoher Aktivität steigt der Fehler stark an. Die klinische Anforderung an die Genauigkeit der Messung der Atemrate liegt bei einer maximalen Abweichung von ± 2 BPM bei sedierten Kindern.

Thermokamera zur Messung der Körpertemperatur: Die Temperaturdetektion mit Thermokamera und Referenztemperatursensor erzielt einen Median der Abweichung von 0,0197 K und einen RMSE von 0,12 K. Die klinische Anforderung liegt bei $\pm 0,1$ K.

Radar zur Messung der Herzrate: Die Herzratenmessung mit einer RGB-Kamera erzielt auch unter guten Lichtverhältnissen am schlafenden Kind ein RMSE von 14,07 BPM und erfüllt damit nicht die erforderliche Genauigkeit. Die vielversprechende Messung der Herzrate mit Radar durfte nicht am Kind getestet werden. Der Softwaretest ergibt einen RMSE von 4,61 BPM, was unter der klinischen Anforderung von ± 5 BPM maximaler Abweichung liegt. Ein mit fusionierten Bilddaten trainiertes **YOLOv3-Neuronales Netz** wurde als beste und hochgenaue Lösung gefunden, um die geeigneten Messzonen automatisch auszuwählen.

8 Zusammenfassung

Frühgeborene sind besonders empfindlich. Es ist erforderlich, sie während ihres Aufenthalts auf einer neonatologischen Station umfassend zu überwachen. Allerdings sind sie leicht durch die herkömmlichen Monitoring-Techniken, die Hautkontakt erfordern, zu schädigen. Daher ergibt sich die Frage, ob es möglich ist, die Vitalparameter von Neu- und Frühgeborenen sicher und ausreichend genau, aber kontaktlos zu erfassen. Die wesentlichen Kriterien für die Verfassung des Kindes sind im klinischen Monitoring die Atemrate, die Herzrate und die Körpertemperatur. In der Literatur wird von unterschiedlichen Ansätzen zur kontaktlosen Überwachung berichtet. Dazu werden Sensoren und Kameras wie RGB, Thermo, Radar und Tiefenkameras eingesetzt. Es gibt bisher kein Gesamtsystem, das für Frühgeborene geeignet ist. Entweder sind die Verfahren nur für Erwachsene entwickelt, nicht unter klinischen Verhältnissen im Inkubator einsetzbar, nicht genau oder nicht zur Auswertung von Echtzeitdaten fähig.

Von den besten Ansätzen des Standes der Technik ausgehend wurden neue Methoden und Algorithmen entwickelt.

Der implementierte Entwurf einer Multi-Sensorbox bot die Basis, alternative Verfahren zur Erfassung der Atemrate per ToF-Kamera, Radar bzw. Thermokamera, die Messung der Herzrate mit RGB-Kamera oder Radar und die Temperatur mittels Thermokamera zu testen. Zur automatischen Detektion der jeweils relevanten Regions of Interest wurden verschiedene Neuronale Netze mit Bilddaten von Neonatologie-Patienten trainiert.

Die vielversprechendste Lösung umfasst die Messung der Atemrate mit der ToF-Kamera und die Temperaturdetektion mit Thermokamera und Referenztemperatursensor, sowie die Messung der Herzrate mit Radar, welche bisher nicht am Kind getestet werden durfte. Dabei wird zur automatischen Detektion der passenden Messzonen ein mit fusionierten Bilddaten trainiertes YOLOv3-Neuronales Netz verwendet.

Aus den Ergebnissen lässt sich schließen, dass es möglich ist, ein kontaktloses Überwachungssystem für Frühgeborene aufzubauen, das bei optimalen Bedingungen die klinischen Anforderungen erfüllt, aber noch nicht robust genug gegenüber Störeinflüssen wie starke Bewegung des Kindes, Intervention und schlechte Lichtverhältnisse ist. Die Robustheit kann durch weiteres Training der Neuronalen Netze erhöht werden. Dabei sollte eine Segmentierung von Hautarealen berücksichtigt werden. Die Herzratendetektion mittels Radar sowie die Kombination von ToF und Radar zur Detektion der Atemrate sollten weiterentwickelt und klinisch getestet werden.

Literaturverzeichnis

- [17] „Schnelle Fourier-Transformation (FFT) und Fensterfunktion“. Zugriff am 17. Januar 2019. 2017. <http://www.ni.com/white-paper/4844/de/%5C#toc2>.
- [Aar+13] L. A. Aarts, V. Jeanne, J. P. Cleary, C. Lieber u. a. „Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit - A pilot study“. In: *Early Human Development* 89.12 (2013), S. 943–948. <http://dx.doi.org/10.1016/j.earlhumdev.2013.09.016>.
- [Abb+11] A. K. Abbas, K. Heimann, K. Jergus, T. Orlikowsky und S. Leonhardt. „Neonatal non-contact respiratory monitoring based on real-time infrared thermography“. In: *BioMedical Engineering OnLine* 5.4 (2011), S. 1124–35. <http://www.biomedical-engineering-online.com/content/10/1/93%7B%5C%7D5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3985994%7B%5C%7Dtool=pmcentrez%7B%5C%7Drend-ertype=abstract>.
- [Abb+12] A. K. Abbas, K. Heimann, V. Blazek, T. Orlikowsky und S. Leonhardt. „Neonatal infrared thermography imaging: Analysis of heat flux during different clinical scenarios“. In: *Infrared Physics and Technology* 55.6 (2012), S. 538–548. <http://dx.doi.org/10.1016/j.infrared.2012.07.001>.
- [AC16] A. Al-Naji und J. Chahl. „Remote respiratory monitoring system based on developing motion magnification technique“. In: *Biomedical Signal Processing and Control* 29 (2016), S. 1–10. <http://dx.doi.org/10.1016/j.bspc.2016.05.002>.
- [AC17] J. K. A. Prochazka H. Charvatova. Vysata und J. Chambers. „Breathing Analysis Using Thermal and Depth Imaging Camera Video Records“. In: (2017).
- [Ada+00] A. K. Adams, R. A. Nelson, E. F. Bell und C. A. Egoavil. „Use of infrared thermographic calorimetry to determine energy expenditure in preterm infants“. In: *American Journal of Clinical Nutrition* 71.4 (2000), S. 969–977.
- [AL14] A. K. Abbas und S. Leonhardt. „Intelligent neonatal monitoring based on a virtual thermal sensor“. In: *BMC Medical Imaging* 14.1 (2014).
- [Ale] AlexeyAB. „Darknet“. <https://github.com/AlexeyAB/darknet>. Zugriff am 9. August 2024.
- [Ale+22] Q. G. Alexander, V. Hoskere, Y. Narazaki, A. Maxwell und B. F. Spencer. „Fusion of thermal and RGB images for automated deep learning based crack detection in civil infrastructure“. In: *AI in Civil Engineering* 1.1 (2022), S. 1–10. <https://doi.org/10.1007/s43503-022-00002-y>.
- [Ant+20a] C. H. Antink, J. C. M. Ferreira, M. Paul, S. Lyra u. a. „Fast body part segmentation and tracking of neonatal video data using deep learning“. en. In: *Med. Biol. Eng. Comput.* 58.12 (Dez. 2020), S. 3049–3061.

- [Ant+20b] E. Antolinos, F. García-Rial, C. Hernández, D. Montesano, J. I. Godino-Llorente und J. Grajal. „Cardiopulmonary Activity Monitoring Using Millimeter Wave Radars“. In: *Remote Sensing* 12.14 (2020). <https://www.mdpi.com/2072-4292/12/14/2265>.
- [arn] arntanguy. „C++ Implementation of Savitzky-Golay filtering based on Gram polynomials“. Zugriff am: 26. Oktober 2019. <https://github.com/arntanguy/gram%20savitzky%20golay>.
- [Art] I. for Artificial Intelligence - University of Bremen. „pico_flexx_driver“. Zugriff am 29. April 2020. https://github.com/code-iai/pico_flexx_driver.
- [Bac+18] V. Bach, S. McCabe, C. Abbiss, L. Barcat und P. Tourneux. „Distal skin vasodilation precedes sleep onset in preterm neonates and children“. Okt. 2018.
- [Beb+06] G. Bebis, A. Gyaourova, S. Singh und I. Pavlidis. „Face recognition by fusing thermal infrared and visible imagery“. In: *Image and Vision Computing* 24.7 (2006), S. 727–742. <http://linkinghub.elsevier.com/retrieve/pii/S0262885606000527>.
- [Bel+22] G. Beltrão, R. Stutz, F. Hornberger, W. A. Martins u. a. „Contactless radar-based breathing monitoring of premature infants in the neonatal intensive care unit“. In: *Scientific Reports* 12.1 (2022), S. 1–15. <https://doi.org/10.1038/s41598-022-08836-3>.
- [Bel99] F. Bellifemine. „Infrared thermometer“. Patent WO 1999039166 A1. 1999.
- [Ben+14] F. Benetazzo, A. Freddi, A. Monteriù und S. Longhi. „Respiratory rate detection algorithm based on RGB-D camera: theoretical background and experimental results“. In: *Healthcare Technology Letters* 1.3 (2014), S. 81–86.
- [Bep+22] F. Beppu, H. Yoshikawa, A. Uchiyama, T. Higashino, K. Hamada und E. Hirakawa. „Body part detection from neonatal thermal images using deep learning“. In: *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*. Lecture notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Cham: Springer International Publishing, 2022, S. 438–450.
- [Ber+14] N. Bernacchia, L. Scalise, L. Casacanditella, I. Ercoli, P. Marchionni und E. P. Tomasini. „Non contact measurement of heart and respiration rates based on Kinect™“. In: *2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*. Juni 2014, S. 1–5.
- [Bet+13] J. Betts, P. Desaix, E. Johnson, J. Johnson u. a. „Anatomy & Physiology“. Open Textbook Library. OpenStax College, Rice University, 2013. <https://books.google.de/books?id=dvVgngEACAAJ>.
- [BGK15] S. L. Bennett, R. Goubran und F. Knoefel. „The detection of breathing behavior using Eulerian-enhanced thermal video“. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2015-November* (2015), S. 7474–7477.

- [BGK16] S. L. Bennett, R. Goubran und F. Knoefel. „Adaptive eulerian video magnification methods to extract heart rate from thermal video“. In: *2016 IEEE International Symposium on Medical Measurements and Applications, MeMeA 2016 - Proceedings* (2016).
- [Bha+76] M. Bhatia, R. Poley, J. Haberman und D. Boon. „Abdominal thermography in infantile and childhood liver disease“. In: *South Med J* (1976), 69(8):1045–8.
- [Bie] U. Bielefeld. „Grundlagen zur Handstrahlung“. Zugriff am 26. Juli 2019. <https://www.uni-bielefeld.de/Universitaet/Einrichtungen/Zentrale%20Institute/IWT/FWG/Handstrahlung/Grundlagen-Handstrahlung.html>.
- [Bit10] S. Bittmann. „Checklisten Pädiatrie und Neonatologie“. In: *Checklisten Pädiatrie und Neonatologie*. Mnchen: Elsevier GmbH, Urban & Fischer Verlag, 2010, S. 173–182.
- [Bla+14] N. Blanik, A. K. Abbas, B. Venema, V. Blazek und S. Leonhardt. „Hybrid optical imaging technology for long-term remote monitoring of skin perfusion and temperature behavior“. In: *Journal of Biomedical Optics* 19.1 (2014), S. 016012. <http://biomedicaloptics.spiedigitallibrary.org/article.aspx?doi=10.1117/1.JBO.19.1.016012>.
- [BLY10] Y. W. Bai, W. T. Li und C. H. Yeh. „Design and implementation of an embedded monitor system for body breath detection by using image processing methods“. In: *ICCE 2010 - 2010 Digest of Technical Papers International Conference on Consumer Electronics* (2010), S. 193–194.
- [BMB] BMBF. „GUARDIAN“. Zugriff am 5. August 2019. <https://www.technik-zum-menschen-bringen.de/projekte/guardian>.
- [Bro+11] T. Brox, A. Bruhn, N. Papenberg und J. Weickert. „High Accuracy Optical Flow Estimation Based on a Theory for Warping“. In: 4.May (2011), S. 25–36.
- [BSF94] Y. Bengio, P. Simard und P. Frasconi. „Learning long-term dependencies with gradient descent is difficult“. In: *IEEE Transactions on Neural Networks* 5.2 (1994), S. 157–166.
- [BTM13] M. Bartula, T. Tigges und J. Muehlsteff. „Camera-based system for contactless monitoring of respiration“. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (2013), S. 2672–2675.
- [Bun] Bundesnetzagentur. „Funkanwendungen auf den ISM-Bändern“. Zugriff am 1. August 2019. <https://emf3.bundesnetzagentur.de/pdf/ISM-BNetzA.pdf>.
- [Cao+18] Y. Cao, B. Xu, Z. Ye, J. Yang u. a. „Depth and thermal sensor fusion to enhance 3D thermographic reconstruction“. In: *Optics Express* 26.7 (2018), S. 8179.
- [Cen+15] A. Cenci, D. Liciotti, E. Frontoni, A. Mancini und P. Zingaretti. „Non-contact monitoring of preterm infants using rgb-d camera“. In: *Proceedings of the ASME 2015 International Design Engineering Technical Conferences Computers and Information in Engineering Conference*. 2015.

- [Cha+19] S. Chaichulee, M. Villarroel, J. Jorge, C. Arteta u. a. „Cardio-respiratory signal extraction from video camera data for continuous non-contact vital sign monitoring using deep learning.“ In: *Physiological measurement* (2019). <http://www.ncbi.nlm.nih.gov/pubmed/31661680>.
- [Cha+21] P. H. Charlton, T. Bonnici, L. Tarassenko, D. A. Clifton u. a. „Biomedical Signal Processing and Control An impedance pneumography signal quality index : Design , assessment and application to respiratory rate monitoring“. In: *Biomedical Signal Processing and Control* 65.September 2020 (2021), S. 102339. <https://doi.org/10.1016/j.bspc.2020.102339>.
- [Che+20] Q. Chen, X. Jiang, X. Liu, C. Lu, L. Wang und W. Chen. „Non-Contact Heart Rate Monitoring in Neonatal Intensive Care Unit using RGB Camera“. In: (2020), S. 5822–5825.
- [Che+22] X. Chen, H. Wang, Y. Liang, Y. Meng und S. Wang. „A novel infrared and visible image fusion approach based on adversarial neural network“. In: *Sensors* 22.1 (2022).
- [Chr+03] I. Christidis, H. Zotter, H. Rosegger, H. Engele, R. Kurz und R. Kerbl. „Infrared Thermography in Newborns: The First Hour after Birth“. In: *Gynäkologisch-geburtshilfliche Rundschau* 43.1 (2003), S. 31–35. <https://www.karger.com/DOI/10.1159/000067168>.
- [cla] claydergc. „Find-peaks“. Zugriff am 26. Oktober 2019. <https://github.com/claydergc/find-peaks>.
- [Con] W. Contributors. „MIMO Radar“. https://en.wikipedia.org/w/index.php?title=MIMO_radar&oldid=1036834970. Zugriff am 27. Oktober 2023.
- [Cor] T. Corporation. „Terravic Facial Infrared Database“. Zugriff am 4. Januar 2019. <http://vcipl-okstate.org/pbvs/bench/Data/04/download.html>.
- [Cor+20] C. Coronel, C. Wiesmeyr, H. Garn, B. Kohn u. a. „3D Camera and Pulse Oximeter for Respiratory Events Detection“. In: *IEEE Journal of Biomedical and Health Informatics* (2020), S. 1–1.
- [Cra+16] M. Crawley, M. G. Smith, I. Gross, M. Khachaturian u. a. „Non-touch optical detection of biological vital signs“. USA Patent US 9305350. 2016.
- [CSS14] B. Y. R. P. Clark, J. K. Stotherst und T. Street. „Department of Physiology , the London Hospital Medical College “. In: 50.1980 (2014), S. 323–333.
- [CYN14] Y. K. Cheong, V. V. Yap und H. Nisar. „A novel face detection algorithm using thermal imaging“. In: *2014 IEEE Symposium on Computer Applications and Industrial Electronics (ISCAIE)*. 2014, S. 208–213.
- [CYN15] Y. K. Cheong, V. V. Yap und H. Nisar. „A novel face detection algorithm using thermal imaging“. In: *ISCAIE 2014 - 2014 IEEE Symposium on Computer Applications and Industrial Electronics* (2015), S. 208–213.
- [dat] databricks. „AdaGrad“. <https://www.databricks.com/de/glossary/adagrad>. Zugriff am 29. Juli 2024.

- [Dav+14] S. Davis, P. Watkinson, A. Guazzi, K. McCormick u. a. „Continuous non-contact vital sign monitoring in neonatal intensive care unit“. In: *Healthcare Technology Letters* 1.3 (2014), S. 87–91. <http://digital-library.theiet.org/content/journals/10.1049/htl.2014.0077>.
- [Den+19] J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia und S. Zafeiriou. „RetinaFace: Single-stage Dense Face Localisation in the Wild“. 2019. <https://arxiv.org/abs/1905.00641>.
- [Dos+22] Y. S. Dosso, D. Kyrollos, K. J. Greenwood, J. Harrold und J. R. Green. „NICUface: Robust neonatal face detection in complex NICU scenes“. In: *IEEE Access* 10 (2022), S. 62893–62909.
- [Dro06] A. D. Droitcour. „Non-contact Measurement Of Heart And Respiration rates With A Single-chip Microwave Doppler Radar“. 2006.
- [DS65] M. Dawkins und J. Scopes. „Non-shivering Thermogenesis and Brown Adipose Tissue in the Human New-born Infant“. In: *Nature* 206 (1965), S. 201–202. <https://doi.org/10.1038/206201b0>.
- [EBB05] J. Erderle, S. Blanchard und J. Bronziner. „Introduction to Biomedical Engineering“. 2005.
- [EFC] EFCNI. „Thermoregulation bei Frühgeborenen“. https://www.efcni.org/wp-content/uploads/2021/06/2021_06_11_Thermoregulation_Factsheet_DE_Final.pdf. Zugriff am 23. August 2024.
- [Ek+99] J. R. Ek, E. F. Bell, R. A. Nelson und M. A. Radhi. „Infrared thermographic calorimetry applied to preterm infants under radiant warmers“. In: 1999.
- [Ele] ElectronicsTutorials. „Passive Bandpassfilter“. <https://www.electronics-tutorials.ws/de/filtern/passiver-bandpassfilter.html>. Zugriff am 16. Juli 2024.
- [Emm] J.-F. Emmrich. „What is Structured-Light Scanning?“ Zugriff am 13. November 2019. <https://blog.medit.com/medit/what-is-structured-light-scanning>.
- [Exp] B. Experte. „Babyphone im Test“. Zugriff am 26. Juli 2019. <https://babyphone-experte.de/digitale-babyphones/>.
- [Eyc+10] L. G. T. Eyck, L. E. Lynam, H. Xia und K.-L. J. Deng. „Temperature sensing fabric“. USA Patent US 7787726 B2. 2010.
- [Fai13] M. D. Fairchild. „Color Appearance Models“. en. 3. Aufl. The Wiley-IS&T Series in Imaging Science and Technology. Nashville, TN: John Wiley & Sons, Aug. 2013.
- [FB08] T. Frey und M. Bossert. „Signal- und Systemtheorie“. In: Vieweg+Teubner, 2008.
- [FFT] FFTW. „FFTW“. <http://fftw.org/index.html>. Zugriff am 11. Februar 2021.
- [Fiz] Fizyr. „Keras-Retinanet“. <https://github.com/fizyr/keras-retinanet>. Zugriff am 9. August 2024.

- [FMH] A. Flemmer, R. Maier und H. Hummler. „Behandlung der neonatalen Asphyxie unter besonderer Berücksichtigung der therapeutischen Hypothermie“. Zugriff am 15. September 2020. https://www.awmf.org/uploads/tx_szleitlinien/024-0231_S2k_Behandlung_der_neonatalen_Aphyxie_unter_besonderer_Ber%C3%BCcksichtigung_der_therapeutischen_Hypothermie_2013-06-abgelaufen.pdf.
- [Fra97] J. Fraden. „Noncontact active temperature sensor“. USA Patent US 5645349 A. 1997.
- [Gab+18] G. Gabrielle, C. Sitthichok, V. Mauricio, J. João u. a. „Localised photoplethysmography imaging for heart rate estimation of pre-term infants in the clinic“. In: *Optical Diagnostics and Sensing XVIII: Toward Point-of-Care Diagnostics*. Optical Diagnostics and Sensing XVIII: Toward Point-of-Care Diagnostics. Hrsg. von C. G. L. San Francisco, United States: SPIE, 20. Feb. 2018, S. 26. <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10501/2289759/Localised-photoplethysmography-imaging-for-heart-rate-estimation-of-pre-term/10.1117/12.2289759%20.full> (besucht am 12. 07. 2021).
- [Gal] O. Gal. „fit_ellipse“. https://www.mathworks.com/matlabcentral/fileexchange/3215-fit_ellipse. Zugriff am 30. Juli 2021.
- [GB10] X. Glorot und Y. Bengio. „Understanding the difficulty of training deep feedforward neural networks“. In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. Hrsg. von Y. W. Teh und M. Titterington. Bd. 9. Proceedings of Machine Learning Research. Chia Laguna Resort, Sardinia, Italy: PMLR, 13–15 May 2010, S. 249–256. <https://proceedings.mlr.press/v9/glorot10a.html>.
- [GBC16] I. Goodfellow, Y. Bengio und A. Courville. „Deep Learning“. <http://www.deeplearningbook.org>. MIT Press, 2016.
- [Ghe+17] M. Ghezzi, L. Tenero, M. Piazza, A. Bodini und G. Piacentini. „Structured Light Plethysmography (SLP): Management and follow up of a paediatric patient with pneumonia“. In: *Respiratory Medicine Case Reports* 22 (2017), S. 67–69. <https://doi.org/10.1016/j.rmcr.2017.06.004>.
- [Gle+] J. Gleichauf, S. Herrmann, C. Niebler und A. Koelpin. „A Time-of-Flight and radar dataset of a neonatal thorax simulator with synchronized reference sensor signals for respiratory rate detection“. <https://figshare.com/s/1be0bb0a7772f15e9d29>. Zugriff am 30. August 2023.
- [Gle+20] J. Gleichauf, J. Vollet, C. Pfitzner, P. Koch und S. May. „Sensor Fusion Approach for an Autonomous Shunting Locomotive“. In: *Informatics in Control, Automation and Robotics*. Hrsg. von O. Gusikhin und K. Madani. Cham: Springer International Publishing, 2020, S. 603–624.
- [Gle+21] J. Gleichauf, S. Herrmann, L. Hennemann, H. Krauss u. a. „Automated Non-Contact Respiratory Rate Monitoring of Neonates Based on Synchronous Evaluation of a 3D Time-of-Flight Camera and a Microwave Interferometric Radar Sensor“. In: *Sensors* 21.9 (2021). <https://www.mdpi.com/1424-8220/21/9/2959>.

- [Gle+23] J. Gleichauf, L. Hennemann, F. B. Fahlbusch, O. Hofmann, C. Niebler und A. Koelpin. „Sensor Fusion for the Robust Detection of Facial Regions of Neonates Using Neural Networks“. In: *Sensors* 23.10 (2023). <https://www.mdpi.com/1424-8220/23/10/4910>.
- [GNK20] J. Gleichauf, C. Niebler und A. Koelpin. „Automatic non-contact monitoring of the respiratory rate of neonates using a structured light camera“. In: *42nd Annual International Conferences of the IEEE Engineering in Medicine and Biology Society (EMBC) (accepted)*. 2020.
- [Gon] C. Gonzalez. „Find-peaks“. <https://github.com/claydergc/find-peaks>. Zugriff am 10. Februar 2021.
- [Gor90] P. A. Gorry. „General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method“. In: *Analytical Chemistry* 62.6 (1990), S. 570–573. <https://doi.org/10.1021/ac00205a007>.
- [Gre+18] G. Green, S. Chaichulee, M. Villarroel, J. Jorge u. a. „Localised photoplethysmography imaging for heart rate estimation of pre-term infants in the clinic“. In: November 2019 (2018), S. 26.
- [Gua+18] L. Guanming, W. Shiyun, K. Kaiting, Y. Jingjie, L. Haibo und L. Xiaonan. „Learning Pyramidal Hierarchical Features for Neonatal Face Detection“. In: *2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*. 2018, S. 8–13.
- [Guo+21] J. Guo, J. Deng, A. Lattas und S. Zafeiriou. „Sample and Computation Redistribution for Efficient Face Detection“. In: *CoRR* abs/2105.04714 (2021). <https://arxiv.org/abs/2105.04714>.
- [Gup11] A. K. Gupta. „Respiration Rate Measurement Based on Impedance Pneumography“. In: *Application Report SBAA181, Texas Instruments* (2011).
- [Han+17] X.-F. Han, J. S. Jin, M.-J. Wang, W. Jiang, L. Gao und L. Xiao. „A review of algorithms for filtering the 3D point cloud“. In: *Signal Processing: Image Communication* 57 (2017), S. 103–112. <http://www.sciencedirect.com/science/article/pii/S0923596517300930>.
- [hao] hao56779997. „fft hilbert transformation“. Zugriff am 28. Oktober 2019. <https://gist.github.com/hao56779997/3121356>.
- [Has+17] M. A. Hassan, A. S. Malik, D. Fofi, N. Saad und F. Meriaudeau. „Novel health monitoring method using an RGB camera“. In: *Biomedical Optics Express* 8.11 (2017), S. 4838. <http://dx.doi.org/10.1364/BOE.8.004838>.
- [Hau+22] J. Hausmann, M. S. Salekin, G. Zamzmi, D. Goldgof und Y. Sun. „Robust Neonatal Face Detection in Real-world Clinical Settings“. 2022. <https://arxiv.org/abs/2204.00655>.
- [He+16a] K. He, X. Zhang, S. Ren und J. Sun. „Deep residual learning for image recognition“. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016-December* (2016), S. 770–778.

- [He+16b] K. He, X. Zhang, S. Ren und J. Sun. „Identity mappings in deep residual networks“. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9908 LNCS (2016), S. 630–645.
- [Hei+13] K. Heimann, K. Jergus, A. K. Abbas, N. Heussen, S. Leonhardt und T. Orlikowsky. „Infrared thermography for detailed registration of thermoregulation in premature infants“. In: *Journal of Perinatal Medicine* 41.5 (2013), S. 613–620.
- [Hen23] L. Hennemann. „Realisierung und Optimierung der Detektion von Körperregionen Neugeborener zur kontaktlosen und robusten Überwachung der Vitalparameter mittels eines Neuronalen Netzes“. Magisterarb. Nuremberg Institute of Technology, 2023.
- [Her21] S. Herrmann. „Simultaneous measurement of the heart and respiratory rate of a neonatal thorax simulator with a microwave interferometric radar sensor“. 2021.
- [HFB12] C. L. Herry, M. Frize und E. Bariciak. „Assessment of Abdominal Skin Temperature Change in Premature Newborns with NEC Compared to Healthy Controls“. In: *5th European Conference of the International Federation for Medical and Biological Engineering*. Hrsg. von Á. Jobbágy. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, S. 191–194.
- [HG] C. Hartmann und J. Gleichauf. „ros_cvb_camera_driver“. Zugriff am 07. Juni 2019. http://wiki.ros.org/ros_cvb_camera_driver.
- [HJ11] R. Huch und K. D. Jürgens. „Mensch Körper Krankheit“. In: München: Elsevier GmbH, Urban und Fischer Verlag, 2011.
- [Hme+18] H. Hmeidi, S. Motamedi-Fakhr, E. K. Chadwick, F. J. Gilchrist u. a. „Tidal breathing parameters measured by structured light plethysmography in children aged 2–12 years recovering from acute asthma/wheeze compared with healthy children“. In: *Physiological Reports* 6.12 (2018), S. 1–13.
- [HR14] J. Heerlein und T. Rügheimer. „elektroniknet.de“. Zugriff im Februar 2018. 2014. <http://www.elektroniknet.de/optoelektronik/licht-sensoren-fuer-dieselbstvermessung-%20113430.html>.
- [Hu+14] H. Hu, T. Wu, L. Yin, M. Kim u. a. „MRI detection of brown adipose tissue with low fat content in newborns with hypothermia“. In: *Magn Reson Imaging* (2014).
- [HU09] J. Hopmeier und G. Ullrich. „Die Transkapnode - ein Meßaufnehmer zur Transcutanen pCO₂-Messung mit Iridium-pH-Sensor“. In: *Biomedizinische Technik/Biomedical Engineering*, 30(s1) (2009), S. 94–95.
- [Ich] A. E. Ichim. „pcl::MedianFilter< PointT > Class Template Reference“. Zugriff am 20. Januar 2021. https://pointclouds.org/documentation/classpcl_1_1_median_filter.html.
- [ICN] ICNIRP. „Guidelines for limiting exposure to time-varying electric, magnetic and electromagnetic fields (100 kHz to 300 GHz“. Zugriff am 26. Juli 2019. https://www.icnirp.org/cms/upload/consultation_upload/ICNIRP_RF_Guidelines_PCD_2018_07_11.pdf.

- [inf] infinitioptics. „What Is a Visible Imaging Sensor (RGB Color Camera)?“ <https://www.infinitioptics.com/glossary/visibleimaging-sensor-400700-nm-colour-cameras>. Zugriff am 24. Januar 2024.
- [Ins] T. Instruments. „PHDC3022DEJT“. <https://www.ti.com/store/ti/en/p/product/?p=PHDC3022DEJT>. Zugriff am 7. April 2022.
- [IRC] IRCAM. „FFT Size“. Zugriff am 18. Januar 2019. <http://support.ircam.fr/docs/AudioSculpt/3.0/co/FFT%20Size.html>.
- [JA18] R. Joseph und F. Ali. „YOLOv3: An Incremental Improvement“. In: (8. Apr. 2018). <https://arxiv.org/abs/1804.02767v1> (besucht am 07. 04. 2021).
- [Jan+15] R. Janssen, W. Wang, A. Moço und G. De Haan. „Video-based respiration monitoring with automatic region of interest detection“. In: *Physiological Measurement* 37.1 (2015), S. 100–114.
- [Jar] Q. Jarosz. „Neuron_Hand-tuned“. <https://commons.wikimedia.org/w/index.php?curid=7616130/media/File:Neuron{ }Hand-tuned.svg>. Zugriff am 12. September 2023.
- [JL16] H. Jiang und E. Learned-Miller. „Face Detection with the Faster R-CNN“. 2016. <https://arxiv.org/abs/1606.03473>.
- [Joc+20] G. Jocher, A. Stoken, J. Borovec, NanoCode012 u. a. „ultralytics/yolov5: v3.1 - Bug Fixes and Performance Improvements“. Version v3.1. Okt. 2020. <https://doi.org/10.5281/zenodo.4154370>.
- [Jor+17] J. Jorge, M. Villarroel, S. Chaichulee, A. Guazzi u. a. „Non-Contact Monitoring of Respiration in the Neonatal Intensive Care Unit“. In: *Proceedings - 12th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2017 - 1st International Workshop on Adaptive Shot Learning for Gesture Understanding and Production, ASLAGUP 2017, Biometrics in the Wild, Bwild 2017, Heterogeneous Face Recognition, HFR 2017, Joint Challenge on Dominant and Complementary Emotion Recognition Using Micro Emotion Features and Head-Pose Estimation, DCER and HPE 2017 and 3rd Facial Expression Recognition and Analysis Challenge, FERA 2017* May (2017), S. 286–293.
- [Jor+18] J. Jorge, M. Villarroel, S. Chaichulee, K. McCormick und L. Tarassenko. „Data fusion for improved camera-based detection of respiration in neonates“. In: November 2019 (2018), S. 36.
- [Jor10] A. Jorch Gerhard und Hübler. „Neonatologie“. In: 2010.
- [JVV16] V. Jeyhani, T. Vuorinen und A. Vehkaoja. „Comparison of Simple Algorithms for Estimating Respiration Rate from Electrical Impedance Pneumography Signals in Wearable Devices“. In: December (2016), S. 0–16.
- [JZF20] C. Jung, K. Zhou und J. Feng. „Fusionnet: Multispectral fusion of RGB and NIR images using two stage convolutional neural networks“. In: *IEEE Access* 8 (2020), S. 23912–23919.
- [Kad+18] S. Kado, S. Member, Y. Monno, K. Moriwaki u. a. „Remote Heart Rate Measurement from RGB-NIR Video Based on Spatial and Spectral Face Patch Selection“. In: (2018), S. 5676–5680.

- [Kan+11] M. Kanti Bhowmik, Kankan, S. Majumder, G. Majumder u. a. „Thermal Infrared Face Recognition – A Biometric Identification Technique for Robust Security system“. In: *Reviews, Refinements and New Ideas in Face Recognition* (2011). <http://www.intechopen.com/books/reviews-refinements-and-new-ideas-in-face-recognition/thermal-infrared-face-recognition-a-biometric-identification-technique-for-ro-bust-security-system>.
- [Kat] A. Kathuria. „What’s new in YOLOv3?“. <https://towardsdatascience.com/yolo-v3-object-detection-53fb7d3bfe6b>. Zugriff am 12. Januar 2022.
- [Kau22] C. Kaula. „Hard- und Softwareintegration eines Sensors in ein kontaktloses Monitoringsystem mittels eines an ROS angebenen Mikrocontrollers zur Überwachung der Temperatur und Feuchtigkeit in einem Inkubator“. 2022.
- [KB06] J. Kannala und S. S. Brandt. „A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses“. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28.8 (2006), S. 1335–1340.
- [KGR11] R. B. Knobel, B. D. Guenther und H. E. Rice. „Thermoregulation and thermography in neonatal physiology and disease“. eng. In: *Biological research for nursing* 13.3 (Juli 2011), S. 274–282. <https://pubmed.ncbi.nlm.nih.gov/21586499/> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3775585/>.
- [Kha+21] F.-T.-Z. Khanam, A. G. Perera, A. Al-Naji, K. Gibson und J. Chahl. „Non-contact automatic vital signs monitoring of infants in a Neonatal Intensive Care Unit based on neural networks“. en. In: *J. Imaging* 7.8 (Juli 2021), S. 122.
- [Kim+09] Y. M. Kim, C. Theobalt, J. Diebel, J. Kosecka, B. Matusik und S. Thrun. „Multi-view image and ToF sensor fusion for dense 3D reconstruction“. In: *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops 2009* (2009), S. 1542–1546.
- [Kim+19] J. D. Kim, W. H. Lee, Y. Lee, H. J. Lee u. a. „Non-contact respiration monitoring using impulse radio ultrawideband radar in neonates“. In: (2019).
- [Kla+14] J. H. Klaessens, M. van den Born, A. van der Veen, J. Sikkens-van de Kraats, F. A. van den Dungen und R. M. Verdaasdonk. „Development of a baby friendly non-contact method for measuring vital signs: First results of clinical measurements in an open incubator at a neonatal intensive care unit“. In: 8935 (2014), 89351P. <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2038353>.
- [KNM17] M. Kopaczka, J. Nestler und D. Merhof. „Face Detection in Thermal Infrared Images: A Comparison of Algorithm- and Machine-Learning-Based Approaches“. In: *Advanced Concepts for Intelligent Vision Systems*. Hrsg. von J. Blanc-Talon, R. Penne, W. Philips, D. Popescu und P. Scheunders. Cham: Springer International Publishing, 2017, S. 518–529.
- [Kno+13] R. B. Knobel, J. Levy, L. Katz, B. Guenther und D. Holditch-Davis. „A pilot study to examine maturation of body temperature control in preterm infants“. eng. In: *Journal of obstetric, gynecologic, and neonatal nursing : JOGNN* 42.5 (2013), S. 562–574. <https://pubmed.ncbi.nlm.nih.gov/24004312/> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3783951/>.

- [Kno+17] R. B. Knobel-Dail, D. Holditch-Davis, R. Sloane, B. D. Guenther und L. M. Katz. „Body temperature in premature infants during the first week of life: Exploration using infrared thermal imaging“. In: *Journal of Thermal Biology* 69 (2017), S. 118–123.
- [Koe+16] A. Koelpin, F. Lurz, S. Linz, S. Mann, C. Will und S. Lindner. „Six-port based interferometry for precise radar and sensing applications“. In: *Sensors (Switzerland)* 16.10 (2016).
- [KS17] A. K. Krishnan und S. Saripalli. „Cross-Calibration of RGB and Thermal Cameras with a LIDAR for RGB-Depth-Thermal Mapping“. In: *Unmanned Systems* 5.2 (2017), S. 59–78.
- [KSJ95] E. Kandel, J. Schwartz und T. Jessell. „Neurowissenschaften“. Basingstoke, England: Spektrum Akademischer Verlag, 1995.
- [Kur+15] S. Kurath-Koller, G. Litscher, A. Gross, T. Freidl u. a. „Changes of locoregional skin temperature in neonates undergoing laser needle acupuncture at the acupuncture point large intestine 4“. eng. In: *Evidence-based complementary and alternative medicine : eCAM* 2015 (2015), S. 571857. <https://pubmed.ncbi.nlm.nih.gov/25922612%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4398935/>.
- [Kyr+21] D. G. Kyrollos, J. B. Tanner, K. Greenwood, J. Harrold und J. R. Green. „Non-contact Neonatal Respiration Rate Estimation Using Machine Vision“. In: *2021 IEEE Sensors Applications Symposium (SAS)*. 2021, S. 1–6.
- [Lae] Laerdal. „SimBaby“. <https://laerdal.com/de/products/simulation-training/obstetrics-paediatrics/simbaby/>. Zugriff am 23. Juli 2024.
- [Lan00] R. Lange. „3D time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology“. Diss. Universität Siegen, 2000. <https://dspace.ub.uni-siegen.de/handle/ubsi/178>.
- [Lax+18] S. Laxminarayan, V. Rakesh, T. Oyama, J. B. Kazman u. a. „Individualized estimation of human core body temperature using noninvasive measurements“. In: *Journal of Applied Physiology* 124.6 (2018), S. 1387–1402.
- [LCI19] E. Lauhkonen, B. G. Cooper und R. Iles. „Mini review shows that structured light plethysmography provides a non-contact method for evaluating breathing patterns in children“. In: *Acta Paediatrica, International Journal of Paediatrics* (2019), S. 1398–1405.
- [Lee+21] W. H. Lee, Y. Lee, J. Y. Na, S. H. Kim u. a. „Feasibility of non-contact cardio-respiratory monitoring using impulse-radio ultra-wideband radar in the neonatal intensive care unit“. In: *PLoS ONE* 15.12 December (2021), S. 1–15. <http://dx.doi.org/10.1371/journal.pone.0243939>.
- [Li+09] C. Li, J. Cummings, J. Lam, E. Graves und W. Wu. „Radar remote monitoring of vital signs“. In: *IEEE Microwave Magazine* 10.1 (2009), S. 47–56.
- [Lin+13] S. Linz, G. Vinci, S. Mann, S. Lindner u. a. „A compact, versatile six-port radar module for industrial and medical applications“. In: *Journal of Electrical and Computer Engineering* 2013 (2013).

- [Lin+14] T. Y. Lin, M. Maire, S. Belongie, J. Hays u. a. „Microsoft COCO: Common objects in context“. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 8693 LNCS.PART 5 (2014), S. 740–755.
- [Lin+17a] T. Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan und S. Belongie. „Feature pyramid networks for object detection“. In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017* 2017-January (2017), S. 936–944.
- [Lin+17b] T. Y. Lin, P. Goyal, R. Girshick, K. He und P. Dollar. „Focal Loss for Dense Object Detection“. In: *Proceedings of the IEEE International Conference on Computer Vision 2017-October* (2017), S. 2999–3007.
- [LLP80] J. Long, J. Lucey und A. Philip. „Noise and hypoxemia in the intensive care nursery“. In: *Pediatrics* (1980).
- [LPC14] T. Lukac, J. Pucik und L. Chrenko. „Contactless recognition of respiration phases using web camera“. In: *2014 24th International Conference Radioelektronika, RADIOELEKTRONIKA 2014 - Proceedings* (2014).
- [Lud15] P. Ludivig. „Face Detetection and Tracking for ROS“. Zugriff am 4. Januar 2019. 2015. https://github.com/phil333/face_detection.
- [Lyo+97] A. J. Lyon, M. E. Pikaar, P. Badger und N. McIntosh. „Temperature control in very low birthweight infants during first five days of life“. In: *Archives of Disease in Childhood - Fetal and Neonatal Edition* 76.1 (1997), F47–F50. <https://fn.bmj.com/content/76/1/F47>.
- [Lyo04] A. Lyon. „Applied physiology: temperature control in the newborn infant“. In: *Current Paediatrics* 14.2 (2004), S. 137–144. <https://www.science-direct.com/science/article/pii/S0957583903001465>.
- [LYT14] M. H. Li, A. Yadollahi und B. Taati. „A non-contact vision-based system for respiratory rate estimation“. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014* August 2015 (2014), S. 2119–2122.
- [mam] mamiweb. „Frühgeburt und Frühchen“. Zugriff: 13.12.2018. <http://www.mamiweb.de/nfs/premiummagazin/container/027/27739/frueh-chen-l.jpg?t=1374845400>.
- [Mat] Mathworks. „sgolayfilt“. <https://de.mathworks.com/help/signal/ref/sgolayfilt.html>. Zugriff am 23. August 2024.
- [Mäu] C. Mäurer. „Berechnung der spezifischen Absorptionsrate (SAR) dielektrischer Körper mit FEKO“. <https://www.iddd.de/umtsno/sarvortragkurz.pdf>. Zugriff am 14. August 2019.
- [May09] S. May. „3D Time-of-Flight Ranging for Robotic Perception in Dynamic Environments“. Diss. 2009, S. 136. <http://scholar.google.com/scholar?hl=en%7B%5C%7DbtnG=Search%7B%5C%7Dq=intitle:3D+Time-of-Flight+Ranging+for+Robotic+Perception+in+Dynamic+Environments%7B%5C%7D1>.

- [McC+] E. McCall, F. Alderdice, H. Halliday, S. Vohra und L. Johnston. „Maßnahmen bei der Geburt zur Vermeidung von Hypothermie bei Frühgeborenen und/oder Neugeborenen mit niedrigem Geburtsgewicht“. Zugriff am 17. September 2020. https://www.cochrane.org/de/CD004210/NEONATAL_massnahmen-bei-der-geburt-zur-vermeidung-von-hypothermie-bei-fruhgeborenen-undoder-neugeborenen-mit.
- [Mic] Microepsilon. „Optisches Präzisions-Mikrometer“. Zugriff am 15. September 2020. https://www.micro-epsilon.de/2D_3D/optical-micro-meter/micrometer/optoCONTROL_2520/.
- [Mit96] Mitchell. „xxx“. In: (1996).
- [MK22] W. Mucha und M. Kampel. „Depth and thermal images in face detection - A detailed comparison between image modalities“. In: *2022 the 5th International Conference on Machine Vision and Applications (ICMVA)*. Singapore Singapore: ACM, Feb. 2022.
- [MM] M. Murphy und G. McLoughlin. „Born too soon: preterm birht in Europe trends, causes and prevention“. Zugriff: 13.12.2018. http://www.euro.who.int/__data/assets/pdf_file/0004/277735/Born-too-soon_preterm-birth-in-Europe-trends,-causes-and-prevention.pdf?ua=1.
- [MMA] S. Merz, H. Mey und F. Antwerpes. „Vitalparameter“. Zugriff: 19.12.2018. <https://flexikon.doccheck.com/de/Vitalparameter>.
- [MP90] W. S. Mcculloch und W. Pitts. „A logical calculus nervous activity“. In: *Bulletin of Mathematical Biology* 52.1 (1990), S. 99–115.
- [MS17] M. Martinez und R. Stiefelhagen. „Breathing rate monitoring during sleep from a depth camera under real-life conditions“. In: *Proceedings - 2017 IEEE Winter Conference on Applications of Computer Vision, WACV 2017* (2017), S. 1168–1176.
- [MSB00] G. Matthews, B. Sudduth und M. Burrow. „A Non-Contact Vital Signs Monitor“. In: *Critical Reviews in Biomedical Engineering* 28.1-2 (2000), S. 173–178.
- [MSD19] A. Marnach, D. Schmiech und A. R. Dlewald. „Verification of algorithm for an I/Q-radar system for breathing detection in an incubator“. In: *Proceedings of the 2019 21st International Conference on Electromagnetics in Advanced Applications, ICEAA 2019* (2019), S. 1275–1278.
- [Muh+20] A. Muhammad, C. Chen, L. Xi, Y. Bin u. a. „Novel Framework: Face Feature Selection Algorithm for Neonatal Facial and Related Attributes Recognition“. In: *IEEE Access* 8 (2020), S. 59100–59113.
- [Mur+09] J. N. Murthy, J. Van Jaarsveld, J. Fei, I. Pavlidis u. a. „Thermal infrared imaging: A novel method to monitor airflow during polysomnography“. In: *Sleep* 32.11 (2009), S. 1521–1527. <http://www.embase.com/search/results?subaction=viewrecord%7B%5C%7Dfrom=export%7B%5C%7Ddid=L355715704%7B%5C%7D5Cnhttp://www.journalsleep.org/ViewAbstract.aspx?pid=27626%7B%5C%7D5Cnhttp://sfx.hul.harvard.edu/sfx%7B%5C%7Dlocal?>

- sid = EMBASE % 7B % 5C& % 7Dissn = 01618105 % 7B % 5C& % 7Did = doi : % 7B % 5C& % 7Datitle = Thermal + infrared + imaging % 7B % 5C % % 7D3A + A + no.
- [Nag+21] Á. Nagy, P. Földesy, I. Jánoki, D. Terbe u. a. „Continuous camera-based premature-infant monitoring algorithms for NICU“. en. In: *Appl. Sci. (Basel)* 11.16 (Aug. 2021), S. 7215.
- [Neg+18] T. Negishi, G. Sun, H. Liu, S. Sato, T. Matsui und S. Member. „Stable Contactless Sensing of Vital Signs Using RGB-Thermal Image Fusion System with Facial Tracking for Infection Screening“. In: (2018), S. 4371–4374.
- [Ngu+16] P. Nguyen, S. Transue, M. Choi, A. C. Halbower und T. Vu. „WiKiSpiro: Non-contact Respiration Volume Monitoring during Sleep“. In: *S3 Proceedings of the Eighth Wireless of the Students, by the Students, and for the Students Workshop*. 2016, S. 27–29.
- [NM08] N. Neophytou und K. Mueller. „Color-Space CAD: Direct Gamut Editing in 3D“. In: *IEEE Computer Graphics and Applications* 28 (2008).
- [Oca] D. T. Ocana. „image_pipeline“. Zugriff am 07. Juni 2019. https://github.com/DavidTorresOcana/image_pipeline.
- [Oka] K. Okada. „opencv_apps“. Zugriff am 4. Januar 2019. https://wiki.ros.org/opencv_apps#face_detection.
- [OKC94] G. O’Hara, J. Korff und P. Crill. „Body temperature thermometer and method for measuring human body temperature utilizing calibration mapping“. USA Patent US 5293877 A. 1994.
- [Ope] OpenCV. „Face Recognition with OpenCV“. Zugriff am 4. Januar 2019. https://www.docs.opencv.org/2.4/modules/contrib/doc/facerec/facerec_tutorial.html#face-recognition.
- [opea] openCV. „how to detect ellipse and get centers of ellipse“. <https://answers.opencv.org/question/38885/how-to-detect-ellipse-and-get-centers-of-ellipse/>. Zugriff am 12. Mai 2022.
- [opeb] openCV. „projectPoints“. https://docs.opencv.org/3.4/d9/d0c/group__calib3d.html#ga1019495a2c8d1743ed5cc23fa0daff8c. Zugriff am 12. Mai 2022.
- [opec] openCV. „Rodrigues“. https://docs.opencv.org/3.4/d9/d0c/group__calib3d.html#ga61585db663d9da06b68e70cfb-f6a1eac. Zugriff am 12. Mai 2022.
- [oped] openCV. „solvePnP“. https://docs.opencv.org/3.4/d9/d0c/group__calib3d.html#ga549c2075fac14829ff4a58bc931-c033d. Zugriff am 12. Mai 2022.
- [Oya+97] A. Oya, H. Asakura, T. Koshino und T. Araki. „Thermographic demonstration of nonshivering thermogenesis in human newborns after birth: its relation to umbilical gases“. 1997. <https://www.degruyter.com/view/j/jpme.1997.25.issue-5/jpme.1997.25.5.447/jpme.1997.25.5.447.xml>.
- [Par+19] J. Y. Park, Y. Lee, Y. W. Choi, R. Heo u. a. „Preclinical Evaluation of a Noncontact Simultaneous Monitoring Method for Respiration and Carotid Pulsation Using Impulse-Radio Ultra-Wideband Radar“. In: *Scientific Reports* 9.1 (2019), S. 1–12.

- [PCLa] PCL. Zugriff am 18. Januar 2021. https://pcl.readthedocs.io/projects/tutorials/en/latest/statistical_outlier.html.
- [PCLb] PCL. „Module sample_consensus“. https://pointclouds.org/documentation/group__sample__consensus.html. Zugriff am 10. Februar 2021.
- [PCLc] PCL. „Plane model segmentation“. Zugriff am 26. Oktober 2019. <http://www.pointclouds.org/documentation/tutorials/planarsegmentation.php>.
- [PCLd] PCL. „Plane model segmentation“. http://www.pointclouds.org/documentation/tutorials/planar_segmentation.html. Zugriff am 10. Februar 2021.
- [Pen+09] J. Penne, C. Schaller, J. Hornegger und K. T. „Journal Article Robust real-time 3D respiratory motion detection using time-of-flight cameras“. In: (2009), S. 8–9.
- [Per+17] C. B. Pereira, K. Heimann, B. Venema, V. Blazek, M. Czaplik und S. Leonhardt. „Estimation of respiratory rate from thermal videos of preterm infants“. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (2017), S. 3818–3821.
- [PES72] P. Perlstein, N. Edwards und J. Sutherland. „Age relationship to thermal patterns on the back of cold-stressed infants“. In: *Biol Neonate* (1972), 20:127–133.
- [Pfi18] C. Pfitzner. „Visual Human Body Weight Estimation with Focus on Medical Applications“. 2018.
- [Pis+14] S. Pisa, P. Bernardi, R. Cicchetti, R. Giusto u. a. „Comparison between UWB and CW radar sensors for breath activity monitoring“. In: *Radar Sensor Technology XVIII* 9077. February 2016 (2014), S. 907711.
- [PL15] C. Pantofaru und D. Lazewatsky. „face_detector“. Zugriff am 4. Januar 2019. 2015. https://wiki.ros.org/face_detector.
- [PLU77] J. J. Pomerance, R. L. Lieberman und C. T. Ukrainski. „Neonatal Thermography“. In: *Pediatrics* 59.3 (1977), S. 345–351. <https://pediatrics.aappublications.org/content/59/3/345>.
- [Pom01] F. Pompei. „Temporal artery temperature detector“. USA Patent US 6292685 B1. 2001.
- [Pro+16] A. Procházka, M. Schätz, O. Vyšata und M. Vališ. „Microsoft Kinect visual and depth sensors for breathing and heart rate analysis“. In: *Sensors (Switzerland)* 16.7 (2016), S. 1–11.
- [Qi+21] D. Qi, W. Tan, Q. Yao und J. Liu. „YOLO5Face: Why Reinventing a Face Detector“. 2021. <https://arxiv.org/abs/2105.12931>.
- [Rab89] L. R. Rabiner. „Tutorial on HMM and Applications“. 1989. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=18626>.
- [Rag95] M. Raguse. „Kinderheilkunde für Kinderkrankenschwestern und -pfleger“. In: Springer Verlag, 1995.

- [Rah+15] T. Rahman, A. T. Adams, R. V. Ravichandran, M. Zhang u. a. „DoppleSleep: A Contactless Unobtrusive Sleep Sensing System Using Short-Range Doppler Radar“. In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp '15. Osaka, Japan: Association for Computing Machinery, 2015, S. 39–50. <https://doi.org/10.1145/2750858.2804280>.
- [Ram16] K. G. Ramsundar. „Eulerian Motion Magnification“. Zugriff am 25. Januar 2019. 2016. <https://github.com/kgram007/Eulerian-Motion-Magnification>.
- [Reh+17] H. Rehouma, R. Noumeir, P. Jovet, W. Bouachir und S. Essouri. „A computer vision method for respiratory monitoring in intensive care environment using RGB-D cameras“. In: *2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*. Nov. 2017, S. 1–6.
- [Reh+18] H. Rehoumaa, R. Noumeira, W. Bouachirb, P. Jovetc und S. Essouri. „3D imaging system for respiratory monitoring in pediatric intensive care environment“. In: (2018).
- [Rei] K. Reimann. „Interview: ”Das Herz eines Neugeborenen ist nicht größer als eine Walnuss““. <http://www.gesundheitsberater-berlin.de/praxis/krankheiten-von-a-z/kardiologie-fur-kinder/interview-das-herz-eines-neugeborenen-ist-nicht-grosser-als-eine-walnuss--2>. Zugriff am 23. Februar 2021.
- [RF18] J. Redmon und A. Farhadi. „YOLOv3: An Incremental Improvement“. 2018.
- [RHK20] M. T. Rudrappa, R. Herschel und P. Knott. „Vital Parameters Detection of Non Stationary Human Subject Using MIMO Radar“. In: *2020 21st International Radar Symposium (IRS)*. 2020, S. 125–130.
- [Ric+10] H. E. Rice, C. L. Hollingsworth, E. Bradsher, M. E. Danko u. a. „Infrared Thermal Imaging (Thermography) of the Abdomen in Extremely Low Birthweight Infants“. In: 2010.
- [Rih+16] S. Rihana, E. Younes, D. Visvikis und H. Fayad. „Kinect2 - Respiratory movement detection study“. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2016-October* (2016), S. 3875–3878.
- [RML14] L. B. C. Rivera, L. O. Maynez und R. C. A. L??zaro. „Determination of temperature in neonates based in a face IR thermography image segmentation algorithm“. In: *2014 11th International Conference on Electrical Engineering, Computing Science and Automatic Control, CCE 2014* (2014).
- [Roc+16] I. Rocco Spremolla, M. Antunes, D. Aouada und B. Ottersten. „RGB-D and Thermal Sensor Fusion“. In: *3.Visigrapp* (2016), S. 610–617.
- [ROS] ROS. „ROS“. Zugriff am 14. Januar 2019. <http://www.ros.org/>.
- [ros] ros. „camera_calibration“. Zugriff am 14. Januar 2019. https://wiki.ros.org/camera_calibration.
- [RPL72] E. Rylander, H. Pribylova und J. Lind. „A thermographic study of infants exposed to cold“. In: *Acta Paediat Scand* (1972), 61:42–8.

- [RS89] G. Ramachandran und M. Singh. „Three-dimensional reconstruction of cardiac displacement patterns on the chest wall during the P, QRS and T-segments of the ECG by laser speckle interferometry“. In: *Medical and Biological Engineering and Computing* 27.5 (1989), S. 525–530. <https://doi.org/10.1007/BF02441473>.
- [Rus+08] R. B. Rusu, Z. C. Marton, N. Blodow, M. Dolha und M. Beetz. „Towards 3D Point cloud based object maps for household environments“. In: *Robotics and Autonomous Systems* 56.11 (2008). Semantic Knowledge in Robotics, S. 927–941. <http://www.sciencedirect.com/science/article/pii/S0921889008001140>.
- [Rus+14] O. Russakovsky, J. Deng, H. Su, J. Krause u. a. „ImageNet Large Scale Visual Recognition Challenge“. 2014. <https://arxiv.org/abs/1409.0575>.
- [Sa21] K. Shi und et al. „Contactless analysis of heart rate variability during cold pressor test using radar interferometry and bidirectional LSTM networks“. In: *Scientific Reports* (2021).
- [Sal+15] C. Salinas, R. Fernández, H. Montes und M. Armada. „A new approach for combining time-of-flight and RGB cameras based on depth-dependent planar projective transformations“. In: *Sensors (Switzerland)* 15.9 (2015), S. 24615–24643.
- [Sal+21] M. S. Salekin, G. Zamzmi, J. Hausmann, D. Goldgof u. a. „Multimodal neonatal procedural and postoperative pain assessment dataset“. In: *Data in Brief* 35 (2021), S. 106796. <https://www.sciencedirect.com/science/article/pii/S2352340921000809>.
- [Sch+13] B. Schleicher, I. Nasr, A. Trasser und H. Schumacher. „IR-UWB radar demonstrator for ultra-fine movement detection and vital-sign monitoring“. In: *IEEE Transactions on Microwave Theory and Techniques* 61.5 (2013), S. 2076–2085.
- [Sch+19] S. Schellenberger, K. Shi, F. Michler, F. Lurz, R. Weigel und A. Koelpin. „Respiration Extraction from Radar Heart Sound Measurements“. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2019, S. 6533–6536.
- [Sch+20] S. Schellenberger, K. Shi, T. Steigleder und et al. „A dataset of clinically recorded radar vital signs with synchronised reference sensor signals“. In: *Scientific Data, Nature* (2020).
- [Sev13] N. Sevi. „Implementierung einer Puls- und Herzschlagmessung durch Verstärkung der Bewegungsänderungen in Videosequenzen“. Bachelorarbeit. 2013.
- [Sha+15] R. Shaoqing, H. Kaiming, G. Ross und S. Jian. „Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks“. In: (4. Juni 2015). <https://arxiv.org/abs/1506.01497v3> (besucht am 21. 04. 2021).
- [Shi+20a] K. Shi, S. Schellenberger, F. Michler, T. Steigleder u. a. „Automatic Signal Quality Index Determination of Radar-Recorded Heart Sound Signals Using Ensemble Classification“. In: *IEEE Transactions on Biomedical Engineering* 67.3 (2020), S. 773–785.

- [Shi+20b] K. Shi, S. Schellenberger, C. Will, T. Steigleder u. a. „A dataset of radar-recorded heart sounds and vital signs including synchronised reference sensor signals“. In: *Scientific Data* 7.1 (2020), S. 1–12.
- [Shi+20c] S. S. Shivakumar, N. Rodrigues, A. Zhou, I. D. Miller, V. Kumar und C. J. Taylor. „PST900: RGB-Thermal Calibration, Dataset and Segmentation Network“. In: *Proceedings - IEEE International Conference on Robotics and Automation* (2020), S. 9441–9447.
- [Sil+19] G. Silva, R. Monteiro, A. Ferreira, P. Carvalho und L. Corte-Real. „Face Detection in Thermal Images with YOLOv3“. In: *Advances in Visual Computing*. Hrsg. von G. Bebis, R. Boyle, B. Parvin, D. Koracin u. a. Cham: Springer International Publishing, 2019, S. 89–99.
- [SIM] SIMCharacters. „Paul“. <https://simcharacters.com/produkte/paul/>. Zugriff am 29. Juli 2024.
- [Sit+17] C. Sitthichok, V. Mauricio, J. João, A. Carlos u. a. „Multi-Task Convolutional Neural Network for Patient Detection and Skin Segmentation in Continuous Non-Contact Vital Sign Monitoring“. In: *2017 12th IEEE International Conference on Automatic Face Gesture Recognition (FG 2017)*. 2017, S. 266–272.
- [SJP19] I. Shopovska, L. Jovanov und W. Philips. „Deep visible and thermal image fusion for enhanced pedestrian visibility“. In: *Sensors (Switzerland)* 19.17 (2019), S. 1–21.
- [SMD19] D. Schmiech, A. Marnach und A. R. Diewald. „Verification and first test measurement of a microwave-based vital sign monitor“. In: *Advances in Radio Science* 17.October (2019), S. 249–256.
- [Spa09] M. Sparshott. „Früh- und Neugeborene pflegen“. 2009.
- [SPM10] L. St-Laurent, D. Prévost und X. Maldague. „Fast and accurate calibration-based thermal / colour sensors registration“. In: section 1 (2010).
- [staa] stackoverflow. „opencv 3, blobdetection, The function/feature is not implemented () in detectAndCompute“. <https://stackoverflow.com/questions/30622304/opencv-3-blobdetection-the-function-feature-is-not-implemented-in-detectand>. Zugriff am 12. Mai 2022.
- [stab] stackoverflow. „What are the main references to the fish-eye camera model in OpenCV3.0.0dev?“ <https://stackoverflow.com/questions/31089265/what-are-the-main-references-to-the-fish-eye-camera-model-in-opencv3-0-0dev>. Zugriff am 5. November 2021.
- [STC16] D. B. Springer, L. Tarassenko und G. D. Clifford. „Logistic regression-HSMM-based heart sound segmentation“. In: *IEEE Transactions on Biomedical Engineering* 63.4 (2016), S. 822–832.
- [Str] B. für Strahlenschutz. „EIRP“. <https://www.bfs.de/SharedDocs/Glossareintraege/DE/A/eirp.html>. Zugriff am 03. Januar 2025.
- [STW85] A. V. Sahakian, W. J. Tompkins und J. G. Webster. „Electrode Motion Artifacts in Electrical Impedance Pneumography“. In: 1.6 (1985), S. 448–451.

- [SW08] A. K. Saxena und G. H. Willital. „Infrared thermography: Experience from a decade of pediatric imaging“. In: *European Journal of Pediatrics* 167.7 (2008), S. 757–764. <https://doi.org/10.1007/s00431-007-0583-z>.
- [SZ14] K. Simonyan und A. Zisserman. „Very Deep Convolutional Networks for Large-Scale Image Recognition“. 2014. <https://arxiv.org/abs/1409.1556>.
- [Tah+72] E. Tahti, J. Lind, L. Osterlund und E. Rylander. „Changes in skin temperature of the neonate at birth“. In: *Acta Paediat Scand* (1972), 61:159–64.
- [Tan+22] L. Tang, J. Yuan, H. Zhang, X. Jiang und J. Ma. „PIAFusion: A progressive infrared and visible image fusion network based on illumination aware“. In: *Information Fusion* 83-84.June (2022), S. 79–92.
- [Tar+14] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. A. Clifton und C. Pugh. „Non-contact video-based vital sign monitoring using ambient light and auto-regressive models“. In: *Physiological Measurement* 35.5 (2014), S. 807–831.
- [Tis19] S. M. Tisha. „LSU Digital Commons Thermal-Kinect Fusion Scanning System for Bodyshape Inpainting and Estimation under Clothing“. In: November (2019).
- [Tsu+17] L. Tsung-Yi, G. Priya, G. Ross, H. Kaiming und D. Piotr. „Focal Loss for Dense Object Detection“. In: (7. Aug. 2017). <https://arxiv.org/abs/1708.02002v2> (besucht am 07.04.2021).
- [Tup12] J. P. Tupin. „(12) Patent Application Publication (10) Pub. No.: US 2012/0059268 A1“. In: 1.19 (2012).
- [Two16] P. Tworzyclo. „Monitoring Breathing Using a Doppler Radar“. In: (2016).
- [VA71] S. Viitanen und K. A. „Skin circulation changes in the newborn during the first minutes of life as measured by thermography and skin thermometer“. In: *Ann Clin Res* (1971), S. 153–158.
- [Vad+22] M. Vadidar, A. Kariminezhad, C. Mayr, L. Kloeker und L. Eckstein. „Robust Environment Perception for Automated Driving: A Unified Learning Pipeline for Visual-Infrared Object Detection“. In: *IEEE Intelligent Vehicles Symposium, Proceedings 2022-June* (2022), S. 367–374.
- [Van+12] J. Van Baar, P. Beardsley, M. Pollefeys und M. Gross. „Sensor fusion for depth estimation, including TOF and thermal sensors“. In: *Proceedings - 2nd Joint 3DIM/3DPVT Conference: 3D Imaging, Modeling, Processing, Visualization and Transmission, 3DIMPVT 2012* (2012), S. 472–478.
- [vci] vcipl-okstate. „IRIS Imaging, Robotics and Intelligent Systems“. Zugriff am 4. Januar 2019. <http://vcipl-okstate.org/pbvs/bench/Data/02/download.html>.
- [Vin+13] G. Vinci, S. Lindner, F. Barbon, S. Mann u. a. „Six-Port Radar Sensor for Remote Respiration Rate and Heartbeat Vital-Sign Monitoring“. In: *IEEE Transactions on Microwave Theory and Techniques* 61.5 (Mai 2013), S. 2093–2100.
- [Vin+15] G. Vinci, T. Lenhard, C. Will und A. Koelpin. „Microwave interferometer radar-based vital sign detection for driver monitoring systems“. In: *2015 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility, ICMIM 2015* May 2019 (2015).

- [Vos+22] F. Voss, S. Lyra, D. Blase, S. Leonhardt und M. Lüken. „A Setup for Camera-Based Detection of Simulated Pathological States Using a Neonatal Phantom“. In: *Sensors* 22.3 (2022).
- [Vos+23] F. Voss, N. Brechmann, S. Lyra, J. Rixen, S. Leonhardt und C. Hoog Antink. „Multi-modal body part segmentation of infants using deep learning“. In: *Biomedical engineering online* 22.1 (2023), S. 28. <https://doi.org/10.1186/s12938-023-01092-0>.
- [Vuk+19] T. Vuković, R. Petrović, M. Pavlović und S. Stanković. „Thermal Image Degradation Influence on R-CNN Face Detection Performance“. In: *2019 27th Telecommunications Forum (TELFOR)*. 2019, S. 1–4.
- [Wan+15] S. Wang, A. Pohl, T. Jaeschke, M. Czaplik u. a. „A novel ultra-wideband 80 GHz FMCW radar system for contactless monitoring of vital signs“. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2015-November* (2015), S. 4978–4981.
- [Wan+22a] H. Wang, W. An, L. Li, C. Li und D. Zhou. „Infrared and visible image fusion based on multi-channel convolutional neural network“. In: *IET Image Processing* 16.6 (2022), S. 1575–1584.
- [Wan+22b] Z. Wang, F. Wang, D. Wu und G. Gao. „Infrared and Visible Image Fusion Method Using Saliency Detection and Convolutional Neural Network“. In: *Sensors* 22.14 (2022). <https://www.mdpi.com/1424-8220/22/14/5430>.
- [WHO] WHO. „Preterm birth“. Zugriff: 13.12.2018. <https://www.who.int/news-room/fact-sheets/detail/preterm-birth>.
- [wik] wikipedia. „Nervenzelle“. <https://de.wikipedia.org/wiki/Nerven-zelle>. Zugriff am 12. Juli 2024.
- [Wil+17a] C. Will, K. Shi, S. Schellenberger, T. Steigleder u. a. „Local Pulse Wave Detection Using Continuous Wave Radar Systems“. In: 1.2 (2017), S. 81–89.
- [Wil+17b] C. Will, K. Shi, R. Weigel und A. Koelpin. „Advanced template matching algorithm for instantaneous heartbeat detection using continuous wave radar systems“. In: *2017 First IEEE MTT-S International Microwave Bio Conference (IMBIOC)*. 2017, S. 1–4.
- [Wil+18] C. Will, K. Shi, S. Schellenberger, T. Steigleder u. a. „Radar-Based Heart Sound Detection“. In: *Scientific Reports* 8.1 (2018).
- [Wol] C. Wolff. „radartutorial.eu“. Zugriff am 18. November 2020. <https://www.radartutorial.eu>.
- [Wol87] D. Wolke. „Environmental neonatology“. In: *Arch Dis Child* (1987).
- [WP17] U. Wijenayake und S. Y. Park. „Real-time external respiratory motion measuring technique using an RGB-D camera and principal component analysis“. In: *Sensors (Switzerland)* 17.8 (2017).

- [Wu+12] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand und W. Freeman. „Eulerian video magnification for revealing subtle changes in the world“. In: *ACM Transactions on Graphics* 31.4 (2012), S. 1–8. <http://dl.acm.org/citation.cfm?doid=2185520.2335416>.
- [Xu+14] S. Xu, L. Sun, G. K. Rohde, A. K. Abbas u. a. „Neonatal non-contact respiratory monitoring based on real-time infrared thermography“. In: *BioMedical Engineering OnLine* 5.4 (2014), S. 1124–35. <http://www.biomedical-engineering-online.com/content/10/1/93%5C%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3985994%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [Yan+07] Q. Yang, R. Yang, J. Davis und D. Nistér. „Spatial-depth super resolution for range images“. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2007).
- [Yan+09] Yan, C. Li, X. Yu, M. D. Weiss und J. Lin. „Verification of a non-contact vital sign monitoring system using an infant simulator“. In: *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009 August 2016* (2009), S. 4836–4839.
- [Yan+15] S. Yang, P. Luo, C. C. Loy und X. Tang. „WIDER FACE: A Face Detection Benchmark“. 2015. <https://arxiv.org/abs/1511.06523>.
- [YIS19] D. Yudin, A. Ivanov und M. Shchendrygin. „Detection of a human head on a low-quality image and its software implementation“. en. In: *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLII-2/W12 (Mai 2019), S. 237–241.
- [YSL18] M. D. Yang, T. C. Su und H. Y. Lin. „Fusion of infrared thermal image and visible image for 3D thermal model reconstruction using smartphone sensors“. In: *Sensors (Switzerland)* 18.7 (2018).
- [Yu10] S. Z. Yu. „Hidden semi-Markov models“. In: *Artificial Intelligence* 174.2 (2010), S. 215–243. <http://dx.doi.org/10.1016/j.artint.2009.11.011>.
- [Zha+16] K. Zhang, Z. Zhang, Z. Li und Y. Qiao. „Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks“. In: *IEEE Signal Processing Letters* 23.10 (Okt. 2016), S. 1499–1503. <https://doi.org/10.1109%5C%2Fslsp.2016.2603342>.
- [Zha+19] X. Zhang, P. Ye, S. Peng, J. Liu, K. Gong und G. Xiao. „SiamFT: An RGB-Infrared Fusion Tracking Method via Fully Convolutional Siamese Networks“. In: *IEEE Access* 7 (2019), S. 122122–122133.
- [Zha+20] H. Zhang, L. Zhang, L. Zhuo und J. Zhang. „Object tracking in RGB-T videos using modal-aware attention network and competitive learning“. In: *Sensors (Switzerland)* 20.2 (2020).

Abbildungsverzeichnis

1.1	Frühgeborenes auf der Intensivstation [mam]	1
2.1	MRT-Bild eines zwei Tage alten Neugeborenen. Das (warme) braune Fettgewebe ist mit den Pfeilen markiert. [Hu+14]	4
2.2	Aufbau eines biologischen Neurons mit Axon. Auf die Dendriten werden Erregungen von anderen Neuronen übertragen und summiert. Abhängig vom Ergebnis der Summation am Axonhügel wird entschieden, ob ein Schwellenpotential überschritten und ein Aktionspotential gebildet wird [wik]. [Jar] [HJ11]	8
2.3	Aufbau eines Künstlichen Neuronalen Netzes mit Ein- und Ausgabeschicht und n versteckten Schichten. Es besteht eine Verbindung jedes Neurons der Vorgängerschicht mit allen Neuronen der Folgeschicht. [Hen23]	9
2.4	Faltung mit einem 3 x 3 Filter (Schrittweite eins). In der Grafik ist die untere Schicht dargestellt, darüber die nächste Schicht n+1. Das rezeptive Feld ist bunt markiert. Mit Padding ist das Verhalten der Randbereiche gemeint. Meist wird Zero-Padding verwendet. [Hen23]	10
2.5	Faltung einer 7 x 7 Feature Map M mit einem 3 x 3 Filter und Schrittweite eins. Die Feature Map M' zeigt das Ergebnis der Faltung. [Hen23]	10
2.6	Durch das weight sharing (siehe oben) müssen weniger Parameter trainiert werden. Unten: alle Neuronen sind miteinander über die Gewichte verbunden. [GBC16]	11
2.7	Beispiel für Max-Pooling angewandt auf einen 10 x 10 Layer (li.) mit Aufteilung in 2 x 2 große Bereiche und die Schicht nach dem Pooling (re.) [Hen23] . . .	12
2.8	Verschiedene Residual Blöcke. Links der original Residual Block von He. et al. [He+16a] und rechts das full pre-activation identity mapping, welches sich durchgesetzt hat [He+16b]. BN steht hierbei für Batch Normalization.	13
2.9	Struktur des RetinaNet [Lin+17a]. Der FPN (b) Backbone setzt auf einer ResNet Architektur (a) auf.	14
2.10	In der Grafik ist ein Residual Block für das ResNet-50 Backbone zu erkennen. In diesem Fall geht die Shortcut Verbindung über drei Layer mit zwei 1x1 Convolutions zur Reduktion und Erhöhung der Dimensionen und einer 3x3 Convolution. [Lin+17a].	14
2.11	Aufbau von YOLOv3 [Kat]	15
3.1	Grundprinzip des reflektorischen, optischen Pulsmessens (li.) und Absorptionsspektrum (re.) [HR14]	17
3.2	Impedanz-Pneumographie [Gup11]	17
4.1	Konzept des Gesamtsystems. Heterogene Sensorik soll in einer Sensorbox im Inkubator integriert werden.	40

4.2 Säuglingsinkubator Isolette 8000 von der Firma Dräger. Bildquelle: Fa. Dräger 41

4.3 Simulation des SAR-Werts eines Erwachsenen bei einem Abstand von 20 cm. Der Wert wird über 10 g gemittelt (re.). (Simulationen von Fabian Michler, LTE FAU, erstellt.) 47

4.4 Simulation des SAR-Werts eines Neugeborenen bei einem Abstand von 20 cm und der Detailansicht der Hautschichten (re.). (Simulationen von Fabian Michler, LTE FAU, erstellt.) 48

4.5 Elektro-mechanischer Ansatz für den Baby-Thorax-Simulator [Gle+] 49

4.6 Simulation der Atmung und des Herzschlages durch Überlagerung zweier Sinus-signale [Gle+21] 50

5.1 Kamerasensoren: Thermokamera Optris PI640 (li.) und JAI GO-2400C-USB Kamera (re.) Quelle: Fa. Optris und Fa. JAI 54

5.2 Radarsensor von Fa. InnoSenT (li.) und Time of Flight Kamera CamBoard pico flexx von Firma PMD (re.) 55

5.3 Ausgangskanäle der einzelnen Ausgänge 56

5.4 I-Q-Diagramm 57

5.5 Distanzverlauf über die Zeit, berechnet aus den I-Q-Signalen. 58

5.6 Kommunikationsstruktur zwischen Temperatur- bzw. Feuchtigkeitssensor und Mikrocontroller und ROS [Kau22] 59

5.7 Gesamter Sensoraufbau (li. Ansicht von unten) mit Simulator-Puppe Paul (Fa. SimCharacters) als Messobjekt (re.). Temperatursensor (vorne blau) mit Isolierung. Die flexx2 ToF-Kamera ist in der Sensorbox ganz links angebracht. 60

5.8 Sensoren für die Referenzmessungen: EKG-Klebeelektroden (re.), Pulsoxymeter (mi.) und rektale Temperatursonde (li.) 62

5.9 Finaler Testaufbau mit der ToF-Kamera und dem Radarsensor. Der Baby-Thorax-Simulator ist unten im Bild zu sehen [Gle+21]. Das Laser-Mikrometer dient als Goldstandard. 63

5.10 Intrinsische Kalibrierung mit dem zugeschnittenen RGB-Kamerabild unter der Verwendung von 196 Bildern (rechts). Das Kalibrieremuster ist ein 5 x 9 Schachbrettmuster mit Quadraten der Seitenlänge 14 mm. Unkalibriertes Bild (links) [Gle+23]. 64

5.11 Empirische Anpassung der intrinsischen Kalibrierung der RGB-Kamera. (a) Originalbild; (b) Änderung des ersten Werts der Projektions-Matrix von 337,75012 zu 400; (c) Änderung des ersten Werts der Projektions-Matrix von 337,75012 zu 410; (d) Änderung des zweiten Werts der Projektions-Matrix von 567,37183 zu 500. [Gle+23] 65

5.12 Empirische Anpassung der intrinsischen Kalibrierung der RGB-Kamera. (e) Änderung des dritten Werts der Projektions-Matrix von 456,18869 zu 470; (f) Projektions-Matrix mit erstem Wert 560, zweiten Wert 500 und drittem Wert 470; (g) Änderung der Verzerrungs-Matrix zu [-0,22825, 0,03338, 0,00101, -0,00106, 0] [Gle+23] 66

5.13 Finale Kalibrierung der Thermokamera mit 144 Bildern. Kalibriertes Bild (re.) und unkalibriertes Bild (li.). Der Kisseneffekt ist komplett entfernt. [Gle+23] 67

5.14 Blob detection im Thermo (li.), ToF (mi.) und RGB-Kamerabild (re.) Bei dem RGB-Bild wurde zur Erhöhung des Kontrastes ein blaues Papier unter das Kreismuster gelegt. 68

5.15 Extrinsische Kalibrierung der RGB- mit der ToF-Kamera; 26 Punktkorrespondenzen wurden verwendet (re.). Die türkisen Punkte sind die projizierten ToF-Punkte (li.). 69

5.16 Extrinsische Kalibrierung der Thermo- mit der ToF-Kamera; 15 Punktkorrespondenzen wurden verwendet (re.). Die türkisen Punkte sind die projizierten ToF-Punkte (li.). 69

5.17 Thermo-ToF-Fusion (li.), RGB-ToF-Fusion (mi.) und Thermo-RGB-Fusion Kamerabild (re.). 69

5.18 Ablauf des Algorithmus zur Atemdetektion mit einer Thermokamera. 72

5.19 Ablauf des Algorithmus zur Atemdetektion mit einer Structured-Light-Kamera [GNK20]. 72

5.20 Segmentierung der Punktwolke. Die bunte Ebene zeigt den Brust- und Bauchbereich 72

5.21 Ablauf der ToF-Atemdetektion. Im Unterschied zur Datenverarbeitung bei der Structured-Light-Kamera wird bei der Ebenendetektion anstelle vom RANSAC der LMEDS Algorithmus angewandt. Hinzu kommt das Statistical-Outlier-Removal-Filter, das Medianfilter und die PCA mit Bestimmung des Schwerpunkts über die Zeit. [Gle+21] 74

5.22 Ausschnitt des ungefilterten (blau) und gefilterten (magenta) Atemsignals bei 20 BPM im *deep* Modus. Das angewandte Savitzky–Golay Filter hat eine Fenstergröße 3 und der Grad des Polynoms liegt bei 2. Durch das Filter wird eine Glättung des Signals bewirkt, sodass die Peaks eindeutig herausgestellt werden. [Gle+21] 75

5.23 Ablauf der Atemdetektion mit der ToF-Kamera und integriertem Neuronalem Netz. 75

5.24 Signalverarbeitung für die Atemraten-Detektion mit dem Radar-Sensor. 76

5.25 Ablauf der Herzrattendetektion mit der RGB-Kamera und integriertem Neuronalem Netz. 78

5.26 Flowchart des Peak2peak-Algorithmus für die Herzraten-Detektion mit dem Radar-Sensor. 80

5.27 Flowchart des FFT-per-Band-Algorithmus für die Herzraten-Detektion mit dem Radar-Sensor. 80

5.28 Flowchart des one-FFT-for-all-Algorithmus für die Herzraten-Detektion mit dem Radar-Sensor. 80

5.29 Flowchart Temperaturdetektion über die Differenz der peripheren und der Körperkerntemperatur. 82

6.1 Aufbau mit Structured-Light-Kamera. 84

6.2 Abschnitt des respiratorischen Signals des SimBabies in Rückenlage in *deep* mode. Die eingestellte Atemrate liegt bei 50 BPM. Daraus resultiert, dass die Fenstergröße des Filters 3 ist und es insgesamt 11 Filterfenster in diesem Abschnitt gibt (vgl. Abschnitt 5.3). Jedes Filterfenster endet an dem markierten magentafarbenen Punkt. Gefiltertes Signal (magenta), ungefiltertes Signal (blau). [GNK20] 85

6.3 Box-Whisker-Plot im *deep* Modus während das SimBaby auf dem Rücken liegt (oben) und in Bauchlage (unten). Die maximale Differenz zur Referenz im Normbereich liegt bei 14 BPM während das SimBaby auf dem Rücken liegt und bei 6 BPM in Bauchlage. [GNK20] 86

6.4 Box-Whisker-Plot im *shallow* Modus während das SimBaby auf dem Rücken liegt (oben) und in Bauchlage (unten). Die maximale Differenz zur Referenz im Normbereich liegt bei 9 BPM in Rückenlage und 20 BPM in Bauchlage. [GNK20] 87

6.5 Box-Plot, bei dem die BPM-Werte des ToF- und des Radar-Algorithmus für Atemraten zwischen 5 BPM und 40 BPM im *deep* Modus verglichen werden [Gle+21]. 89

6.6 Box-Plot, bei dem die BPM-Werte des ToF- und des Radar-Algorithmus für Atemraten zwischen 45 BPM und 80 BPM im *deep* Modus verglichen werden [Gle+21]. 90

6.7 Box-Plot, bei dem die BPM-Werte des ToF- und des Radar-Algorithmus für Atemraten zwischen 5 BPM und 40 BPM im *normal* Modus verglichen werden [Gle+21]. 91

6.8 Box-Plot, bei dem die BPM-Werte des ToF- und des Radar-Algorithmus für Atemraten zwischen 45 BPM und 80 BPM im *normal* Modus verglichen werden [Gle+21]. 92

6.9 Ungefilterter Abschnitt des Atemsignals bei 10 BPM im *deep* Modus, gemessen mit der ToF-Kamera, dem Radar-Sensor und dem Laser-Mikrometer. Im Kontext der ToF-Kamera bedeutet "ungefiltert", dass der Savitzky-Golay-Filter noch nicht angewandt wurde. Zu Beginn ist das Synchronisationssignal zu erkennen. Das ToF-Kamerasignal ist stark verrauscht und hat große Amplitudensprünge. Das Radar- und Laser-Mikrometer-Signal liegen nahe aneinander. [Gle+21] . 94

6.10 Noch ungefilterter Abschnitt des Atemsignals bei 45 BPM im *deep* Modus, gemessen mit der ToF-Kamera, dem Radar-Sensor und dem Laser-Mikrometer. Zu Beginn ist das Synchronisationssignal zu erkennen. Zu Beginn des ToF-Signals ist das Rauschen klar sichtbar. Später ist der Hub des Simulator-Thorax höher als das Rauschen. Die Amplitude ähnelt der des Laser-Mikrometers. Die Amplitude des Radar-Sensors ist leicht verschoben und ungleichmäßig. Dies kann an den Effekten des Ellipsen-Fittings liegen. [Gle+21] 95

6.11 Variierende Radaramplitude, abhängig davon, ob Ellipse-Fitting (gelb), der Mittelwert (rot), gleitender Durchschnitt (lila) oder ein gefensterter Durchschnitt (grün) verwendet wird. In diesem Beispiel wird eine Atemrate von 40 BPM im *deep* Modus verwendet. [Gle+21] 96

6.12 Box-Plot zum Vergleich der Atemraten von der Time-of-Flight-Kamera (blau) und der eingestellten Atemrate (normal) des Simulators nach Frequenzen ausgewertet. 97

6.13 Box-Plot zum Vergleich der Atemraten von der Time-of-Flight-Kamera (blau) und der eingestellten Atemrate RDS des Simulators nach Frequenzen ausgewertet. 98

6.14 Es wird anstelle des Thorax die Windel des Simulators detektiert, da die Windel den geringsten Abstand zur Kamera hat. 98

6.15 Trainingsverlauf Phase 1 RGB-Daten für RetinaNet, AP per Epoche 102

6.16 Trainingsverlauf Phase 1 RGB-Daten für YOLOv3, AP per Epoche 102

6.17 Trainingsverlauf Phase 2 RGB-Daten für RetinaNet, AP per Epoche 103

6.18 Trainingsverlauf Phase 2 RGB-Daten für YOLOv3, AP per Epoche 103

6.19 Trainingsverlauf Phase 1 Thermo-Daten für RetinaNet, AP per Epoche 104

6.20 Trainingsverlauf Phase 1 Thermo-Daten für YOLOv3, AP per Epoche 104

6.21 Trainingsverlauf Phase 2 Thermo-Daten für RetinaNet, AP per Epoche 105

6.22 Trainingsverlauf Phase 2 Thermo-Daten für YOLOv3, AP per Epoche 105

6.23 Trainingsverlauf Phase 1 Fusions-Daten für RetinaNet, AP per Epoche 106

6.24 Trainingsverlauf Phase 1 Fusions-Daten für YOLOv3, AP per Epoche 106

6.25 Trainingsverlauf Phase 2 Fusions-Daten für RetinaNet, AP per Epoche 107

6.26 Trainingsverlauf Phase 2 Fusions-Daten für YOLOv3, AP per Epoche 107

6.27 Visuelle Auswertung von Beispiels Fusions-Bildern mit RetinaNet (links) und YOLOv3 (rechts). Mit RetinaNet werden confidence scores von 0.995 für den Kopf und 0.973 für den Torso erhalten. Für YOLOv3 ergeben sich confidence scores von 0.95 für den Kopf und 0.94 für die Intervention. [Gle+23] 112

6.28 Box-Plot der prozentualen Abweichungen der Atemdetektion mit der ToF-Kamera sortiert nach aufsteigendem Median. 113

6.29 Box-Plot der Atemdetektion mit der Thermokamera von Proband 9 aufgeteilt nach Frequenzbereichen. (Rot: Frequenzbereich der manuellen Referenz, blau: Ergebnisse der Thermokamera) 114

6.30 Box-Plot der Atemdetektion mit der Thermokamera von Proband 15 aufgeteilt nach Frequenzbereichen. (Rot: Frequenzbereich der manuellen Referenz, blau: Ergebnisse der Thermokamera) 115

6.31 Box-Plot der prozentualen Abweichungen der Herzrattendetektion mit der RGB-Kamera. Es wurde mit aufgezeichneten Datensätzen verschiedener Situationen getestet. 116

6.32 Verlauf der Herzrate mit RGB-Kamera (rot) vs. Referenz vom EKG (blau) von Proband 17, 15:31. 117

6.33 Beispielhafte Box-Plots der Temperaturdetektion unter verschiedenen Bedingungen. 120

7.1 Neonat mit breiter Binde im RGB-(li.) und Thermokamerabild (re.). 130