Robuste und flexible Methoden zur quantitativen Analyse multidimensionaler Bilddaten in Radiologie und Materialwissenschaft

Zur Erlangung des akademischen Grades eines

DOKTORS DER INGENIEURWISSENSCHAFTEN (Dr.-Ing.)

von der KIT-Fakultät für Maschinenbau des
Karlsruher Instituts für Technologie (KIT)
angenommene

DISSERTATION

von

Julian Grolig, M.Sc.

Tag der mündlichen Prüfung:	24. Juni 2026
Hauptreferentin:	Prof. Dr. Britta Nestler
Korreferentin:	PD Dr. med. Lena Wucherpfennig



Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz (CC BY-SA 4.0): https://creativecommons.org/licenses/by-sa/4.0/deed.de

Kurzfassung

Viele drängende gesellschaftliche Herausforderungen, wie verbesserte medizinische Diagnostik und die Entwicklung neuer Werkstoffe, erscheinen grundlegend verschieden, lassen sich jedoch in der Ära der heutigen beispiellosen Rechenleistung und Künstlichen Intelligenz (KI) durch gemeinsame und effiziente Lösungsstrategien angehen. Beispielsweise kommen zur Diagnose von Lungenerkrankungen radiologische Bildgebungstechniken zum Einsatz, die eine Analyse von biologischen Geweben ermöglichen, die flexibel, durchblutet und porös sind. Ebenso profitiert die Entwicklung und Optimierung poröser Funktionsmaterialien stark von biologisch inspirierten Algorithmen und radiologischen Bildgebungstechniken. Früher benötigten disziplinspezifische Probleme jeweils eigene Analysemethoden, was zu redundanter Entwicklungsarbeit und langen Durchlaufzeiten führte. Heute ermöglichen die Verfügbarkeit von Daten, Fortschritte im strukturierten Forschungsdatenmanagement (FDM) und datengesteuertes Modellieren erstmals interdisziplinäre, effiziente Lösungsstrategien für diese gemeinsamen Probleme.

Ziel dieser Arbeit ist, eine wiederverwendbare informationsgetriebene Bildanalysestrategie zu etablieren, die die Entwicklungszeit in Radiologie und Materialwissenschaften durch flexible und robuste Methoden sowie ein strukturiertes FDM verkürzt. Der hier vorgestellte informationsgetriebene Bildanalysezyklus verknüpft Daten, Methoden und Dokumentation in einem iterativen Workflow und automatisiert die Protokollierung, wodurch die Weiterverwendung und Weiterentwicklung von Analyseergebnissen nachhaltig gesichert werden.

Die Segmentierung spiegelt einen zentralen Aspekt der Bildanalyse wider. In dieser Arbeit werden für drei verschiedene Datensätze aus Radiologie und Materialwissenschaften flexible Segmentierungsverfahren entwickelt, die keinerlei annotierte Trainingsdaten benötigen. Die ersten beiden Methoden kombinieren verborgene zufällige Markov-Felder mit einem populationsbasierten Optimierer bzw. einem U-Net-Modell. Diese lassen sich an unterschiedliche Bildmodalitäten anpassen und liefern für alle drei Testdatensätze überzeugende Ergebnisse. Die dritte Methode demonstriert, wie ein maßgeschneiderter Algorithmus die Segmentierungsqualität weiter steigern kann, ohne an Flexibilität zu verlieren.

Für eine robuste und automatische Quantifizierung der Perfusion in kontrastmittelverstärkten MRT-Bildern (DCE-MRT) der Lunge wird eine bestehende Pipeline weiterentwickelt. Die bestehende Pipeline wird um ein maschinelles Lernmodell zur automatischen Atemerkennung, die quantitative Bestimmung der Perfusionsverzögerung und einen Qualitäts-Score für DCE-MRT-Bilder erweitert. Diese Ergänzungen ermöglichen eine vollautomatische, robuste Quantifizierung der Lungen-Perfusion. Die weiterentwickelte Pipeline und ihre Bestandteile werden anhand von zwei Studien validiert. Die enthaltene Segmentierungsmethode erzeugt hierbei gute und robuste Segmentierungsergebnisse. In einer Studie mit PatientInnen mit Mukoviszidose kann das Potenzial der neuen quantitativen Perfusionsverzögerungsmetrik gezeigt werden. Die Validierung innerhalb einer multizentrischen Studie mit verschiedenen MRT-Scannern zeigt starke Korrelationen der berechneten quantitativen Perfusionswerte mit klinischen Parametern sowie quantitativen CT-Werten und einen leichten Einfluss des MRT-Scanners.

Für eine strukturierte Dokumentation der Forschungsdaten innerhalb von radiologischen Bildverarbeitungsprojekten wird die virtuelle Forschungsumgebung Kadi4Mat verwendet. Hierzu wird ein Vorgehen entwickelt, das sowohl Daten als auch Methoden automatisch abspeichert und diese miteinander verknüpft.

Durch die Kombination flexibler und robuster Bildanalyse-Methoden mit einer informationsgetriebenen Infrastruktur wird der Entwicklungsaufwand durch das Konzept erheblich reduziert und kann ohne größere Anpassungen in zukünftigen radiologischen und materialwissenschaftlichen Projekten eingesetzt werden.

Abstract

Many pressing societal needs, such as improved medical diagnostics and materials development, may appear fundamentally different but offer themselves to joint, and efficient solution strategies in the current era of unprecedented computational capabilities and Artificial Intelligence (AI). For example, radiological imaging techniques are used in the diagnosis of lung diseases, enabling the analysis of biological tissues that are flexible, perfused, and porous. Similarly, the development and optimization of porous functional materials greatly benefit from bio-inspired algorithms and radiological imaging techniques. Previously, discipline-specific problems required dedicated analysis methods, resulting in redundant development effort and long turnaround times. Today, the availability of data, advances in structured Research Data Management (RDM), and data-driven modeling enable interdisciplinary, efficient solution strategies to these common problems for the first time.

The aim of this work is to establish a reusable, information-driven image analysis strategy that shortens development time in radiology and materials science through flexible, robust methods and structured RDM. The developed information-driven image analysis cycle links data, methods, and documentation in an iterative workflow and automates logging, thereby ensuring sustainable reuse and the further development of the analysis results.

Image segmentation represents a central aspect of image analysis. In this thesis, flexible segmentation methods are developed for three distinct datasets from radiology and materials science. These methods require no annotated training data. The first two methods combine Hidden Markov Random Fields with a population-based optimizer and a U-Net model, respectively. These methods can be adapted

to various imaging modalities and produce convincing results across all three test datasets. The third method demonstrates how a tailored algorithm can further enhance segmentation quality without sacrificing flexibility.

For robust and automatic quantification of perfusion in contrast-enhanced MRI (DCE-MRI) of the lung, an existing pipeline is extended. The pipeline is expanded to include a machine learning model for automatic breathing detection, the quantitative determination of perfusion delay, and a quality score for DCE-MRI images. These additions enable fully automatic and robust lung perfusion quantification. The enhanced pipeline and its components are validated using two studies. The included segmentation method yields accurate and robust segmentation results. In a study of cystic fibrosis patients, the potential of the new quantitative perfusion delay metric is demonstrated. Validation within a multicenter trial using different MRI scanners shows strong correlations between computed quantitative perfusion values and clinical parameters and quantitative CT measurements, with only a slight influence from the MRI scanner.

For structured documentation of research data within radiological image analysis projects, the virtual research environment Kadi4Mat is employed. A workflow is developed that automatically saves both data and methods and links them together.

By combining flexible and robust image analysis methods with a information-driven infrastructure, the development effort is significantly reduced. This approach can be applied in future radiology and materials science projects without major adaptations.

Danksagung

Diese Dissertation entstand im Zuge einer Kooperation zwischen dem Institut für Angewandte Materialien – Mikrostruktur-Modellierung und Simulation (IAM-MMS) bzw. dem Institut für Nanotechnologie (INT) des Karlsruher Instituts für Technologie (KIT) und der Klinik für Diagnostische und Interventionelle Radiologie (DIR) des Universitätsklinikums Heidelberg (UKHD). Während dieser Zeit konnte ich als wissenschaftlicher Mitarbeiter und Doktorand in beiden Arbeitsgruppen sehr viele wertvolle Erfahrungen sammeln, für die ich sehr dankbar bin.

Zunächst möchte ich mich bei Frau Prof. Dr. Britta Nestler für die Betreuung meiner Promotion sowie die Möglichkeit, an ihrem Institut zu forschen, bedanken. Ebenso bedanke ich mich bei Frau PD Dr. Lena Wucherpfennig für die Bereitschaft zur Übernahme des Korreferats. Mein Dank richtet sich ebenfalls an Herrn Prof. Dr. Hans-Ulrich Kauczor für die Möglichkeit in der DIR an spannenden radiologischen Forschungsthemen zu forschen.

In der Arbeitsgruppe IAM-MMS bzw. INT gilt ein besonderer Dank meinen Gruppenleitern Herrn Dr. Arnd Koeppe, Herrn Dr. Michael Selzer und Herrn Dr. Lars Griem für viele interessante Diskussionen, Feedback und Anregungen zur Ausrichtung der Forschung und das Korrekturlesen dieser Arbeit. Außerdem möchte ich mich bei meinen ehemaligen und aktuellen KollegInnen des CIDS- und Kadi-Team bedanken, für wertvollen Ideenaustausch und die familiäre Atmosphäre während und außerhalb der Arbeit. Dieses besondere Arbeitsumfeld hat die tägliche Arbeit am Institut sehr bereichert. Mein Dank gilt auch allen weiteren

MitarbeiterInnen des Instituts für die gemeinsame Arbeit sowie dem Sekretariatsteam, insbesondere Frau Inken Heise, für die Hilfe rund um organisatorische Themen.

Meinen ehemaligen und aktuellen KollegInnen in der DIR danke ich für einen regelmäßigen fruchtbaren und interessanten Austausch. Insbesondere auch die gemeinsamen Konferenzbesuche wie das jährliche DZL Annual Meeting waren immer ein Highlight und zeigten den starken Zusammenhalt und den großen Spaß innerhalb des Teams. Ein großer Dank geht auch an meinen ehemaligen Betreuer Prof. Dr. Mark O. Wielpütz, der sich immer viel Zeit nahm, um mir als ursprünglich fachfremde Person medizinische Lungengrundlagen zu erklären und wichtige Impulse zur Forschungsausrichtung lieferte. Ein ganz besonderer Dank gilt meinem Betreuer Herrn PD Dr. Simon Triphan, der mir die Grundlagen des Lungen-MRT näher gebracht hat und mir regelmäßig bei der Forschungsausrichtung und bei Implementierungsfragen mit Rat und Tat zur Seite stand und ebenso durch das Korrekturlesen dieser Arbeit wichtiges Feedback lieferte. Ebenfalls möchte ich mich beim Sekretariatsteam der DIR für die Unterstützung bei organisatorischen Themen danken sowie bei Frau Dr. Oyunaa v. Stackelberg, dass sie mich damals zu Beginn für dieses interdisziplinäre Projekt begeistert hat.

Der größte Dank geht jedoch an meine Familie. Von ganzem Herzen danke ich meiner Frau Melanie für ihre uneingeschränkte Unterstützung, Geduld und Verständnis während der gesamten Promotionszeit. Trotz ihrer weit fortgeschrittenen Schwangerschaft konnte ich mich während des Endspurts der Promotion stets auf sie verlassen und bekam sogar noch Feedback von ihr durch das Probelesen dieser Arbeit. Ein tiefer Dank geht auch an meine Eltern und meinen Bruder, dass sie mich in allen Entscheidungen immer unterstützen. Sie haben einen großen Anteil daran, dass ich zu dem Mensch geworden bin, der ich heute bin.

Karlsruhe, im April 2026 *Julian Grolig*

Inhaltsverzeichnis

Kurzfassung i

Abstract iii

Danksagung v

1 Einleitung 1
1.1 Einordnung und Motivation 1
1.2 Zielsetzung und Beitrag der Arbeit 5
1.3 Gliederung der Arbeit 6

2 Theoretische Grundlagen 9
2.1 Lunge und Atemwege 9
2.1.1 Aufbau und Funktion 9
2.1.2 Ausgewählte Lungenerkrankungen 12
2.2 Bildgebung 15
2.3 Computertomographie 16
2.3.1 Quantitatives CT 17
2.4 Magnetresonanztomographie 18
2.4.1 Grundlagen 19
2.4.2 Besonderheiten des Lungen-MRT 27
2.4.3 DCE-MRT der Lunge 29
2.4.4 Auswertung des Perfusions-MRT 31
2.5 Maschinelles Lernen 32
2.5.1 Neuronale Netze 34
2.5.2 Populationsbasierte Metaheuristiken 38

2.5.3 Markov Random Fields 47

3 Konzept der informationsgetriebenen Bildanalyse 51
3.1 Bildgebung 53
3.2 Bildverarbeitung 53
3.2.1 Segmentierung 54
3.3 Quantifizierung 57
3.4 Verbesserungsmaßnahmen 57
3.5 Forschungsdatenplattform und Daten-Workflows 58
3.6 Flexible und robuste Methoden 59
3.6.1 Flexible Methoden 59
3.6.2 Robuste Methoden 60

4 Bildverarbeitung: Flexible Methoden zur Segmentierung von multidimensionalen Bilddaten 63
4.1 Stand der Forschung 64
4.2 Datensätze 67
4.3 EHOHMRF 68
4.3.1 Aufbau des EHOHMRF-Algorithmus 68
4.3.2 EHO mit und ohne Gradientenupdate 71
4.3.3 Untersuchung verschiedener Verteilungsfunktionen 72
4.3.4 Einfluss des Nachbarschaftsterms 76
4.3.5 Gewichtung des Nachbarschaftsterms 77
4.3.6 Anwendung auf *RealPUFoam*-Datensatz 78
4.3.7 Anwendung auf Lungen-MRT-Datensatz 81
4.3.8 Bewertung der EHOHMRF-Methode 84
4.4 HMRF-UNet 87
4.4.1 Neurartige Fuzzy HMRF-Lossfunktion für neuronale Netze 87
4.4.2 Aufbau des HMRF-UNet 90
4.4.3 Daten Vorverarbeitung 91
4.4.4 Optimierung & Training des HMRF-UNet 93
4.4.5 Einfluss des Nachbarschaftsterms 94
4.4.6 Verwendung des HMRF-UNet zum Vortrainieren 99

4.4.7 Anwendung auf *RealPUFoam*-Datensatz 103
4.4.8 Anwendung auf Lungen-MRT-Datensatz 107
4.4.9 Bewertung des HMRF-UNet 111
4.5 Algorithmische Lungensegmentierung 113
4.5.1 Segmentierungs-Pipeline 113
4.5.2 Optimierte Tracheasegmentierung 116
4.5.3 Anwendung auf Lungen-MRT-Datensatz 117
4.5.4 Bewertung der algorithmischen Lungensegmentierung . . 120

5 Quantifizierung: Robuste automatische Perfusionsquantifizierung 123
5.1 Stand der Forschung 124
5.2 Halbautomatische Perfusionsquantifizierung 131
5.2.1 Einlesen von DICOMs 133
5.2.2 Manuelle Atmungsentfernung 133
5.2.3 Lungensegmentierung 134
5.2.4 Detektion der Pulmonalarterie 135
5.2.5 Atmungskorrektur 136
5.2.6 Perfusionsquantifizierung 137
5.3 Alternative Registrierungsmethoden 138
5.4 Automatische Atmungserkennung 140
5.4.1 Daten-Vorverarbeitung 141
5.4.2 Auswahl einer passenden Modellarchitektur 142
5.4.3 Ausgewählte Modellarchitektur 144
5.4.4 Validierung des Modells 145
5.5 Charakterisierung des Kontrastmitteldurchlaufs 147
5.6 Automatisches Herausfiltern veratmeter Zeitpunkte 148
5.7 Quantifizierung der Perfusionsverzögerung 150
5.8 DCE-MRT Qualitätsscore 151
5.8.1 Messung der Qualität der Bolusinjektion 153
5.8.2 Bestimmung der Atmungshäufigkeit 154
5.9 EigenLung-Distanz Metrik 155
5.10 Validierung der algorithmischen Lungensegmentierung 158

5.11 Monozentrische Perfusionsquantifizierung in CF 160
5.11.1 Klinische Einordnung 160
5.11.2 Studienkohorte 161
5.11.3 Auswertung der MRT Messungen 162
5.11.4 Korrelation mit klinischen Parametern 164
5.11.5 Zusammenhang von Perfusionsdefekten und quantitativen Pefusionswerten 166
5.11.6 Vergleich von PatientInnen mit und ohne BAD 168
5.11.7 Einordnung und Zusammenfassung der Ergebnisse 172
5.12 Multizentrische Perfusionsquantifizierung in COPD 172
5.12.1 Studienkohorte 173
5.12.2 Validierung der algorithmischen Lungensegmentierung . . 175
5.12.3 Vergleich verschiedener QDP-Varianten 179
5.12.4 Vergleich der Registrierungsmethoden 182
5.12.5 Korrelation mit klinischen Parametern & QCT 185
5.12.6 Quantitative Perfusionswerte 188
5.12.7 Einfluss des MRT-Scanners auf die Quantifizierung 190
5.12.8 MRT-Qualitätsscore 195
5.12.9 Einordnung und Zusammenfassung der Ergebnisse 198

6 Daten-Workflows für die Bildverarbeitung 203
6.1 Stand der Forschung 204
6.2 FDM für die medizinische Bildverarbeitung 207
6.3 Aufbereitung und Exportieren der Bilddaten 209
6.4 FDM für ML-Projekte 211
6.5 FDM für die Dokumentation der COSYCONET-Ergebnisse . . . 215
6.6 Einordnung des entwickelten FDM-Konzeptes 219

7 Zusammenfassung und Ausblick 221
7.1 Zusammenfassung 221
7.2 Ausblick 225

Literaturverzeichnis 229

Abbildungsverzeichnis **261**

Tabellenverzeichnis **265**

Abkürzungs- und Symbolverzeichnis **269**

A Anhang **281**
A.1 Elastix Registrierungsparameter 282
A.2 ANTs Registrierungsparameter 285
A.3 Untersuchte Hyperparameter Atmungserkennung 285
A.4 COSYCONET Studiendetails 287
A.5 Validierung Segmentierung 288

B Veröffentlichungen **289**

1 Einleitung

1.1 Einordnung und Motivation

Laut Weltgesundheitsorganisation (WHO) waren im Jahr 2021 drei der fünf häufigsten Todesursachen auf Lungenerkrankungen zurückzuführen. Selbst wenn der Sonderfall der COVID-19 Pandemie ausgeklammert wird, befanden sich mit der chronisch obstruktiven Lungenerkrankung (COPD) und den unteren Atemwegsinfektionen zwei Lungenerkrankungen unter den vier häufigsten Todesursachen. [1] Durch die aktuellen gesellschaftlichen Probleme wie den Klimawandel und die verstärkte Luftverschmutzung könnte die Anzahl von Lungenerkrankungen in Zukunft sogar noch weiter ansteigen [2]. Aus diesem Grund ist weitere Forschung hinsichtlich der zugrundeliegenden Ursachen sowie der Diagnose- und Behandlungsmöglichkeiten essenziell.

Ein zentraler Aspekt zur Eindämmung des Klimawandels und der Klimaerwärmung ist die Reduktion von klimaschädlichen Gasen. Eine Möglichkeit hierfür ist die Entwicklung von leichten Elektrofahrzeugen. [3] Sowohl im Leichtbau, als auch in der Batterieentwicklung spielen poröse Materialien eine wichtige Rolle. Die Verkürzung der Entwicklungszeiten von neuen Materialien ist hierbei ein wichtiger Baustein, um die Treibhausgasemissionen und deren Auswirkungen auf das Klima schnellstmöglich zu reduzieren.

Auf den ersten Blick scheinen Lungenerkrankungen und Materialentwicklung zwei völlig unterschiedliche Themen zu sein. Bei näherer Betrachtung erkennt man jedoch, dass die Lunge im wesentlichen auch eine Art Verbundmaterial darstellt,

das aus vielen unterschiedlichen Bestandteilen, wie den Atemwegen, dem Lungenparenchym und den Blutgefäßen besteht. Ebenso finden in der Lunge und in Materialien ähnliche Prozesse wie beispielsweise die Diffusion oder die Alterung statt. Eine weitere Gemeinsamkeit der beiden Bereiche ist, dass zur Bestimmung von Eigenschaften eine Schädigung bzw. Zerstörung des Objektes, wie es bei einer Lungenbiopsie oder dem Zugversuch vorkommt, im Normalfall vermieden werden soll und daher häufig nichtinvasive bzw. zerstörungsfreie Messmethoden eingesetzt werden. Eine in beiden Bereichen verwendete zerstörungsfreie Messmethode ist die Bildgebung mit anschließender quantitativer Bildanalyse. Damit bedienen sich beide Bereiche aus einem gemeinsamen Methodenbaukasten, der Methoden der Bildgebung, Bildverarbeitung und Quantifizierung beinhaltet.

Um die genannten Probleme im Bereich der Lungenerkrankungen und der Materialentwicklung zu lösen, gilt es Methoden zu entwickeln, die sowohl bei der Erforschung von Lungenerkrankungen als auch neuer Materialien helfen. Eine Neuentwicklung von Methoden ist jedoch meist sehr zeitintensiv. Neben der Neuentwicklung von Methoden gibt es laut Ralyté [4] drei weitere Vorgehen zur Entwicklung wissenschaftlicher Prozesse. Dies sind die Methodenanpassung, die Funktionserweiterung und die selektive Auswahl von Teilfunktionen. Die Weiterverwendung ist ein essenzieller Bestandteil moderner Forschung, da sie es ermöglicht, bereits entwickeltes Prozesswissen wiederzuverwenden bzw. auf neue Forschungsfelder zu übertragen. Dadurch kann die Entwicklungszeit signifikant verkürzt werden.

Hierbei lässt sich grundsätzlich zwischen der Weiterverwendung einer Methode innerhalb der gleichen Disziplin (intradisziplinär) und der Weiterverwendung in anderen Disziplinen (interdisziplinär) unterscheiden. Der intradisziplinäre Methodenaustausch wird vor allem durch die Robustheit einer Methode gegenüber äußeren Einflüssen, wie beispielsweise einer Veränderung der Aufnahmebedingungen, ermöglicht. Robustheit beschreibt die Eigenschaft einer Methode gegenüber verschiedenen Eingangs- und Einflussgrößen invariant zu sein und dennoch vergleichbare Ergebnisse zu liefern. Im Gegensatz dazu erfordert der interdisziplinäre Austausch wissenschaftlicher Methoden insbesondere eine leichte Anpassbarkeit

und Flexibilität der Methode. Die Flexibilität ist definiert als die Eigenschaft, gegenüber dem Einsatzgebiet der Methode offen zu sein und mit einfachen Mitteln an neue Einsatzbereiche anpassbar zu sein. Die Verwendung von fremden Methoden ist häufig insbesondere dann nützlich, wenn am eigenen Standort keine oder nur wenig Expertise für ein Thema vorliegt, aber dennoch Daten ausgewertet werden sollen. In diesem Fall kann eine automatische Auswertung mithilfe einer bereits erprobten Methode die Analyse der gewünschten Daten ermöglichen. Jedoch sollte, insbesondere im Fall fehlender eigener Expertise in dem Bereich, besonderer Wert auf die genaue Validierung und Erprobung der genutzten Methode gelegt werden.

Eine alternative Möglichkeit, die fehlende Expertise zu kompensieren, ist die Bildung eines interdisziplinären Teams, indem Experten auf dem gewünschten Gebiet hinzugezogen werden. Interdisziplinäre Forschung wird von Forschenden aus mindestens zwei unterschiedlichen Disziplinen durchgeführt. Hierbei werden disziplinspezifische Konzepte, Wissen und Methoden der beteiligten Bereiche innerhalb eines Projektes verwendet. Die Forschenden ergänzen sich dabei durch das Einbringen ihrer individuellen Expertise in das gemeinsame Forschungsprojekt. [5] Eine interdisziplinäre Zusammenarbeit profitiert dann am meisten, wenn die entwickelten Methoden in den Disziplinen aller Projektbeteiligten verwendet werden können oder aber auf Fragestellungen aller Disziplinen transferierbar sind. Dies erfordert jedoch meist, dass es ähnliche Fragestellungen oder Bedarfe gibt.

Sowohl in den Materialwissenschaften als auch in der Radiologie spielt die Analyse von Bildern eine wichtige Rolle. Dabei werden zunächst Daten in Form von Bildern mit verschiedenen Bildgebungstechniken gemessen. Im Anschluss werden diese Bilder analysiert, um Informationen über das gemessene Objekt zu erhalten. Die Informationen dienen dann in einem finalen Schritt dazu, weitere Maßnahmen zu ergreifen bzw. Handlungsempfehlungen zu erstellen. Eine weitere Gemeinsamkeit der Materialwissenschaften und der Radiologie ist das Ziel, eine möglichst vollständige Automatisierung einer Methode zu erreichen. Die wesentlichen Vorteile einer vollständigen Automatisierung liegen in der Vergleichbarkeit

der Ergebnisse, da es keine subjektiven, menschlichen Einflussgrößen gibt. Zusätzlich erspart eine Automatisierung auch Zeit und entlastet damit Berufsgruppen wie z. B. RadiologInnen, die unter einer großen Arbeitslast leiden. Des Weiteren ermöglicht eine vollautomatische Methode auch die zuvor erwähnte Analyse von Datensätzen, die aufgrund von fehlender Expertise am eigenen Standort, sonst nicht analysierbar wären. Weitere Gemeinsamkeiten sind die mehrdimensionalen (Bild-)Daten, sowie das Ziel, die gewählten Methoden auf verschiedene Datensätze (Erkrankungen, Materialien, etc.) und Bildgebungstechniken (z.B. Scanner) zu übertragen.

Das Auffinden von Objekten bzw. Strukturen in einem Bild, Segmentierung genannt, spielt eine zentrale Rolle in der Bildverarbeitung. In den Materialwissenschaften lassen sich aus den Segmentierungen digitale Zwillinge von Materialien erstellen, die dann mithilfe verschiedener computergestützer Verfahren, wie z. B. der Simulation oder Methoden des maschinellen Lernens (ML) weiter untersucht werden können. Durch eine quantitative Analyse lassen sich hieraus Materialeigenschaften herleiten. Mithilfe der datengetriebenen Analyse, lassen sich hierdurch neue Materialien entwickeln. Auch in der Radiologie ist die Segmentierung der erste zentrale Schritt zur Durchführung von quantitativen Analysen. Für die Bewertung der Lungengesundheit ist die Quantifizierung der Perfusion in der kontrastmittelunterstützten Magnetresonanztomographie (MRT) der Lunge ein vielversprechendes Verfahren. Während die Perfusionsquantifizierung in anderen Organen weit verbreitet und etabliert ist, ist die Anwendung in der Lunge bisher jedoch nur auf kleinere oder mononzentrische Studien beschränkt.

In dieser Arbeit werden robuste und flexible Methoden zur quantitativen Analyse multidimensionaler Bilddaten entwickelt. Durch den Fokus auf eine große Flexibilität und Robustheit der entwickelten Methoden, lassen sich diese disziplinübergreifend verwenden oder ermöglichen den breiten Einsatz unter verschiedensten Einsatzbedingungen. Hierdurch wird das Problem des zeitintensiven Entwicklungsprozesses von Methoden am Beispiel der Materialentwicklung und

der Erforschung von Lungenerkrankungen für zwei wesentliche Forschungsbereiche gelöst, um gegen die zu Beginn genannten gesellschaftlichen Probleme effizienter vorgehen zu können.

1.2 Zielsetzung und Beitrag der Arbeit

Das übergeordnete Ziel dieser Arbeit ist die Entwicklung von robusten und flexiblen Methoden für die quantitative Analyse von multidimensionalen Bilddaten. Die entwickelten flexiblen Methoden sollen disziplinübergreifend nutzbar sein und damit die Notwendigkeit von Neuentwicklungen für jede einzelne Disziplin obsolet machen. Dagegen besteht das Ziel für die robusten Methoden darin, die Anwendbarkeit der Methode unter verschiedensten äußeren Bedingungen zu ermöglichen und damit einen breiteren Einsatz innerhalb des Anwendungsgebiets zu gewährleisten. Die Methodenentwicklung wird dabei durch die Erarbeitung eines Konzeptes zur automatischen Dokumentation der Forschungsergebnisse ergänzt.

Im ersten Anwendungsfall besteht das Ziel in der Entwicklung einer flexiblen Methode, die ein gemeinsames Problem löst und damit in beiden Bereichen eingesetzt werden kann. Hierzu soll für die Segmentierung von Objekten in mehrdimensionalen Bildern eine Methode entwickelt werden, die sich leicht an unterschiedliche Anwendungen anpassen lässt. Dafür sollen verschiedene Segmentierungsmethoden entwickelt und validiert werden und anhand von spezifischen Bilddaten aus der Radiologie und den Materialwissenschaften getestet werden.

Der zweite Anwendungsfall beschäftigt sich mit der Weiterentwicklung einer bestehenden Methode zur Perfusionsquantifizierung von MRT-Bildern der Lunge. Die zugrundeliegende physikalische Methode ist zwar in der Forschung bereits teilweise verbreitet, findet jedoch im klinischen Alltag kaum Anwendung. Der Grund hierfür liegt in der bisher fehlenden Erprobung und Validierung in heterogenen Populationen wie multizentrischen Studien. Das Ziel für diesen Anwendungsfall ist die vollständige Automatisierung einer Pipeline zur Quantifizierung der

Perfusion in Lungen-MRT Bildern. Außerdem soll die Robustheit der Pipeline erhöht werden, damit diese in multizentrischen Studien zum Einsatz kommen kann. Hierzu werden die Einflüsse von unterschiedlichen MRT-Scannern genau untersucht.

Abschließend soll für die Dokumentation von Forschungsdaten in medizinischen Forschungsprojekten ein einheitliches Vorgehen entwickelt werden, das die Nutzung von medizinischen Bilddaten erleichtert und alle einzelnen Schritte bei der Verarbeitung dieser Daten dokumentiert. Dies soll die Reproduzierbarkeit sowie Nachvollziehbarkeit der erlangten Ergebnisse unterstützen und damit ein größeres Vertrauen in die verwendeten Methoden schaffen.

Diese Arbeit besitzt vier wesentliche wissenschaftliche Beiträge zur interdisziplinären Bildanalyse:

1. Ein Konzept zur informationsgetriebenen Bildanalyse.
2. Die Entwicklung und Erprobung von verschiedenen Methoden zur unüberwachten Segmentierung von Objekten in multidimensionalen Bilddaten.
3. Eine robuste und vollautomatisierte Pipeline zur automatischen Perfusionsquantifizierung in Lungen-MRT Bildern, die den breiten Einsatz der Methode in multizentrischen Studien ermöglicht.
4. Daten-Workflows zur automatischen Protokollierung von medizinischen Forschungsdaten in multizentrischen Forschungsprojekten, bei denen sowohl algorithmische Methoden, als auch Methoden des maschinellen Lernens zum Einsatz kommen.

1.3 Gliederung der Arbeit

Diese Arbeit gliedert sich in insgesamt sieben Abschnitte, die im Zusammenspiel das Vorgehen zur Entwicklung von robusten und flexiblen Methoden zur

quantitativen Analyse multidimensionaler Bilddaten beschreiben. Im nachfolgenden Kapitel 2 werden die nötigen theoretischen Grundlagen für diese Arbeit gelegt. Zunächst wird ein Grundverständnis für den Aufbau und die Funktion der Lunge sowie spezifischer Lungenerkrankungen vermittelt. Anschließend werden Grundlagen der Bildgebung behandelt. Die in dieser Arbeit verwendeten Bildgebungstechniken Computertomographie und Magnetresonanztomographie werden im Anschluss nochmals einzeln im Detail behandelt und es wird auf die Besonderheiten des Lungen-MRT eingegangen. Das Grundlagenkapitel schließt mit einer Beschreibung von ausgewählten Methoden des maschinellen Lernens. Hier werden neben klassischen neuronalen Netzen auch populationsbasierte Metaheuristiken und zufällige Markov-Felder behandelt. Kapitel 3 stellt das entwickelte Konzept der informationsgetriebenen Bildanalyse vor, das eine wesentliche Grundlage für die nachfolgenden Kapitel darstellt. Zwei Bestandteile dieses informationsgetriebenen Bildanalysezyklus werden in den Kapiteln 4 und 5 näher betrachtet, indem für zwei unterschiedliche Anwendungsfälle flexible und robuste Methoden entwickelt werden. Kapitel 4 beschäftigt sich mit der Entwicklung und Validierung von verschiedenen Methoden zur unüberwachten Segmentierung von generischen, multidimensionalen Bilddaten. Die entwickelten Methoden werden jeweils an den zwei Anwendungsfällen Lungensegmentierung und der Segmentierung von porösen Schaumstrukturen getestet und erprobt. In Kapitel 5 wird die Weiterentwicklung einer automatischen und robusten Pipeline zur Perfusionsquantifizierung in Lungen-MRT Bildern beschrieben. Hierbei werden die einzelnen Schritte und Algorithmen innerhalb der Pipeline zunächst detailliert erläutert. Anschließend wird die entwickelte Pipeline anhand einer monozentrischen und einer multizentrischen Studie erprobt und getestet. Im Anschluss wird in Kapitel 6 das Konzept zur Dokumentation von medizinischen Forschungsdaten charakterisiert, das einen zentralen Aspekt des informationsgetriebenen Bildanalysezyklus darstellt. Abschließend werden in Kapitel 7 die Ergebnisse zusammengefasst und Empfehlungen für weitere Forschungsarbeiten auf diesem Gebiet gegeben, die auf dieser Arbeit aufbauen könnten.

2 Theoretische Grundlagen

2.1 Lunge und Atemwege

Die Lunge ist eines der wichtigsten Organe im menschlichen Körper. Durch den Austausch von Kohlendioxid und Sauerstoff, reguliert sie den Sauerstoffgehalt im Blut und ermöglicht damit die Versorgung des gesamten Körpers mit Sauerstoff. Die folgenden Beschreibungen über deren Aufbau, als auch die ablaufenden Prozesse innerhalb der Lunge und der beteiligten anderen Systeme beruhen auf dem Buch von Bösch [6] und dem Buchkapitel von Lebiedz [7] und sollen ein fundamentales Verständnis für die Lunge, ihre Funktion und über mögliche Krankheiten liefern. Für ein detaillierteres Verständnis wird auf die beiden oben genannten Werke und entsprechende Fachliteratur verwiesen.

2.1.1 Aufbau und Funktion

Die Lunge lässt sich übergeordnet in die linke und die rechte Lungenhälfte unterteilen. Die Lungenhälften lassen sich wiederum in einzelne Lungenlappen, wie in Abbildung 2.1 (a) dargestellt, aufteilen. Die rechte Lunge besteht aus dem rechten Oberlappen (ROL), dem rechten Mittellappen (RML) und dem rechten Unterlappen (RUL). Die linke Lunge kann entweder in die drei Lappen linker Oberlappen (LOL), Lingula und linker Unterlappen (LUL) unterteilt werden, oder aber in nur zwei Lappen, indem die Lingula zum linkem Oberlappen gezählt wird. [6] Neben der Lunge sind die dazugehörigen Atemwege der zweite wesentliche Bestandteil des Atmungssystems. Die Atemwege übernehmen den Gastransport von und in die Lungenhälften. Anatomisch lassen sich die Atemwege in die oberen und unteren

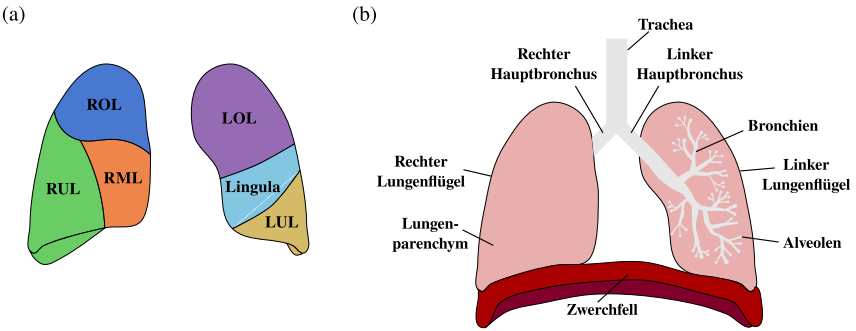


Abbildung 2.1: Aufbau der Lunge und Atemwege: (a) Unterteilung der Lungenhälften in Lappen (b) Beschreibung der Bestandteile der Lunge und Atemwege

Atemwege aufteilen. Die oberen Atemwege bestehen aus dem Nasenrachenraum und der Pharynx, eine gemeinsame Verbindung des Atem- und Speiseweges, der den Nasenrachenraum mit der Luftröhre (Trachea) verbindet. Die unteren Atemwege sind in Abbildung 2.1 (b) schematisch dargestellt. Die Trachea spaltet sich in den linken und rechten Hauptbronchus auf, welche sich dann selbst stufenweise weiter in einzelne Bronchien aufteilen. Am Ende der kleinsten Bronchien befinden sich die Alveolen, die eine zentrale Rolle beim Gasaustausch einnehmen. [6, 7]

Der dritte wesentliche Bestandteil der Atmung sind die in Abbildung 2.2 (a) dargestellten pulmonalen Blutgefäße. Von der rechten Herzkammer wird sauerstoffarmes Blut über die Lungenarterien, auch Pulmonalarterien genannt, in die Lunge transportiert. Nach dem dort stattfindenden Gasaustausch transportieren die Lungenvenen das sauerstoffreiche Blut von der Lunge über den linken Herzvorhof zur linken Herzkammer, die das Blut über die Aorta in die Körperarterien verteilt, die dann wiederum die Organe mit sauerstoffreichem Blut versorgen. Der Gasaustausch in der Lunge ist ein Wechselspiel zwischen der Ventilation, der Diffusion und der Perfusion. Die Ventilation beschreibt den Transport von Sauerstoff (O_2) in die Lungen und den Abtransport von Kohlenstoffdioxid (CO_2) aus den Lungen. Die Ventilation wird durch die Aktivität der Atemmuskulatur ausgelöst. Die Atemmuskulatur besteht aus dem Zwerchfell, der Interkostalmuskulatur, den Unterrippenmuskeln und der Atemhilfsmuskulatur. Durch eine Kontraktion des

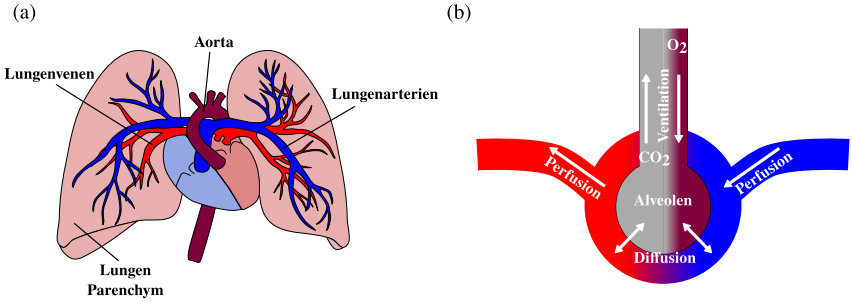


Abbildung 2.2: Pulmonale Blutgefäße und Gasaustausch: (a) Beschreibung der pulmonalen Blutgefäße (b) Schematische Darstellung des Gasaustausches

Zwerchfells im Zusammenspiel mit den anderen Atemmuskeln wird der Brustkorb gedehnt und die Lunge expandiert (Inspiration). Durch die Expansion der Lunge wird ein Unterdruck erzeugt, der Atemluft in die Lungen zieht und diese dann durch die Bronchien bis in die Alveolen transportiert. An den Alveolen findet der Gasaustausch statt, der in Abbildung 2.2 (b) skizziert ist. Durch die Perfusion (Durchblutung) der Lungenarterien wird sauerstoffarmes Blut von der rechten Herzkammer zu den Alveolen transportiert. Die Alveolen sind von den feinsten Blutgefäßen der Lungenarterien und -venen netzartig umgeben. Durch eine Partialdruckdifferenz zwischen Blutgefäßen und Alveolen wird ein Diffusionsprozess ausgelöst. Die Diffusion sorgt dafür, dass der Sauerstoff aus den Alveolen in die Lungenvenen wandert und umgekehrt Kohlenstoffdioxid von den Lungenarterien in die Alveolen diffundiert. Die Diffusionskapazität hängt im Wesentlichen von der Gesamtaustauschfläche ab, die aufgrund der großen Oberfläche aller Alveolen mit 50–100 m^2 in einer gesunden Lunge sehr groß ist. Das nun sauerstoffreiche Blut wird durch die Perfusion in den Lungenvenen zurück zum Herzen über den linken Vorhof in die linke Herzkammer transportiert, von wo aus es durch die Aorta und die Körperarterien zu den Organen fließt und diese mit Sauerstoff versorgt. Die Exspiration geschieht dagegen automatisch, ausgelöst durch elastische Rückstellkräfte im Thorax. Durch die Exspiration zieht sich die Lunge zusammen und presst die Luft, die nun vorwiegend aus dem in den Alveolen vorhandenen CO_2 besteht, durch die Bronchien, die Trachea und die oberen Atemwege wieder heraus. [6, 7]

2.1.2 Ausgewählte Lungenerkrankungen

Erkrankungen der Lunge sind sehr vielschichtig und können eine Vielzahl unterschiedlicher Ursachen haben. Je nach Erkrankung sind lediglich einzelne Bestandteile der Lunge betroffen, in anderen Fällen wirkt sich die Krankheit sogar auf das komplette Herz-Lungen-System aus. Es existiert eine große Zahl verschiedener Lungenerkrankungen. Da es in dieser Arbeit insbesondere um die konkrete Methodenanwendung an Lungen-Bilddaten von an chronisch obstruktiver Lungenerkrankung und an Mukoviszidose erkrankten Personen geht, werden im Folgenden lediglich diese beiden Erkrankungen kurz beschrieben. Für einen vertiefenden Einblick in weitere Lungenerkrankungen wird auf die entsprechende Fachliteratur verwiesen [6, 8].

Mukoviszidose

Mukoviszidose, auch zystische Fibrose (CF, vom Englischen Cystic Fibrosis) ist eine der häufigsten genetischen Erkrankungen in der nordeuropäischen Bevölkerung [9] mit einer Prävalenz von 7,4 Fällen pro 100 000 EinwohnerInnen [10]. CF ist eine progressive, vererbliche, respiratorische Erkrankung, die durch eine Mutation im Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) Gen verursacht wird. Der CFTR Gendefekt führt zu einer Fehlfunktion der Chloridkanäle im Körper, die wiederum zu zähen, wasserarmen Sekreten/Schleim sämtlicher Körperdrüsen führt. [6, 11] Zystische Fibrose betrifft nicht nur die Lunge und Atemwege, sondern führt auch zu Fehlfunktionen in anderen Organen wie der Bauchspeicheldrüse oder Leber, was beispielsweise Kleinwüchsigkeit, männliche Unfruchtbarkeit oder Unterernährung verursachen kann [10]. In der Lunge wird durch die Fehlfunktion der Chloridkanäle der Selbstreinigungsmechanismus (Mukoziliäre Clearance) gestört, der Fremdpartikel wie Staub, aber auch Schleim aus den Lungen abtransportiert. Die nicht abtransportierten Fremdpartikel lösen in der Lunge chronische Infektionen und Entzündungen aus und schädigen nachhaltig die Lungenstruktur [6, 12]. Außerdem kann durch den zähen Schleim eine sogenannte Mukostase (engl. mucus plugging) entstehen, eine Verstopfung

der Atemwege mit Sekret, was ebenfalls zu chronischen Infektionen und Entzündungen führt [11]. In einigen Fällen entwickeln CF PatientInnen eine Dilatation der Bronchialarterien (BAD, vom Englischen Bronchial Artery Dilatation), kleine spiralförmige Arterien, die der Aorta entspringen und die Lunge selbst mit Nährstoffen versorgen. Das Vorhandensein von BAD kann zu einer massiven Lungenblutung führen, die oft lebensbedrohlich ist. [13]

Lange Zeit war CF eine Krankheit mit einer sehr geringen Lebenserwartung und erkrankte Personen starben oft bereits im Kindesalter. Die Forschung in den letzten Jahrzehnten hat jedoch zu einer deutlichen Verbesserung der Lebenserwartung geführt, sodass erkrankte Personen mittlerweile eine durchschnittliche Lebenserwartung von mehr als 50 Jahren besitzen. [10] Die größten Fortschritte sind das CF-Neugeborenenscreening, das eine frühzeitige Diagnose und damit frühzeitige Therapie ermöglicht, und die bei vielen CF-PatientInnen anwendbare CFTR-Modulator Therapie, die die Lungenfunktion und die Funktion anderer Organe deutlich verbessern kann. [9] Die Diagnose von CF geschieht meist mithilfe des erwähnten Neugeborenenscreenings und wird dann durch einen Schweißtest, bei dem die Schweißchloridkonzentration gemessen wird, bestätigt. Anschließend wird häufig ein Gentest zur Bestimmung der CFTR-Mutation durchgeführt, um zu bestimmen, ob eine CFTR-Modulator Therapie anwendbar ist. Für das Monitoring der Krankheit werden Lungenfunktionstests, Gasauswaschverfahren und Bildgebung eingesetzt [9, 10]. Lange Zeit galt für die Bildgebung die Computertomographie (CT) als Goldstandard, jedoch führt diese zu einer erhöhten Strahlenbelastung, besonders bei regelmäßigen Kontrollbesuchen. Mittlerweile wurde gezeigt, dass die Magnetresonanztomographie (MRT) erfolgreich für die Diagnose und Kontrolle bei CF PatientInnen eingesetzt werden kann [14].

Chronisch obstruktive Lungenerkrankung

Die chronisch obstruktive Lungenerkrankung (COPD, vom Englischen Chronic Obstructive Pulmonary Disease) ist die weltweit dritthäufigste Todesursache. Im Jahr 2020 betrug die weltweite Prävalenz für Menschen älter als 40 Jahre

10,6 %. [15] Die COPD ist charakterisiert durch eine Obstruktion der Atemwege. Auslöser für diese Obstruktion sind verstärkte Entzündungsreaktionen der Atemwege und der Lunge, die durch das Inhalieren von Schadstoffen wie Zigarettenrauch, schädlichen Gasen und Stäuben oder Luftverschmutzung verursacht werden. [6] Charakteristisch für die COPD sind chronische respiratorische Symptome (z. B. Husten oder Kurzatmigkeit), die durch Veränderungen der Atemwege wie Fibrose oder Bronchitis und/oder durch verstärkte Emphysembildung in den Alveolen hervorgerufen werden. Diese Ursachen und Symptome führen dann häufig zu dauerhaften und fortschreitenden Atemwegsobstruktionen. [16] Die Diagnose von COPD geschieht mittels des Lungenfunktionstests Spirometrie, bei dem die erkrankte Person in das Mundstück eines Spirometers atmet. Hierbei wird dann unter anderem das forcierte exspiratorische Volumen in einer Sekunde (FEV1) als auch die forcierte Vitalkapazität (FVC) gemessen. Eine Diagnose für COPD wird gestellt, wenn der Quotient beider Werte $\mathrm{FEV1}/\mathrm{FVC} < 0,7$ beträgt. [16] Mehrere Studien kamen jedoch zur Schätzung, dass ca. 70 % aller COPD Fälle undiagnostiziert bleiben, da PatientInnen beispielsweise ihre Symptome bei ihrem Arztbesuch nicht schildern oder bei einem Arztbesuch keine Spirometrie durchgeführt wird. [17]

Für die Klassifizierung des Schweregrades der Krankheit wird in der Praxis meist der GOLD-Grad (vom Englischen Global Initiative for Chronic Obstructive Lung Disease) verwendet. Dieser unterteilt COPD-PatientInnen in vier Gruppen GOLD 1–4 , wobei hier eine höhere Zahl eine schwerere Erkrankung bedeutet. [16] Zu den Therapiemöglichkeiten der COPD zählt die Rauchentwöhnung, die sportliche Betätigung und die Teilnahme an Patientenschulungsprogrammen. Die medikamentöse Behandlung umfasst oft die Behandlung mit Bronchodilatoren, die die Bronchien erweitern. Bei erkrankten Personen mit einem Sauerstoffmangel (Hypoxämie) kommt häufig auch eine Sauerstofflangzeittherapie zum Einsatz, während erkrankte Personen mit Atempumpversagen eine nichtinvasive Beatmungstherapie erhalten. In schwerwiegenden COPD Fällen ist die letzte Therapiemöglichkeit eine Lungentransplantation. [6] Zur Beurteilung des Krankheitszustands spielen bildgebende Verfahren wie CT und MRT eine wesentliche Rolle, da hier, neben

anatomischen und strukturellen Eigenschaften, auch funktionelle Eigenschaften, wie die Perfusion oder Ventilation untersucht werden können [18].

2.2 Bildgebung

Ein Bild ist normalerweise eine Darstellung im zweidimensionalen (2D) Raum. Hierbei besteht das Bild aus Bildzeilen und Bildspalten und ein einzelnes Element in diesem Gitter wird als Pixel bezeichnet. Ein Bild bezeichnet in manchen Fällen jedoch auch ein dreidimensionales (3D) Volumen. Ein 3D Bild kann im Wesentlichen auf zwei verschieden Arten erzeugt werden. Entweder werden einzelne 2D Bilder, die Schichten, gemessen und diese dann zu einem 3D Bild gestapelt, oder aber das 3D Bild wird direkt durch Messung eines kompletten Volumens erzeugt. In einem 3D Bild werden die kleinsten Elemente als Volumenelemente (Voxel) bezeichnet. Zur Vereinfachung wird im Folgenden immer der Begriff Voxel für die kleinsten Elemente in 2D und 3D Bildern verwendet. Bilder können mit verschiedenen Techniken erzeugt werden. Die am häufigsten verwendete Methode im Alltag ist die Kamera. Im wissenschaftlichen Kontext werden Bilder jedoch meist per Mikroskopie, Röntgen, Computertomographie oder Magnetresonanztomographie erzeugt. Die Bildgebung und Bildverarbeitung sind zentrale Techniken in vielen wissenschaftlichen Bereichen. Dabei steht das Ziel im Mittelpunkt die Eigenschaften eines Objektes, Materials oder einer zu untersuchenden Person zu messen, um basierend darauf weiterführende Maßnahmen zu ergreifen.

Für die Aufnahme eines Bildes werden verschiedene Techniken verwendet. Eine große Gruppe bilden die Verfahren der Mikroskopie, zu der unter anderem die Lichtbildmikroskopie, die Transmissionselektronenmikroskopie (TEM), die Rasterelektronenmikroskopie (REM) oder die Rasterkraftmikroskopie (AFM, vom Englischen Atomic Force Microscopy) gehören. Ein großer Nachteil dieser Methoden ist, dass mit ihnen keine zerstörungsfreie Bildgebung eines kompletten 3D Volumens möglich ist. [19, 20] Eine Röntgenaufnahme ist ein 2D Bild eines 3D Volumens, bei der Röntgenstrahlen durch ein Objekt geschossen werden und auf der anderen Seite von einem Detektor wieder aufgefangen werden. Durch die

Interaktion mit dem Objekt werden die Röntgenstrahlen teilweise vom Material/Gewebe absorbiert, wodurch sich die gemessenen Röntgenstrahlen von den Ausgangsstrahlen unterscheiden. Aus dem vom Detektor aufgefangenen Signal lässt sich dann ein 2D Bild erzeugen. Ein Nachteil dieser Methode ist jedoch, dass das Signal eines gemessenen Röntgenstrahles, immer die Absorption aller Materialien/Gewebe entlang der Strahlrichtung enthält, wodurch das 2D Bild der Summe der Materialabsorptionen des 3D Objektes entspricht. [21] Für die Aufnahme von 3D Bildern sind die Computertomographie und die Magnetresonanztomographie zwei der wichtigsten Verfahren. Da das CT und MRT eine wesentliche Rolle für diese Arbeit spielen, ist diesen jeweils nachfolgend ein eigener Abschnitt gewidmet.

2.3 Computertomographie

Die Computertomographie nutzt ebenso wie das Röntgengerät Röntgenstrahlen, kann jedoch Schichtbilder und damit ein komplettes Volumen messen und davon ein 3D Bild generieren. Das Design eines CT-Scanners unterscheidet sich je nach Anwendung. In der klinischen Diagnostik besteht der CT-Scanner aus einer Röhre, in deren Inneren ein Vakuum herrscht. Eine Elektronenquelle schießt einen Elektronenstrahl auf eine ringförmige Röntgenquelle. Diese angeregte Röntgenquelle emittiert nun Röntgenstrahlen, die die zu untersuchende Person durchstrahlen und auf der anderen Seite von einem ringförmigen Röntgendetektor aufgefangen werden. Im Gegensatz dazu bestehen industrielle CT-Scanner, die beispielsweise für die Messung von Materialien eingesetzt werden, meist aus einer Röntgenquelle und einem gegenüberliegenden bogenförmigen Röntgendetektor. Während der Messung dreht sich nun entweder das System aus Röntgenquelle und -detektor oder aber das Messobjekt selbst. Sowohl für klinische als auch industrielle CT-Scanner wird das 3D Bild mithilfe eines mathematischen Rekonstruktionsverfahrens generiert, das das gemessene Signal verschiedener Röntgenstrahlen kombiniert und damit einen Intensitätswert für jeden Voxel im 3D Bild berechnet. [22] Dabei beschreibt die Intensität im CT die Abschwächung des Röntgenstrahls durch

das Gewebe/Material. Dadurch sind die Intensitäten im CT, bei passender Normierung, bereits quantitative Werte, die zum Vergleich zweier Messungen verwendet werden können. Die Standardgröße im CT sind die Hounsfield-Einheiten (HU, vom Englischen Hounsfield Unit), die die Abschwächung der Röntgenstrahlung quantifizieren. [23] Im Bereich des CT gab es verschiedene Weiterentwicklungen, die es erlauben Bilder mit immer größerer Auflösung zu messen. Das Photon-Counting-CT und das Mikro-CT (μCT) sind in der Lage im Mikrometer Bereich zu messen, während das Nano-CT und Synchrotron-CT sogar eine Messung im Nanometer Bereich ermöglichen. [24, 25]

2.3.1 Quantitatives CT

Das quantitative CT (QCT) beschreibt die Berechnung von quantitativen Werten, die anhand von CT-Bildern berechnet werden. Eine weit verbreitete quantitative CT-Metrik sind die Parametric Response Maps (PRM), die zunächst für das MRT entwickelt [26] und dann später auch für das CT eingesetzt wurden [27]. Die Grundidee der PRM-Technik in der Lunge ist, dass zwei verschiedene CT-Aufnahmen, eine Aufnahme in Inspiration und eine zweite in Exspiration, aufgenommen werden. In beiden Aufnahmen wird die Lunge segmentiert und anschließend werden beide Bilder mithilfe einer deformierbaren Registrierung aufeinander registriert. Anschließend werden beide CT-Bilder einem Thresholding unterzogen. Im inspiratorischen Bild werden alle Voxel kleiner einem Schwellwert von -950 HU als Emphysem klassifiziert. Im exspiratorischen Bild werden alle Voxel mit einem Wert von weniger als -856 HU als Gaseinschluss definiert. Die zwei Schwellwerte werden auf ein gemeinsames Histogramm angewendet und es findet anschließend eine Gruppierung in drei Klassen statt. Die erste Klasse $\text{PRM}_{\text{normal}}$ beinhaltet normale/gesunde Voxel, die zweite Klasse PRM_{fSAD} entspricht Voxeln mit funktionaler Erkrankung der kleinen Atemwege (fSAD, vom Englischen Functional Small Airways Disease) und die dritte Klasse PRM_{emph} entspricht Emphysem-Voxeln. Die Einteilung in die verschiedenen PRM-Klassen und das Vorgehen der PRM-Methode sind in Abbildung 2.3 skizziert.

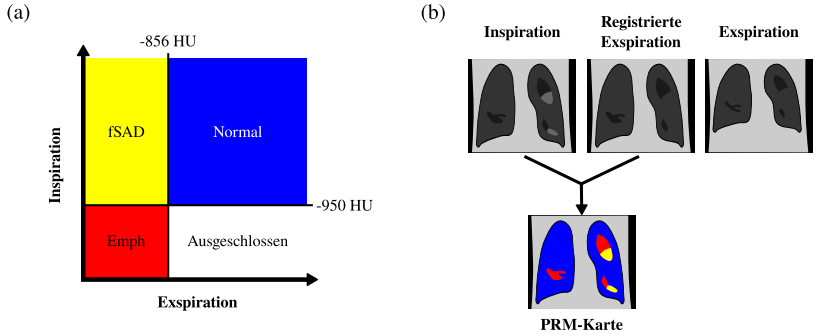


Abbildung 2.3: Erklärung der PRM-Methode: (a) Einteilung in die drei PRM-Klassen (b) Berechnung der PRM-Karte aus inspiratorischen und exspiratorischen CT-Bild der Lunge.

Die drei PRM-Klassen sind gut geeignet, um drei COPD-Phänotypen zu definieren, die PatientInnen mit schwerpunktmäßiger Emphysem-geprägter Erkrankung, die PatientInnen mit schwerpunktmäßiger Erkrankung der kleinen Atemwege und PatientInnen, die eine gemischte Form aufweisen. [27] Weitere quantitative Metriken umfassen meist die genaue Messung der Atemwege, indem diese automatisch segmentiert werden. Hierdurch lassen sich die Durchmesser einzelner Atemwege messen und das Gesamtvolumen der Atemwege bestimmen. Zusätzlich kann nach einer Segmentierung der Lunge auch das Lungenvolumen berechnet werden. Andere QCT-Werte sind die mittlere Lungendichte (MLD) oder der Emphysem-Index (EI). [28–30]

2.4 Magnetresonanztomographie

In diesem Abschnitt soll ein Grundverständnis für die Magnetresonanztomographie gelegt werden. Hierzu wird sowohl auf allgemeine Aspekte des MRT, als auch auf spezifische Besonderheiten für das Lungen-MRT, die morphologische Lungenbildgebung und auf das Perfusions-MRT eingegangen. Für eine zusätzlich vertiefende Auseinandersetzung mit dem Thema wird auf Standardwerke, wie das von Weishaupt et al. [31] verwiesen. Die Informationen in diesem Unterkapitel beziehen sich, wenn nicht anders erwähnt, auf dieses Buch von Weishaupt et al.

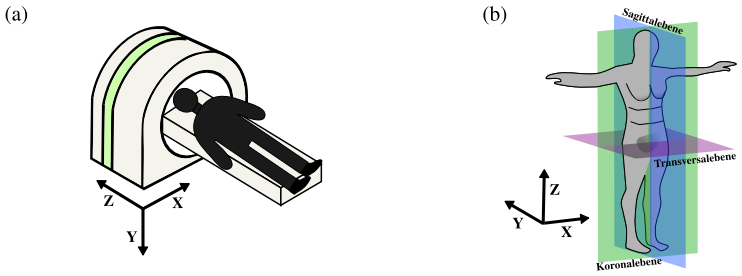


Abbildung 2.4: Koordinatensystem und Ebenen im MRT:
(a) Patienten Koordinatensystem (b) Bildebenen des MRT

2.4.1 Grundlagen

Die wesentlichen Bestandteile eines MRT-Scanners sind der Magnet, das Gradientensystem, bestehend aus verschiedenen Gradientenspulen, das Hochfrequenzsystem und ein Computer. Der Magnet erzeugt das permanente Hauptmagnetfeld B_0 des Scanners. Mithilfe der Gradientenspulen können zusätzliche, temporäre Magnetfelder erzeugt werden, die der Ortskodierung und Schichtauswahl dienen. Das Hochfrequenzsystem hat zwei wesentliche Komponenten, den Sender und den Empfänger. Mit dem Sender können Hochfrequenzpulse (HF-Pulse) für die Auslenkung von Spins erzeugt werden. Der Empfänger besteht aus mehreren Empfangsspulen und dient dem Empfang der induzierten Wechselspannungen. Mithilfe des Computers wird das ganze System gesteuert und die Wechselspannungen in Signale und Bilder umgewandelt. [31] Das üblicherweise verwendete Koordinatensystem für das MRT ist in Abbildung 2.4 (a) dargestellt. Beim MRT werden für eine einfachere Betrachtung der Bilder durch die RadiologInnen aus den 3D-Daten Schnittbilder durch den Körper angefertigt. Dabei wird zur Beschreibung der Bildausrichtung meist der Name der Schnittebene verwendet. Die drei Schnittebenen sind in Abbildung 2.4 (b) dargestellt und werden als Koronalebene, Transversalebene und Sagittalebene bezeichnet.

Das klassische klinische MRT nutzt die Protonen von Wasserstoffatomen für die Bildgebung. Ein Wasserstoffatom besitzt ein Proton, das aus einer positiven

elektrischen Ladung besteht und einen Spin besitzt. Dieser Spin, eine Drehbewegung des Protons um sich selbst, erzeugt aufgrund der rotierenden elektrischen Ladung ein magnetisches Moment M. Wird nun in einem äußeren Magnetfeld B_0 eingeschaltet, so wirkt dieses auf die Spins und verändert deren Rotationsachse. Dies führt zu einer Ausweichbewegung der Spins, der sogenannten Präzessionsbewegung. Diese Präzessionsbewegung geschieht mit einer zum äußeren Magnetfeld proportionalen Frequenz, der Larmorfrequenz ω_0. Die Spins richten sich langsam entlang der B_0-Richtung aus, wodurch die sogenannte Längsmagnetisierung M_Z entsteht. Wird nun Energie mit der Larmorfrequenz in Form eines HF-Pulses eingebracht, so kann der Längsmagnetisierung-Vektor bei richtig gewählter Leistung und Dauer in die XY-Ebene ausgelenkt werden, wodurch aus der Längsmagnetisierung die Transversalmagnetisierung M_{XY} wird. Das Kreisen in der XY-Ebene induziert in den Empfangsspulen eines MRT-Scanners eine Wechselspannung mit der Larmorfrequenz, die in das MR-Signal umgewandelt wird. [31]

Relaxation

Das MR-Signal, das durch das Kreisen in der XY-Ebene erzeugt wird, nimmt mit der Zeit wieder ab. Verantwortlich für die Abnahme des Signals sind die Spin-Gitter Relaxation (T1-Relaxation) und die Spin-Spin-Relaxation (T2-Relaxation). Die T1-Relaxation beschreibt den Vorgang des Zurückkippens der Magnetisierung von der XY-Ebene in die Z-Richtung. Dieser Vorgang kann durch eine Exponentialgleichung mit dem Koeffizienten T1 beschrieben werden, daher der Name T1-Relaxation. Neben der Stärke eines Spins besitzt dieser auch eine Phase (Winkel in der XY-Ebene). Nach der Anregung durch den HF-Puls sind alle Spins in Phase. Da manche Spins jedoch schneller präzedieren als andere, ergibt sich zwischen einzelnen Spins ein Phasenunterschied. Diesen Vorgang bezeichnet man als Dephasierung. Da das MR-Signal aus der Summe der Vektoren der einzelnen Spins besteht, führt eine Dephasierung der Spins dazu, dass sich die einzelnen Magnetisierungsvektoren gegenseitig aufheben, wodurch sich das MR-Signal verringert. Die Dephasierung wird hierbei durch zwei verschiedene Effekte ausgelöst, die lokalen Magnetfeldänderungen durch benachbarte Spins und Inhomogenitäten

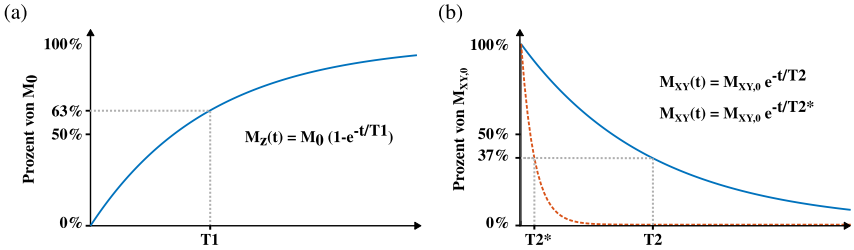


Abbildung 2.5: Schematische Darstellung der (a) T1-Relaxation und (b) T2- bzw. T2*-Relaxation.

des äußeren Magnetfeldes durch den Scanner oder den gemessenen Körper. Die Dephasierung durch die lokalen Magnetfeldänderungen kann ebenso durch eine Exponentialgleichung beschrieben werden, die einen Koeffizienten T2 besitzt (T2-Relaxation), während die Dephasierung aufgrund der Magnetfeldinhomogenitäten des Scanners und des Körpers durch die gleiche Exponentialgleichung mit einem Koeffizienten T2* beschrieben werden kann (T2*-Relaxation). Die T1- und T2- bzw. T2*-Relaxation laufen gleichzeitig ab und sind voneinander unabhängig. [31] Eine Visualisierung der T1-, T2- und T2*-Relaxation ist in Abbildung 2.5 schematisch dargestellt.

Bildkontrast

Der Bildkontrast im MRT wird durch die verschiedenen Eigenschaften von unterschiedlichen Geweben erzeugt. Das Signal eines Gewebes ist abhängig von dessen Protonendichte, dessen T1-Zeit und dessen T2-Zeit. [31] Daneben gibt es drei weitere wesentliche Parameter, die den Bildkontrast beeinflussen. Die Repetitionszeit (TR) beschreibt die vergangene Zeit zwischen zwei Anregungen, also die Zeit, die Spins haben um sich wieder in die Z-Richtung zu bewegen. Die verfügbare Längsmagnetisierung für die nächste Anregung, wird bestimmt über den Anteil des Magnetisierunsvektors in Z-Richtung. Daher kann bei einem kurz gewählten TR, ein starker T1-Kontrast erzeugt werden, da Gewebe mit kurzem T1, deutlich mehr Längsmagnetisierung für die nächste Anregung besitzt, als Gewebe mit einer langen T1 Zeit. Das Gewebe mit kurzem T1 erzeugt hierdurch bei der

zweiten Anregung ein deutlich stärkeres Signal, als das Gewebe mit langer T1 Zeit. Eine ständige Auslenkung der Spins um 90°, würde, insbesondere bei einem kurzen TR, dazu führen, dass ein kleinerer Anteil an Längsmagnetisierung zur Verfügung steht. Mithilfe des zweiten Parameters, dem Flipwinkel, lässt sich dieser Effekt beeinflussen. Um den zuvor genannten Effekt zu minimieren wird häufig statt einem Flipwinkel von 90°, ein kleinerer Flipwinkel verwendet. Der dritte Parameter, die Echozeit (TE) bestimmt dagegen den T2-Kontrast eines Bildes. Hierbei beschreibt TE die Zeitspanne zwischen der Anregung und dem Auslesen des MR-Signals. Mithilfe der Gradientenspulen werden die Spins zunächst dephasiert und damit das MR-Signal zerstört. Das MRT-Bild kann jedoch erst gemessen werden, wenn die Spins wieder in Phase sind, hierzu werden erneut die Gradientenspulen genutzt, um diesmal die Dephasierung umzukehren. Das Echo findet dann statt, wenn alle Spins wieder in Phase sind. Eine längere Echozeit erzeugt einen stärkeren T2-Kontrast des MRT-Bildes. [31]

Bilderzeugung

Die Erzeugung eines MRT-Bildes kann entweder durch die Akquisition einzelner Schichten oder durch die Akquisition eines kompletten Volumens (engl. Slab) geschehen. Für die Messung einzelner Schichten besitzt der MRT-Scanner eine Gradientenspule, die einen Gradienten des Magnetfelds in Z-Richtung erzeugt. Der steigende Gradient des Magnetfeldes in Z-Richtung führt auch zu steigenden Larmorfrequenzen der Spins entlang der Z-Richtung. Da Spins nur angeregt werden, wenn die Anregungsfrequenz in etwa der Larmorfrequenz entspricht, können gezielt Spins in einer bestimmten Schicht angeregt werden und damit dann nur eine Schicht gemessen werden. Zusätzlich zur Schichtwahl findet eine Ortskodierung mithilfe von zwei weiteren Gradientenspulen, eine in X- und eine zweite in Y-Richtung, statt. Für die Ortskodierung wird eine Kombination aus Phasenkodierung und Frequenzkodierung verwendet. Welche Kodierungsvariante für welche Achse verwendet wird, hängt von der jeweiligen Messung ab. Wird eine Phasenkodierung in der Y-Richtung verwendet, so wird nach der Anregung

durch den HF-Puls ein Y-Gradient von oben nach unten geschaltet. Dieser Gradient führt dazu, dass die Larmorfrequenzen der Spins von oben nach unten sinken und dadurch die oberen Spins schneller präzedieren als die Unteren. Diese Phasenverschiebung bleibt auch dann bestehen, wenn der Gradient wieder abgeschaltet wird und damit die Larmorfrequenzen entlang der Y-Richtung wieder gleich sind. Die unterschiedlichen Phasen zwischen den einzelnen Zeilen ermöglichen die genaue Identifikation einer jeden Bildzeile. Für die Identifikation einer jeden Spalte in X-Richtung wird die Frequenzkodierung verwendet. Hierfür wird mit einer Gradientenspule ein X-Gradient erzeugt, der im Gegensatz zum Phasenkodierungsgradienten während der Messung dauerhaft angeschaltet bleibt. Durch den Gradienten erhält man erneut eine Sortierung der Larmorfrequenzen der Spins, wodurch jede Spalte anhand ihrer Larmorfrequenz eindeutig identifiziert werden kann.

Die gemessenen Signale liegen im sogenannten K-Raum, in dem jede Zeile einer Messung einem Phasengradienten entspricht. Das Zentrum des K-Raums enthält dabei schwerpunktmäßig Kontrastinformationen, während die äußeren Teile räumliche Details enthalten. Da das empfangene Signal ein Mischsignal aus den Signalen aller Spins ist, muss für die Generierung eines MRT-Bildes eine Auftrennung der Signale mithilfe des Wissens aus der Kodierung geschehen. Hierzu verwendet man eine zweidimensionale Fourier-Transformation. Mithilfe einer ersten Fourier-Transformation lässt sich das Signal anhand der einzelnen Frequenzen durch die Frequenzkodierung aufteilen. Für jede Frequenz besteht dieses Teilsignal, jedoch weiterhin aus allen Signalen jeder Zeile in Y-Richtung. Da sich die Phasen, die die Y-Position bestimmen, nicht mithilfe einer einzigen Messung unterscheiden lassen, werden mehrere Messsequenzen mit unterschiedlichen Phasenkodierungen durchgeführt. Hierbei korreliert die Bildqualität mit der Anzahl der aufgenommenen unterschiedlichen Phasenkodierungen. Anschließend kann mit einer zweiten Fourier-Transformation entlang der Phasenrichtung, eine Aufteilung in die einzelnen Phasen geschehen. Hierdurch lässt sich jetzt das Signal eines jeden Voxels im Bild eindeutig identifizieren. Soll statt der Messung einer einzelnen Schicht ein komplettes Slab gemessen werden, so muss auch eine Kodierung in der dritten Richtung (Z-Richtung im obigen Beispiel) durchgeführt

werden. Hierzu wird der Z-Gradient für eine weitere Phasenkodierung verwendet und es findet eine dritte Fourier-Transformation statt. Je mehr Messungen mit unterschiedlichen Phasen durchgeführt werden, desto besser ist die Auflösung und desto geringer die Schichtdicke in Z-Richtung. [31] Alle in dieser Arbeit verwendeten MRT-Sequenzen beruhen auf der Messung eines kompletten Slabs. Hierbei wird eine Fourier-Transformation in drei Richtungen durchgeführt.

Techniken zur Messbeschleunigung

Zur Verkürzung der Messzeit können verschiedene Messtechniken verwendet werden. Eine Möglichkeit sind parallele Bildgebungstechniken wie z. B. GRAPPA (vom Englischen GeneRalized Autocalibrating Partial Parallel Acquisition). GRAPPA nutzt das Wissen über die Position und Sensitivität der Empfangsspulen, um die Anzahl der Phasenkodierungsmessungen zu reduzieren. Hierdurch müssen weniger Linien im K-Raum gemessen werden und die benötigte Messzeit wird reduziert. [32] Eine weitere Technik, die für dynamische MRT Aufnahmen verwendet wird, ist das Keyhole-Imaging (deutsch: Schlüsselloch-Bildgebung). Das Zentrum des K-Raums enthält die Kontrastinformationen, während der äußere Teil des K-Raums räumliche Details wie die Morphologie enthält. Die Idee beim Keyhole-Imaging ist, dass der äußere Teil des K-Raums nicht bei jeder Messung gemessen wird, sondern nur einmal zu Beginn der Messung, in der der komplette K-Raum gemessen wird. Bei den restlichen Messungen wird nur das Zentrum des K-Raums gemessen, der alle notwendigen Daten enthält, um dynamische Änderungen des Kontrastes zu messen. Die restlichen Bildinformationen wie Kanten und Details werden dagegen aus der ursprünglichen Referenzmessung verwendet. Hierdurch müssen pro Zeitschritt deutlich weniger Linien im K-Raum gemessen werden und die temporäre Auflösung kann dadurch erhöht werden. Dabei wird angenommen, dass sich während der Messung der Kontrast ändert, aber die Bilddetails gleich bleiben. [33] Eine Weiterentwicklung der Keyhole-Technik ist die TRICKS-Technik (vom Englischen Time-Resolved Imaging of Contrast Kinetics), bei der die Referenzmessung des äußeren K-Raums zu regelmäßigen Zeitpunkten der Messsequenz neu gemessen und aktualisiert wird. [34]

Gradientenecho

Bei einem Gradientenecho (GRE) wird nach der Anregung durch einen HF-Puls ein negativer Gradient geschaltet, der zu einer (beschleunigten) Dephasierung der Spins führt. Anschließend schaltet man den negativen Gradienten aus und erzeugt einen doppelt so langen, positiven Gradienten gleicher Stärke, der zu einer Rephasierung der Spins führt. Der Zeitpunkt, in dem alle Spins wieder in Phase sind, bezeichnet man als Echo und die Zeit, die zwischen der Anregung und dem Auslesen bzw. dem Echo vergangen ist, als Echozeit. Der wesentliche Vorteil von Gradientenechos ist, dass aufgrund der kurzen Repetitionszeiten Bilder sehr schnell akquiriert werden können. Für die Erzeugung eines T1-Kontrastes im Bild muss übriges Signal aus der Voranregung zerstört werden. Hierzu wird ein spezieller Gradient geschaltet, der das verbliebene Signal zerstört (engl. spoiled). Sequenzen mit diesem Spoiling werden daher als gespoilte Gradientenechos bezeichnet.

Für eine Kombination von T1 und TR gibt es genau einen Flipwinkel, den sogenannten Ernst-Winkel, der ein maximales Signal liefert. Daher lässt sich durch die Veränderung von TR und dem Flipwinkel eine unterschiedlich starke T1-Gewichtung erzeugen. Verwendet man für ein Gradientenecho kleinere Flipwinkel, so führt dies zunächst zu einem kleineren MR-Signal, da M_{XY} geringer ist. Da jedoch für die nächste Anregung mehr M_Z zur Verfügung steht, ist das MR-Signal der kompletten Sequenz im Endeffekt größer, als für die Wahl eines größeren Flipwinkels. [31] Die in dieser Arbeit verwendeten morphologischen MRT Sequenzen sind T1-gewichtete gespoilte Gradientenechos mit kurzem TR, kurzem TE und einem kleinen Flipwinkel.

Die Unterschiede dieser unterschiedlichen Flipwinkel auf das MRT-Bild lassen sich in Abbildung 2.6 erkennen. Während im morphologischen VIBE Bild mit kleinerem Flipwinkel in Abbildung 2.6 (a) eine gute Unterscheidung von Lunge, Herz und umliegendem Gewebe möglich ist, sind die Unterschiede im TWIST Bild vor Ankunft des Kontrastmittels in Abbildung 2.6 (b) deutlich schwieriger zu erkennen. Der Grund für die größeren Flipwinkel im TWIST-Bild lässt sich

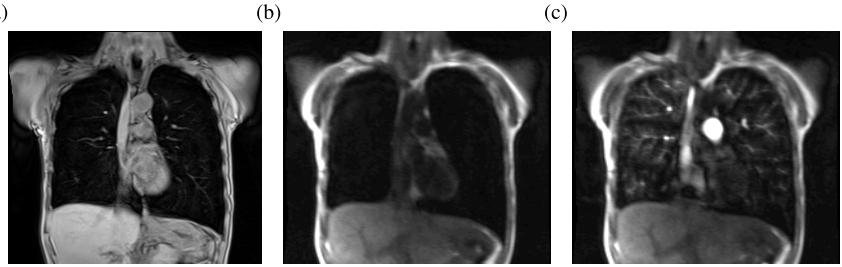


Abbildung 2.6: Vergleich zweier T1-gewichteter Bilder mit unterschiedlichen Flipwinkeln. (a) Koronale VIBE (Flipwinkel 5°) (b) TWIST (Flipwinkel 20°) vor Kontrastmittelankunft (c) TWIST nach Kontrastmittelankunft.

in Abbildung 2.6 (c) erkennen. Hier führt das Kontrastmittel in der Lunge zu einem erhöhten Signal des Lungenparenchyms und der Lungengefäße und damit kann die Perfusion gut abgebildet werden. Die Stärke des MR-Signals hängt bei Gradientenechos von der T1- und T2*-Relaxationszeit ab. Um die Messzeit des GRE zu verkürzen, wird häufig die oben beschriebene Beschleunigung durch parallele Bildgebung verwendet.

Zeitaufgelöste MR-Angiographie

Die zeitaufgelöste MR-Angiographie ist eine dynamische Messung, bei der meist die zuvor beschriebenen gespoilten Gradientenechos verwendet werden. Bei der MR-Angiographie werden komplette 3D Bilder zu verschiedenen Zeitpunkten aufgenommen, um den Blutfluss in den Blutgefäßen zu beobachten. Häufig kommen hier auch Gadolinium-haltige Kontrastmittel zum Einsatz, die den T1-Kontrast zusätzlich verstärken. Zur Verkürzung der Messzeit und der Erhöhung der zeitlichen Auflösung kommen hier entweder die oben erwähnte Keyhole-Technik, oder aber die TRICKS-Technik zum Einsatz. [35]

2.4.2 Besonderheiten des Lungen-MRT

Das MRT der Lunge besitzt drei wesentliche Herausforderungen. Zunächst ist die Protonendichte in der Lunge relativ gering, wodurch das Signal in der Lunge deutlich schwächer ausfällt als in den meisten anderen Organen. Außerdem führen die häufigen Übergänge von Gewebe zu Luft zu sogenannten Suszeptibilitätsartefakten, die die Bildsignale infolge von Magnetfeldinhomogenitäten verzerren und aufgrund der T2*-Dephasierung schwächen. Die dritte Herausforderung stellen die verschiedenen Bewegungen verursacht durch Atmung und Herzschlag dar. Je nach Pathologie können Lungenerkrankungen entweder zu einer höheren Protonendichte und damit einem stärkeren Signal (Plus-Pathologie), oder aber zu einer geringeren Protonendichte mit einem schwächeren Signal (Minus-Pathologie) führen, wie es beispielsweise bei chronisch obstruktiven Lungenerkrankungen der Fall ist. Ein weiterer Grund für unterschiedlich starke Signale ist der mögliche Einfluss von Lungenerkrankungen auf T2*, der sich durch strukturelle Änderungen in der Lunge ergibt. [36]

Aufgrund des schwachen Signals in der Lunge und den gleichzeitig vorhandenen Bewegungen ist die Aufnahme von MRT-Bildern in der Lunge deutlich schwerer als in anderen Organen. Zwei wichtige Parameter im MRT sind das Signal-Rausch-Verhältnis (SNR, vom Englischen Signal-To-Noise Ratio) und das Kontrast-Rausch-Verhältnis (CNR, vom Englischen Contrast-To-Noise Ratio). Um das geringe SNR in der Lunge zu erhöhen, gibt es grundsätzlich zwei Möglichkeiten. Entweder kann ein kurzes TE gewählt werden, damit die Signalverluste durch die T2*-Relaxation verringert werden. Oder es kann länger gemessen und damit mehr gemittelt werden, wobei hier der limitierende Faktor die Dauer eines Atemanhalts ist. [37]. Daher muss immer zwischen einer höheren Bildauflösung, einer vertretbaren Messzeit und einem höheren SNR abgewogen werden.

Morphologische Lungen-Bildgebung

Ein klassisches MRT-Protokoll sieht eine Reihe von morphologischen Bildern vor, die unterschiedliche Kontraste besitzen und daher jeweils in der Lage sind unterschiedliche Pathologien zu erkennen. T1-gewichtete Bilder sind besonders geeignet, um das Mediastinum (Bereich rund um das Herz), Lungenknoten oder Lungenrundherde darzustellen. T2-gewichtete Bilder werden dagegen verwendet, um pulmonale Infiltrationen, Schleim- oder Blutansammlungen darzustellen. [36] Die in dieser Arbeit verwendeten morphologischen MRT-Bilder sind T1-gewichtete 3D Gradientenechos, die mithilfe von parallelen Bildgebungstechniken unterstützt werden (z. B. SIEMENS VIBE Sequenz). Durch die Beschleunigung der parallelen Bildgebung lässt sich das komplette Lungenvolumen mit einer Schichtdicke von 4–5 mm innerhalb eines inspiratorischen Atemanhalts aufnehmen. [37]

Funktionale Bildgebung

Das MRT besitzt im Vergleich zum Röntgen oder der Computertomographie neben der Strahlungsfreiheit einen zweiten wesentlichen Vorteil, die Stärke der funktionalen Bildgebung. Mit dem MRT lassen sich mehrere, aufeinander folgende Messungen durchführen, wodurch Prozesse wie Bewegungen im Thorax, die Ventilation, die Perfusion oder die Diffusion gemessen werden können. [38] Für die Perfusion spielt die hypoxische pulmonale Vasokonstriktion (HPV, auch Euler-Liljestrand Mechanismus genannt) eine wesentliche Rolle, da sie die Lungenperfusion mit der Lungenventilation verbindet. Da die HPV auf die Sauerstoffwerte im Blut reagiert, wird Blutfluss aus Bereichen mit blockierten Lungenbereichen in Bereiche umgeleitet, die noch funktional sind, um eine optimale Sauerstoffversorgung zu garantieren. [39, 40] Aus diesem Grund kann man in chronisch obstruktiven Erkrankungen, wie der COPD, von Perfusionsdefekten auf eine gestörte Ventilation schließen. Für die Messung der Ventilation mithilfe des MRT gibt es verschiedene Methoden. Bei den gasbasierten Methoden unterscheidet man hierbei zwischen dem MRT mit hyperpolarisierten Gasen (z. B. ^{3}He-MRT,

^{129}Xe-MRT), dem sauerstoffverstärkten MRT (O_2-MRT) und dem MRT mit fluorierten Gasen (F_2-MRT). Diese Methoden können die lokale Ventilation in der Lunge quantifizieren, besitzen jedoch große technische Ansprüche und Kosten und sind daher nicht breit verfügbar. [40] Eine zweite große Gruppe stellen die Signal-Oszillations-Methoden dar, die eine Kombination aus Protonen-MRT und Bildregistrierung verwenden, um sowohl die Ventilation, als auch die Perfusion zeitgleich zu erfassen. Beispiele für Signal-Oszillations-Methoden sind Phase-Resolved Functional Lung (PREFUL) [41] oder das Matrix-Pencil Decomposition MRT (MP-MRT) [42]. Der wesentliche Vorteil dieser Methoden ist, dass hier kein Kontrastmittel notwendig ist, jedoch lassen sich hiermit immer nur ein paar Schichten und nicht das komplette Lungenvolumen messen. Außerdem mangelt es noch an Verifikationsmethoden, um die Techniken zu validieren.

Die am breitesten verfügbare und am besten validierte Methode für Perfusions-Bildgebung ist das dynamische kontrastmittelverstärkte MRT (DCE-MRT, vom Englischen Dynamic Constrast-enhanced MRI), das die Injektion eines Kontrastmittels benötigt. Der wesentliche Vorteil des DCE-MRT ist, dass die hierfür verwendeten MRT-Sequenzen weit verbreitet sind und keine besonderen Hardwarekomponenten benötigt werden. Daher kann das DCE-MRT in fast allen Krankenhäusern mit einem MRT-Gerät durchgeführt werden und entspricht damit der klinischen Versorgungsrealität. [40, 43] In dieser Arbeit wurde für Perfusions-MRT Aufnahmen die DCE-MRT Technik verwendet, da diese, insbesondere in anderen Organen bereits intensiv validiert wurde und es daher Referenzliteratur für das DCE-MRT gibt.

2.4.3 DCE-MRT der Lunge

Für die Messung eines DCE-MRT kommt ein Gadolinium-basiertes Kontrastmittel (GBCA) zum Einsatz, das intravenös injiziert wird. Durch GBCA wird die T1-Relaxation beschleunigt und damit das Signal in T1-gewichteten Bildern lokal erhöht. Für die Ausbildung eines klar definierten Bolus mit schnellem Anstieg und schneller Abnahme, wird das Kontrastmittel mit einer hohen Injektionsrate

injiziert und direkt nach dem Kontrastmittel ein sogenannter Chaser in Form einer Kochsalzlösung injiziert, damit das Kontrastmittel nicht in der Armvene hängen bleibt. Beim DCE-MRT kommt als MRT-Sequenz ein T1-gewichtetes gespoiltes Gradientenecho kombiniert mit parallelen Bildgebungstechniken und der Keyhole-Technik zum Einsatz, zum Beispiel die in dieser Arbeit verwendete TWIST (vom Englischen Time-resolved Angiography With Interleaved Stochastic Trajectories) Sequenz. Die TWIST Sequenz ist eine 3D+t Messung, die in einem inspiratorischen Atemanhalt gemessen wird und das komplette 3D Volumen der Lunge zu verschiedenen Zeitpunkten aufnimmt. Es wird eine ungefähre Messzeit von ca. 35 s verwendet, von der ca. 5 s zu Beginn ohne Kontrastmittel und anschließend 30 s mit Kontrastmittel gemessen werden sollten, um zu garantieren, dass die komplette Boluspassage vorhanden ist. Um ein optimales Signal des Kontrastmittels zu erhalten, wird ein relativ großer Flipwinkel zwischen 30–40° verwendet, der eine stärkere T1-Gewichtung des Bildes erzeugt und damit eine gute Unterscheidung zwischen gut und schlecht perfundierten Bereichen ermöglicht. Der Nachteil des großen Flipwinkels ist jedoch, dass dieser zu einem geringen Kontrast-zu-Rausch-Verhältnis zwischen Lungenparenchym und der Thoraxwand führt, da das Signal der Lunge aufgrund dessen langer T1-Zeit, weiter geschwächt wird. [37, 40]

Für eine quantitative Auswertung des DCE-MRT ist es wichtig, dass die zeitliche Auflösung 1,5 s oder weniger beträgt, damit diese kleiner als die mittlere Transitzeit des Kontrastbolus innerhalb der Lunge (ca. 3–4 s) ist. [44]. In Abbildung 2.7 ist der zeitliche Verlauf eines beispielhaften mittleren Lungensignals innerhalb einer DCE-MRT Messung dargestellt. Man erkennt, dass eine Gesamtmesszeit von ca. 30–35 s notwendig ist, um die komplette Boluspassage abzubilden. Nach ca. 5 s kommt das Kontrastmittel in der Lunge an, was am steilen Anstieg der Kurve zu diesem Zeitpunkt zu erkennen ist. Neben dem Hauptberg ergibt sich ein zweiter Berg durch die Rezirkulation des Kontrastmittels. Außerdem wird die Wichtigkeit der hohen zeitlichen Auflösung von unter 1,5 s deutlich, um die Kontrastmitteldynamik richtig darzustellen.

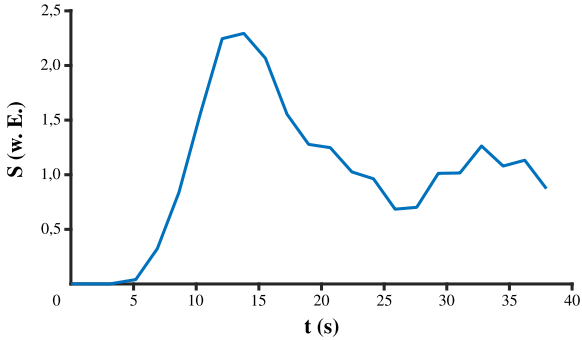


Abbildung 2.7: Zeitlicher Verlauf des mittleren Lungensignals einer beispielhaften DCE-MRT Messung.

2.4.4 Auswertung des Perfusions-MRT

Für die Auswertung von Perfusions-MRT Messungen gibt es je nach verwendeter MRT-Technik unterschiedliche Methoden. Die am weitesten verbreitete Methode ist die manuelle qualitative Bewertung einer sogenannten Maximumintensitätsprojektion (MIP) des Subtraktionsbildes durch RadiologInnen. Hierfür wird für jeden Voxel des Subtraktionsbildes jeweils der maximale Wert entlang der Zeitachse gewählt, wodurch aus dem 3D+t Bild ein 3D Bild wird. [40] Um die Vergleichbarkeit zwischen den Bewertungen verschiedener RadiologInnen zu garantieren, wurde ein semiquantitatives Scoring-System entwickelt, das eine lappenweise Evaluation von Perfusionsdefekten ermöglicht. [45, 46] Obwohl eine gute Intraobserver-Variabilität (Abweichung der Entscheidung bei Bewertung durch gleiche Person zu verschiedenen Zeitpunkten) und Interobserver-Variabilität (Abweichung der Entscheidung bei Bewertung durch zwei verschiedene Personen) für den Perfusions-Score gezeigt werden konnte [45], besteht das wesentliche Problem in der groben Einteilung des Scores. Außerdem ist eine gute Erfahrung der RadiologInnen in der Bewertung von DCE-MRT Bildern der Lunge notwendig, was aufgrund der Seltenheit des Lungen-MRT im klinischen Alltag nicht selbstverständlich ist. [40] Das Problem der groben Einteilung des Perfusions-MRT Score kann durch die Verwendung von quantitativen Ansätzen behoben werden.

Bei den Signal-Oszillations-Methoden kommt hier beispielsweise eine lineare Methode der kleinsten Quadrate zum Einsatz, die aus den Messungen voxelweise Karten für Perfusion und Ventilation generiert [42, 47, 48]. Im Gegensatz dazu werden beim F_2-MRT quantitative $T2^*$-Karten [49] und beim O_2-MRT quantitative T1-Karten erzeugt [50]. Die wesentlichen Nachteile dieser zuvor genannten quantitativen Ansätze ist, dass diese oft nur wenige Schichten und nicht die komplette Lunge abbilden oder die quantitativen Werte oft für die ganze Lunge bestimmt werden. Im Gegensatz hierzu ist die voxelweise Perfusionsquantifizierung basierend auf den Prinzipien der Tracer-Kinetik bzw. die Indikatorverdünnungsmethode [51] ein etabliertes und validiertes Verfahren in vielen Organen. Hierbei werden mithilfe eines Entfaltungsansatzes quantitative Karten für den pulmonalen Blutfluss (PBF), das pulmonale Blutvolumen (PBV) und die mittlere Transitzeit (MTT) berechnet. Die Methode wurde bereits mehrfach erfolgreich für die Perfusionsquantifizierung im Gehirn [52], aber insbesondere auch in der Lunge [53–57] verwendet. Die Perfusionsquantifizierung in dieser Arbeit beruht auf dieser Methode und wird nochmals detailliert in Kapitel 5 beschrieben.

2.5 Maschinelles Lernen

Das maschinelle Lernen beschäftigt sich mit Algorithmen, die von Daten lernen [58]. Das Buch von Goodfellow et al. [58] gilt als eines der Standardwerke für die Beschreibung der Grundlagen des Deep Learning, aber auch des maschinellen Lernens. Die nachfolgenden Erklärungen zu ML und zu neuronalen Netzen beruhen daher auf diesem Werk, geben jedoch nur einen Überblick über die relevanten Aspekte für diese Arbeit. Für ein tiefergreifendes Verständnis wird auf das Werk von Goodfellow et al. [58] oder weitere Grundsatzwerke verwiesen.

Im Wesentlichen beruht das maschinelle Lernen auf vier Grundbausteinen: Der zu lösenden Aufgabe, der Erfahrung, der Performanz und der Lösung. Die möglichen Aufgaben sind sehr vielfältig. Klassische Beispiele sind die Klassifikation, zu der auch die Segmentierung zählt, die Regression, aber auch die Datengenerierung. Um die Güte einer Lösung zu bewerten, muss die Performanz gemessen werden.

Zur Messung der Performanz wird eine problemspezifische Optimierungsfunktion verwendet. Soll diese Funktion maximiert werden, spricht man von einer Fitnessfunktion F, bei einer Minimierung spricht man dagegen von einer Verlustfunktion (engl. Loss-Function) L. Beim Sammeln von Erfahrung unterscheidet man zwischen zwei wesentlichen Lernverfahren. Beim unüberwachten Lernen hat der Algorithmus keinerlei Information über die gesuchte Lösung, sondern muss sich diese Lösung selbst erarbeiten. Hierzu wird meist die Wahrscheinlichkeitsverteilung der Daten erlernt, die dann für die eigentlichen Aufgaben wie beispielsweise Datengenerierung oder Klassifikation verwendet werden. Im Gegensatz dazu wird beim überwachten Lernen dem Algorithmus/Modell die optimale Lösung des Problems mit übergeben. Das Modell erhält also neben den Eingangsdaten auch die zu erwarteten Ausgabedaten. Es gibt jedoch auch Mischformen wie das teilüberwachte Lernen, wo die Lösungen nur für Teilschritte der Aufgabe zur Verfügung stehen. In allen Fällen wird das Lernen über ein Optimierungsverfahren gesteuert.

Für die Zusammenstellung eines ML-Algorithmus benötigt man einen Datensatz, eine Optimierungsfunktion, ein Optimierungsverfahren und ein Modell. Der Datensatz und die Optimierungsfunktion sollten hierbei das zu lösende Problem bestmöglich abbilden. Die verwendeten Optimierungsverfahren sind oft eng mit dem gewählten Modell verbunden. Häufig verwendete Varianten sind der stochastische Gradientenabstieg, bei dem der Gradient der Loss-Funktion für einen Teildatensatz berechnet wird, da die Berechnung des Gradienten für den kompletten Datensatz zu teuer wäre. Passt man nun die Lösung entlang dieses Gradienten an, so bewegt man sich in Richtung eines lokalen oder globalen Optima der Optimierungsfunktion. Andere weitverbreitete Optimierungsverfahren sind die heuristischen und metaheuristischen Optimierungsverfahren (siehe auch Abschnitt 2.5.2). Dabei sind die Modelle des maschinellen Lernens äußerst vielfältig. Nachfolgend werden drei typische Modellklassen des maschinellen Lernens näher vorgestellt, die eine wesentliche Rolle in dieser Arbeit spielen: Neuronale Netze, populationsbasierte Metaheuristiken und Markov Random Fields.

2.5.1 Neuronale Netze

Ein neuronales Netz (NN) besteht aus mindestens drei Schichten; einer Eingabeschicht, einer oder mehreren versteckten Schichten und einer Ausgabeschicht, wobei jede Schicht durch eine Funktion $f^{(n)}$ beschrieben wird und aus einem oder mehreren Neuronen besteht, die miteinander verbunden sind. Sind diese Verbindungen nur vorwärts gerichtet, also sind Neuronen einer Schicht nur mit den Neuronen der folgenden Schicht verbunden, so spricht man von einem vorwärts gerichteten neuronalen Netz, auch Multilayer-Perzeptron (MLP) genannt. Das Ziel eines neuronalen Netzes ist die Approximation einer Funktion $f^*(\mathbf{y};\theta)$, die einen Eingangswert $\mathbf{y}$ auf einen Ausgangswert $\mathbf{x}$ abbildet. Hierbei ist θ die Menge aller Gewichtsmatrizen $\mathbf{W}$ und Bias-Vektoren $\mathbf{b}$. Jedes Neuron besitzt einen Gewichtsvektor $\mathbf{w}$, der die Eingangssignale der Neuronen der vorherigen Schicht gewichtet, und einen Bias-Wert b, der zusätzlich dazu addiert wird. Damit berechnet sich die lineare Transformation des Eingangssignals mit

$$\mathbf{z} = \mathbf{y}^\top \mathbf{w} + b \; . \tag{2.1}$$

Um auch nichtlineare Zusammenhänge abbilden zu können wird das linear transformierte Signal nun noch mit einer Aktivierungsfunktion g weiterverarbeitet. Der Ausgabewert $\mathbf{a}$ des Neurons berechnet sich dann mit

$$\mathbf{a} = g(\mathbf{z}) \tag{2.2}$$

und die Funktion der Schicht mit

$$f^{(n)}(\mathbf{y};\mathbf{W}_{:,\mathbf{n}};\mathbf{b}_\mathbf{n}) = g\left(\mathbf{W}_{:,n}^\top \, \mathbf{y} + \mathbf{b}_n\right) \; . \tag{2.3}$$

Je nach Aufbau des Netzes sind diese Funktionen nun zu einer oder mehreren Kette(n) verbunden und bilden dann gemeinsam die globale Funktion $f^*(\mathbf{y};\theta)$.

Für die Aktivierungsfunktionen kommen je nach Position des Neurons und Anwendung unterschiedliche Funktionen zum Einsatz. Für verdeckte Schichten wird meist eine Rectified Linear Unit (ReLU) Funktion verwendet, die definiert ist als

$$g(\mathbf{z}) = max(0; \mathbf{y}) \ . \tag{2.4}$$

Im Gegensatz hierzu wird für die Ausgabeschicht meist eine Sigmoid-Funktion oder eine Softmax-Funktion verwendet. Die Sigmoid Funktion ist definiert als

$$g(\mathbf{z}) = \frac{1}{1 + \exp(-\mathbf{x})} \tag{2.5}$$

und eignet sich besonders für eine binäre Klassifikation, da es die Eingabewerte auf das Intervall $[0, 1]$ abbildet. Die Softmax-Funktion wird hingegen verwendet, wenn es mehr als zwei mögliche Klassen gibt. Der Ausgabewert der Softmax-Funktion für eine Klasse *i* ist gegeben mit

$$softmax(\mathbf{z})_i = \frac{\exp(z_i)}{\sum_j \exp(z_j)} \tag{2.6}$$

und garantiert, dass sich die Summe der Ausgabewerte aller Klassen auf 1 summiert.

Faltende Neuronale Netze

Faltende neuronale Netze (CNNs, vom Englischen Convolutional Neural Networks) sind ein Spezialfall von neuronalen Netzen, die insbesondere bei mehrdimensionalen Daten wie Bildern zum Einsatz kommen. Wie der Name schon sagt, ist der wesentliche Bestandteil eines CNN die Faltungsoperation. Die Faltung zweier Funktionen ist definiert als

$$h(t) = (u * v)(t) = \int u(x)v(t-x)\,dx \ , \tag{2.7}$$

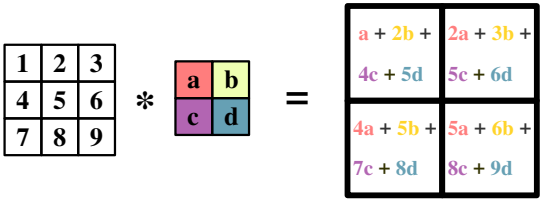


Abbildung 2.8: Visualisierung der Faltungsoperation eines Bildes mit einem Kernel. Das Ergebnis der Faltung ist eine alternative Version des ursprünglichen Bildes.

wobei $*$ der Faltungsoperator ist. Für den Anwendungsfall mit einem zweidimensionalen Bild $\mathbf{Y}$, das aus den Pixeln $s_{i,j}$ mit Reihenindex i und Spaltenindex j besteht, die nur ganze Zahlen annehmen können, lässt sich eine diskrete Faltungsoperation definieren. Die diskrete Faltung eines Bildes $\mathbf{Y}$ mit einem Kernel $\mathbf{K}$ lässt sich schreiben als

$$H(i,j) = (\mathbf{K} * \mathbf{Y})(i,j) = \sum_m \sum_n (\mathbf{Y}(i-m, j-n)\mathbf{K}(m,n)) \ . \tag{2.8}$$

Bei CNNs wird die diskrete Faltungsoperation eingesetzt, um ein Bild mit einem Kernel mit beliebigen Werten zu falten. Das Ergebnis dieser Operation ist eine alternative, transformierte Version des ursprünglichen Bildes. Die Faltung eines Bildes mit einem Kernel ist in Abbildung 2.8 dargestellt.

Wenn man verschiedene Kernel auf dasselbe Bild anwendet, lassen sich hiermit verschiedene Repräsentationen des Bildes generieren, die in einem CNN meist als einzelne Channel abgespeichert werden. Somit wird aus einem Grauwertbild der Form $H \times W$ ein Tensor der Form $H \times W \times C$. Neben der Faltung gibt es noch zwei weitere wichtige Bausteine eines CNN. Genauso wie bei den normalen neuronalen Netzen liefert eine Faltung nur eine affine Transformation. Um eine Nichtlinearität zu integrieren, gibt es auch hier Aktivierungsfunktionen, die auf die Ergebnisse der Faltung angewendet werden. Hier kommen die gleichen klassischen Aktivierungsfunktionen (z.B. ReLU, Sigmoid oder Softmax) zum Einsatz, wie bei den neuronalen Netzen auch. Den zweiten wichtigen Baustein

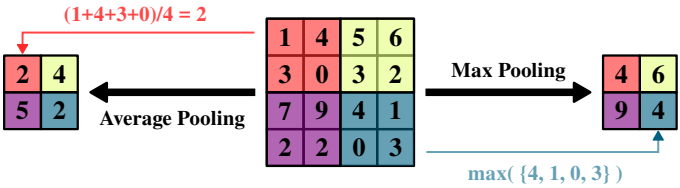


Abbildung 2.9: Darstellung der Average- und Max-Pooling Operation für Stride=2. Das Ergebnis des Pooling führt zu einer räumlich verkleinerten Darstellung des Bildes.

bilden die sogenannten Pooling-Schichten. Beim Pooling werden die Werte mehrerer Pixel innerhalb eines Channels zusammengefasst. Hierdurch lässt sich eine räumliche Reduktion der Bildgröße erreichen, indem Bildbereiche zusammengefasst werden. Das Zusammenfassen geschieht meist durch die Berechnung eines Mittelwertes (Average-Pooling) oder durch die Berechnung eines Maximalwertes (Max-Pooling) innerhalb eines Fensters, das mit einer Verschiebung (engl. Stride) über das Bild geschoben wird. Die Pooling-Operation ist in Abbildung 2.9 für das Average- und Max-Pooling beispielhaft dargestellt. Neben der Verkleinerung der räumlichen Dimensionen besteht der zweite wesentliche Vorteil des Poolings darin, dass die Repräsentationen unveränderlich gegenüber leichten Verschiebungen des Bildes werden. In CNNs werden diese drei Komponenten meist so kombiniert, dass die räumliche Größe der Tensoren mit tieferen Schichten abnimmt, während die Anzahl der Channel zunimmt.

Für die Optimierung der Gewichte wird bei CNNs meist das stochastische Gradientenabstiegsverfahren verwendet. Ausgehend vom Ergebnis des Modells und insbesondere des berechneten Loss-Wertes wird eine Fehlerrückführung (engl. Backpropagation) durchgeführt. Hierbei wird der berechnete Fehler (Loss) bis an den Beginn des Netzes zurückgeführt. Damit kann der individuelle Beitrag einer jeden Schicht zu diesem Fehler berechnet und basierend darauf deren Gewichte angepasst werden. Durch die Anpassung der Gewichte soll beim nächsten Durchlauf das Ergebnis verbessert werden. Während des Trainings wird der Fehler meist für Teilgruppen (engl. Minibatches) des Datensatzes berechnet und damit

dann der Gradient bestimmt. Dies ermöglicht, das auch große Datensätze verarbeitet werden können. Für die Stabilisierung des Trainings eines neuronalen Netzes können verschiedene Techniken verwendet werden. Dabei sind die Batch-Normalisierung und das Dropout-Verfahren zwei der am häufigsten verwendeten Methoden. Bei der Batch-Normalisierung werden die Daten innerhalb eines Minibatches normalisiert, sodass der Mittelwert innerhalb des Minibatches null und die Standardabweichung eins beträgt. Dies kann zu einem schnelleren und stabileren Training führen. Dagegen ist das Dropout-Verfahren eine Regularisierungstechnik, bei der einzelne Neuronen, Channel oder Schichten während einzelner Epochen deaktiviert werden. Dadurch kann eine zu starke Anpassung an die Trainingsdaten vermieden werden.

2.5.2 Populationsbasierte Metaheuristiken

Populationsbasierte Metaheuristiken sind von der Natur inspirierte Suchstrategien, die komplexe Optimierungsprobleme lösen können. Sie lassen sich in zwei wesentliche Gruppen unterteilen, die evolutionären Algorithmen (EA) und die Schwarmintelligenz-Algorithmen (SIA). [59, 60] Beide Gruppen vereint ein ähnliches grundsätzliches Vorgehen, welches in Abbildung 2.10 dargestellt ist. Zunächst wird eine Population bestehend aus Individuen initialisiert. Jedes Individuum stellt dabei eine Lösung des Optimierungsproblems dar.

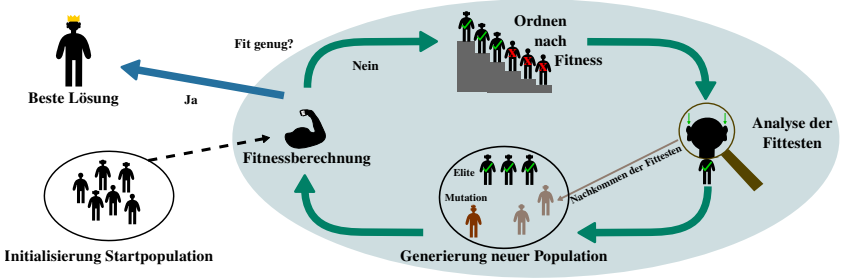


Abbildung 2.10: Darstellung der einzelnen Schritte von populationsbasierten Metaheuristiken.

Die Güte der Lösung wird mithilfe einer Fitnessfunktion gemessen, die für jedes Problem spezifisch definiert werden muss. In jeder Generation/Epoche wird die Fitness jedes Individuums berechnet. Anschließend findet eine Anpassung der Population statt, bei der neue Individuen basierend auf der aktuellen Population generiert werden. Meist gibt es hier elitäre Strategien, die garantieren, dass der Großteil der neuen Individuen basierend auf den fittesten Individuen der aktuellen Generation erzeugt wird. Um jedoch die Vielfalt der Population zu garantieren, werden oft auch neue Individuen unabhängig der besten Lösungen, z. B. mithilfe von Mutationsprozessen, generiert. Die Population der neuen Generation wird dann aus einer Mischung aus den besten Individuen der alten Population und den neu generierten Individuen zusammengestellt. Im Anschluss wird erneut die Fitness der Population berechnet und der Prozess wird so lange wiederholt, bis ein Abbruchkriterium erfüllt ist. [59, 60]

Der wesentliche Unterschied zwischen Schwarmintelligenz-Algorithmen und evolutionären Algorithmen ist, dass sich bei evolutionären Algorithmen die einzelnen Individuen nicht beeinflussen, sondern eine Durchmischung rein durch eine Reproduktion und Kombination von Individuen entsteht. Im Gegensatz dazu gibt es bei den SIAs eine lokale Interaktion zwischen einzelnen Individuen, indem das Verhalten einzelner Individuen durch das Verhalten anderer Individuen in der Umgebung beeinflusst wird. [59] Dieser Unterschied für die Bildung einer neuen Generation ist in Abbildung 2.11 dargestellt. Bei evolutionären Algorithmen wird die neue Generation gebildet, indem sich die fittesten Individuen fortpflanzen und zusätzlich weitere Individuen durch Mutation gebildet werden. In Abbildung 2.11 (a) sind Rot und Blau die fittesten Individuen, die ihre Gene in die nächste Generation weitergeben. Aus ihnen wird das blau-rot-gestreifte und das lila Individuum gebildet. Das grüne Individuum dagegen mutiert und bildet einen grün-schwarzen Gradienten aus. Die übrigen Lösungen aus der aktuellen Generation werden verworfen und durch die drei neu gebildeten Individuen ersetzt, damit die Populationsgröße konstant bleibt. Bei den Schwarmintelligenz-Algorithmen in Abbildung 2.11 (b) entsteht eine neue Generation, indem lokale Interaktionen zwischen den Individuen stattfinden. Der dunkelblaue Fisch sendet an die anderen

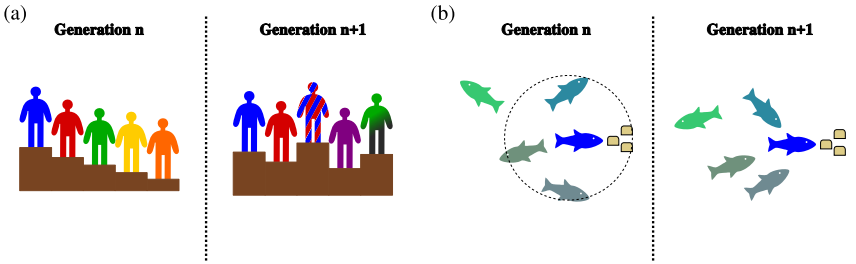


Abbildung 2.11: Darstellung der Generierungsprinzipien neuer Individuen für (a) Evolutionäre Algorithmen (b) Schwarmintelligenz-Algorithmen.

Fische in einem lokalen Umkreis ein Signal, dass er bei der Futtersuche erfolgreich war. Aufgrund des empfangenen Signals passen alle Fische in dem lokalen Umkreis ihre Schwimmrichtung in der neuen Generation an. Der grüne Fisch, der sich nicht im lokalen Umkreis des blauen Fisches befindet, empfängt das Signal des dunkelblauen Fisches nicht und ändert daher seine Schwimmrichtung nicht in Richtung der Futterquelle.

Nachfolgend werden die Schwarmintelligenz-Algorithmen detailliert behandelt, da die evolutionären Algorithmen in dieser Arbeit keine Rolle spielen. Für weiterführende Informationen zum Thema der evolutionären Algorithmen wird auf entsprechende Werke verwiesen [60, 61]. Die Vielfalt an Schwarmintelligenz-Algorithmen in der Literatur ist extrem groß. [62] Zu den bekanntesten Vertretern zählen die Partikel-Schwarm-Optimierung (PSO) [63], die Ameisenkolonie Optimierung (ACO) [64] und der Kuckuck-Suchalgorithmus (CS) [65].
Die Partikel-Schwarm-Optimierung beschreibt die Bewegung von Partikeln im Raum. Hierbei besitzt jeder Partikel immer das Wissen über die beste selbst erreichte Lösung, sowie die lokale (in der eigenen Gruppe) bzw. global beste Lösung. In jedem Zeitschritt wird die Geschwindigkeit und Beschleunigung jedes Partikels in Richtung der eigenen und der lokal/global besten Lösung verändert. [63]
Bei der Ameisenkolonie Optimierung wird das Suchverhalten von Ameisenkolonien nachgestellt. Jede Lösung des Problems wird durch einen Punkt im Raum dargestellt, der über Kanten mit anderen Punkten verbunden ist. Die Ameisen bewegen sich in diesem Lösungsraum entlang der Kanten und schütten auf dem

Weg zu einer Lösung ein Pheromon aus. Die Stärke des Pheromons wird mithilfe eines Pheromon-Modells bestimmt, welches die Lösungsgüte mithilfe der Optimierungsfunktion bestimmt und für bessere Lösungen einen stärkeren Pheromongehalt zuweist. Pfade mit einer stärkeren Pheromonkonzentration werden dann wahrscheinlicher von Ameisen angesteuert, wodurch diese eher Punkte mit hoher Lösungsgüte ansteuern werden. [64]
Der Kuckuck-Suchalgorithmus stellt das Verhalten des Kuckucks nach, der seine eigenen Eier in Nester fremder Vögel legt, damit diese die Brut und Aufzucht übernehmen. Beim CS-Algorithmus stellen die Eier die möglichen Lösungen dar. Zunächst werden zufällig Nester von Brutvögeln mit Eiern generiert. Anschließend wählt ein Kuckuck ein zufälliges Nest aus und legt dort sein Ei, welches einer zufällig generierten Lösung entspricht, in das Nest, sofern die Güte des eigenen Eis größer als die Güte des fremden Eis ist. Der Brutvogel kann mit einer zuvor definierten Wahrscheinlichkeit das fremde Ei entdecken. Wenn dies geschieht, wird das Nest verlassen und durch ein zufällig neues Nest mit Ei ersetzt. [65]

In den letzten Jahren hat die Elefantenherden Optimierung (EHO), ein weiterer Schwarmintelligenz-Algorithmus, jedoch viel Aufmerksamkeit generiert, da er in verschieden Benchmark-Tests den anderen SIAs überlegen war [66]. Aus diesem Grund wurde die EHO für die Lösung von Optimierungsproblemen in dieser Arbeit ausgewählt und soll im folgenden detailliert erläutert werden.

Elefantenherden Optimierung

Bei der Elefantenherden Optimierung wird das Wanderverhalten von Elefantenherden in der Savanne imitiert. Jede Elefantenherde wird hierbei von einer erfahrenen Elefantenkuh, der Matriarchin, geleitet. Die Matriarchin hat über Jahre hinweg Wissen über die besten Futter- und Wasserquellen gesammelt und führt ihre Herde zu diesen Quellen. Jeder Elefant einer Herde hat eine Position, die einer Lösung im Lösungsraum des Optimierungsproblems entspricht. Mit jeder Iteration verändert die Herde ihre Position, angeführt von der Matriarchin. Zusätzlich verlassen junge Elefantenbullen die Herden und werden durch neue Elefantenkälber ersetzt.

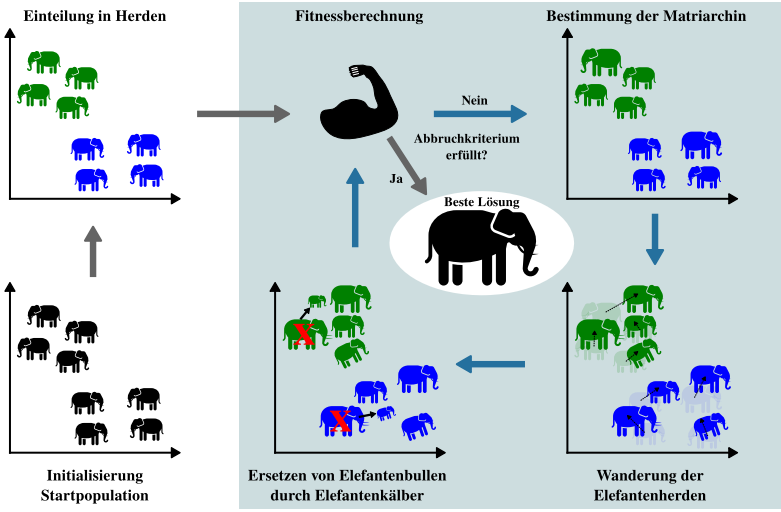


Abbildung 2.12: Darstellung der einzelnen Schritte des EHO-Algorithmus

Der EHO wurde erstmals von Wang et al. [67] vorgestellt. Mittlerweile wurden verschiedene Ergänzungen und Alternativen für Einzelschritte des EHO vorgeschlagen, die sehr gut im Review von Li et al. [66] zusammengefasst sind. Dieses Werk dient auch als Grundlage für die nachfolgende detaillierte Beschreibung des EHO und seiner Varianten.

Die einzelnen Schritte des EHO Algorithmus sind in Abbildung 2.12 dargestellt. Zunächst wird eine Population von Elefanten der Größe N initialisiert, die dann anschließend in einzelne Herden aufgeteilt wird. Eine Herde $c_i \in \mathcal{C}$ besteht dabei aus einer festen Anzahl N_c von Elefanten. Anschließend wird die Fitness eines jeden Elefanten in der Population berechnet. Die Fitnessberechnung basiert auf einer problemspezifischen Fitnessfunktion F. In jeder Herde wird die Population absteigend nach Fitness sortiert. Der fitteste Elefant in jeder Herde ist die Matriarchin der Herde. Anschließend findet die Bewegungsphase statt, in der die Position eines jeden Elefanten angepasst wird. Wie die Position angepasst wird, unterscheidet sich dabei nach der Rolle des Elefanten in der Herde. Die Matriarchin und alle Mitglieder der Herde ändern ihre Position, der schwächste Elefant

in der Herde dagegen wird durch einen neuen Elefanten ersetzt. Anschließend wird erneut die Fitness der Elefantenpopulation berechnet und es findet wieder eine Aktualisierung der Elefantenpositionen statt. Dieser Prozess wird so lange wiederholt, bis ein Abbruchkriterium erreicht ist. Der EHO zeichnet sich durch ein schnelles Konvergenzverhalten aus, das jedoch die Gefahr birgt in lokalen Optima hängenzubleiben. Dieses Problem kann jedoch durch verschiedene Adaptionen behoben werden. [66]

In der ursprünglichen Version von Wang et al. [67] werden die neuen Positionen mit Gleichungen (2.9) bis (2.12) berechnet. Die mittlere Position der gesamten Herde $\mathbf{m}^{(t)}_{center,c_i}$ wird berechnet durch

$$\mathbf{m}^{(t)}_{center,c_i} = \frac{1}{n_i} \sum_{d=1}^{D} \sum_{j=1}^{N_c} \mathbf{m}^{(t)}_{c_i,j} \,, \tag{2.9}$$

wobei D der Dimension des Suchraums und $\mathbf{m}^{(t)}_{c_i,j}$ der Position eines Elefanten $j \in [1, N_c]$ zum Zeitpunkt t entspricht. Die neue Position der Matriarchin $\mathbf{m}^{(t+1)}_{best,c_i}$ berechnet sich basierend auf der mittleren Position der gesamten Herde und einem Gewichtungsfaktor $\beta \in [0,1]$:

$$\mathbf{m}^{(t+1)}_{best,c_i} = \beta \mathbf{m}^{(t)}_{center,c_i} \,. \tag{2.10}$$

Die neue Position der anderen Herdenmitglieder $\mathbf{m}^{(t+1)}_{c_i,j}$ wird basierend auf der Position der Matriarchin, einem Gewichtungsfaktor $\gamma \in [0,1]$ und einem zufälligen Wert $r \sim \mathcal{U}(0,1)$ der Uniformverteilung $\mathcal{U}$ mit Grenzen $(0,1)$ berechnet

$$\mathbf{m}^{(t+1)}_{c_i,j} = \mathbf{m}^{(t)}_{c_i,j} + \gamma \left(\mathbf{m}^{(t)}_{best,c_i} - \mathbf{m}^{(t)}_{c_i,j} \right) r \,. \tag{2.11}$$

Das schwächste Mitglied einer Herde wird durch einen neuen Elefanten ersetzt, dessen Position sich mit folgender Formel berechnet:

$$\mathbf{m}^{(t+1)}_{worst,c_i} = \mathbf{m}_{min} + (\mathbf{m}_{max} - \mathbf{m}_{min} + \mathbf{1})\, r \,. \tag{2.12}$$

Hierbei entsprechen $\mathbf{m}_{min}$ und $\mathbf{m}_{max}$ jeweils den globalen Minima und Maxima des D-dimensionalen Lösungsraums und $\mathbf{1}$ ist ein Einsvektor.

Duan et al. haben mehrere Anpassungen des EHO vorgeschlagen, um Schwachstellen des EHO bezüglich der Balance zwischen Exploration und Exploitation zu beheben. [68] Hierzu wurden sowohl alternative Formeln zur Berechnung der neuen Position von Elefanten, als auch eine Kombination mit einem gradientenbasierten Optimierer vorgeschlagen und getestet. Die Position der Matriarchin ändert sich in dieser abgewandelten Version abhängig von der aktuellen Generation:

$$\mathbf{m}_{best,c_i}^{(t+1)} = \begin{cases} \mathbf{m}_{best,c_i}^{(t)} + C(\sigma) \ , & \frac{t}{t_{\max}} < 0.5 \\ \mathbf{m}_{a,c_i}^{(t)} + 2(r-0.5)(\mathbf{m}_{b,c_i}^{(t)} - \mathbf{m}_{c,c_i}^{(t)}) \ , & \frac{t}{t_{\max}} \geq 0.5 \end{cases} . \tag{2.13}$$

Bis zur Hälfte der maximalen Generationen wird die neue Position aus der Summe der aktuellen Position und einer zufälligen Nummer $C(\sigma)$, basierend auf einer Standard Cauchy-Verteilung mit Standardabweichung σ berechnet. Später berechnet sich die aktuelle Position der Matriarchin anhand der Positionen $\mathbf{m}_{a,c_i}$, $\mathbf{m}_{b,c_i}$ und $\mathbf{m}_{c,c_i}$ von drei zufällig gewählten Herdenmitgliedern a, b und c. Die Position der neuen Elefantenkälber wird berechnet mit

$$\mathbf{m}_{worst,c_i}^{(t+1)} = \mathbf{m}_{worst,c_i}^{(t)} + \boldsymbol{\delta}_m r_1 + f_1 \exp^{-\frac{2t}{t_{\max}}} \ , \tag{2.14}$$

wobei $r_1 \sim \mathcal{U}(0,1)$ und $f_1 \sim \mathcal{U}(0,1)$ zufällige, uniforme Werte sind und $\boldsymbol{\delta}_m = 0.1(\mathbf{m}_{max} - \mathbf{m}_{min})$. Für alle anderen Herdenmitglieder berechnet sich die neue Position entweder nur basierend auf der Position der Matriarchin, oder basierend aus einer Kombination der Position der Matriarchin der eigenen Herde und der Position der global besten Matriarchin $\mathbf{m}_{Gbest}^{(t)}$. Welche Variante gewählt wird, wird über einen Schwellwert $PSR \in [0,1]$ gesteuert.

$$\mathbf{m}_{c_i,j}^{(t+1)} = \begin{cases} \mathbf{m}_{c_i,j}^{(t)} + r_2 \left(\dfrac{\mathbf{m}_{best,c_i}^{(t)} + \mathbf{m}_{Gbest}^{(t)}}{2} - \mathbf{m}_{c_i,j}^{(t)} \right) \\ \quad + r_3 \left(\dfrac{\mathbf{m}_{best,c_i}^{(t)} - \mathbf{m}_{Gbest}^{(t)}}{2} - \mathbf{m}_{c_i,j}^{(t)} \right) , & r < PSR \\ \mathbf{m}_{c_i,j}^{(t)} + \gamma r_4 \left(\mathbf{m}_{best,c_i}^{(t)} - \mathbf{m}_{c_i,j}^{(t)} \right) , & r \geq PSR \end{cases} \tag{2.15}$$

Hierbei ist $\gamma \in [0,1]$ ein konstanter Gewichtungsfaktor und r_1, r_2, r_3 und r_4 zufällige uniforme Zahlen aus $\mathcal{U}(0,1)$.

Die gradientenbasierte Optimierung findet nach der Berechnung der neuen Positionen statt und passt die neuen Positionen aller Elefanten unabhängig von ihrer Herde nochmals basierend auf einer Schrittweite $\Delta\mathbf{m}$ an.

$$\Delta\mathbf{m} = \frac{\left(\mathbf{m}_{Gbest}^{(t)} - \mathbf{m}_{k_1}^{(t)} \right) + \boldsymbol{\delta}}{2} \tag{2.16}$$

mit

$$\boldsymbol{\delta} = 2\, r \left(\left| \frac{\mathbf{m}_{k1}^{(t)} + \mathbf{m}_{k2}^{(t)} + \mathbf{m}_{k3}^{(t)} + \mathbf{m}_{k4}^{(t)}}{4} - \mathbf{m}_{n}^{(t)} \right| \right) , \tag{2.17}$$

Hierbei sind $k_1 - k_4$ zufällige ganze Zahlen im Bereich $[0, N]$ und r eine zufällige uniforme Zahl $r \sim \mathcal{U}(0,1)$. Die neue Position $\mathbf{m}_n^{(t+1)}$ von Elefant n berechnet sich mit

$$\mathbf{m}_n^{(t+1)} = \mathbf{r}_a \cdot \left[\mathbf{r}_b \cdot \mathbf{M1}_n^{(t)} + (\mathbf{1} - \mathbf{r}_b) \cdot \mathbf{M2}_n^{(t)} \right] + (\mathbf{1} - \mathbf{r}_a) \cdot \mathbf{M3}_n^{(t)} , \tag{2.18}$$

wobei $\mathbf{r}_a$ und $\mathbf{r}_b$ Vektoren mit zufälligen uniformen Einträgen im Bereich $[0,1]$ sind. Die Teilterme $\mathbf{M1}_n^{(t)} - \mathbf{M3}_n^{(t)}$ sind definiert durch die Gleichungen (2.19) bis (2.21):

$$\mathbf{M1}_n^{(t)} = \mathbf{m}_n^{(t)} - r\, \rho_1 \cdot \frac{2\Delta\mathbf{m} \cdot \mathbf{m}_n^{(t)}}{\left(\mathbf{y}_{u_n}^{(t)} - \mathbf{y}_{v_n}^{(t)} + \varepsilon \right)} + r\, \rho_2 \, (\mathbf{m}_{Gbest}^{(t)} - \mathbf{m}_n^{(t)}) \tag{2.19}$$

$$\mathbf{M2}_n^{(t)} = \mathbf{m}_{Gbest}^{(t)} - r\,\rho_1 \cdot \frac{2\Delta\mathbf{m}\cdot\mathbf{m}_n^{(t)}}{\left(\mathbf{y}_{u_n}^{(t)} - \mathbf{y}_{v_n}^{(t)} + \varepsilon\right)} + r\,\rho_2\,\left(\mathbf{m}_{k_1}^{(t)} - \mathbf{m}_{k_2}^{(t)}\right) \tag{2.20}$$

$$\mathbf{M3}_n^{(t)} = \mathbf{m}_n^{(t)} - \rho_1\left(\mathbf{M2}_n^{(t)} - \mathbf{M1}_n^{(t)}\right) \tag{2.21}$$

mit

$$\rho_1 = 2\,r\,\gamma - \gamma \tag{2.22}$$

$$\rho_2 = 2\,r\,\gamma - \gamma \tag{2.23}$$

$$\mathbf{y}_{u_n}^{(t)} = \mathbf{y}_n^{(t)} + \Delta\mathbf{m} \tag{2.24}$$

$$\mathbf{y}_{v_n}^{(t)} = \mathbf{y}_n^{(t)} - \Delta\mathbf{m} \tag{2.25}$$

$$\mathbf{y}_n^{(t)} = \frac{\mathbf{m}_n^{(t)} + \mathbf{z}_n^{(t+1)}}{2} \tag{2.26}$$

$$\mathbf{z}_n^{(t+1)} = \mathbf{m}_n^{(t)} - r\,\frac{2\Delta\mathbf{m}\cdot\mathbf{m}_n^{(t)}}{\left(\mathbf{m}_{Gworst}^{(t)} - \mathbf{m}_{Gbest}^{(t)} + \varepsilon\right)}\,. \tag{2.27}$$

Hierbei ist ε ein kleiner konstanter Gewichtungsfaktor im Bereich 0,0–0,1. Ein zusätzlicher Schritt im gradientenbasierten EHO von Duan et al. [68] ist ein lokaler Ausbruch Operator (LEO, vom Englischen local escaping operator), der es ermöglicht aus lokalen Optima zu entkommen. LEO wird mit einer Wahrscheinlichkeit von 50 % nach dem erfolgten Gradientenupdate aus Gleichung (2.18) angewendet. Die neue Position nach LEO-Update $\mathbf{m}_{LEO}^{(t+1)}$ ist gegeben durch

$$\begin{aligned}\mathbf{m}_{LEO}^{(t+1)} = \mathbf{m}_{ref}^{(t+1)} &+ f_2\left(u_1\,\mathbf{m}_{Gbest}^{(t+1)} - u_2\,\mathbf{m}_k^{(t)}\right)\\ &+ f_3\,\rho_1\left[u_3\left(\mathbf{M2}_n^{(t)} - \mathbf{M1}_n^{(t)}\right) + u_2\left(\mathbf{m}_{k_1}^{(t)} - \mathbf{m}_{k_2}^{(t)}\right)\right]\end{aligned} \tag{2.28}$$

mit

$$\mathbf{m}_{ref}^{(t+1)} = \begin{cases}\mathbf{m}_n^{(t+1)}, & r < 0.5\\ \mathbf{m}_{Gbest}^{(t+1)}, & r \geq 0.5\end{cases} \tag{2.29}$$

und der zufälligen uniformen, reellen Zahl $f_2 \sim \mathcal{U}(-1,1)$ sowie einer zufälligen reellen Zahl f_3 aus der Standardnormalverteilung. Für $u_1 - u_3$ gilt

$$u_1 = \begin{cases} 2r, & r_5 < 0.5 \\ 1, & r_5 \geq 0.5 \end{cases} \tag{2.30}$$

$$u_2 = u_3 = \begin{cases} r, & r_5 < 0.5 \\ 1, & r_5 \geq 0.5 \end{cases} \tag{2.31}$$

und $\mathbf{m}_k^{(t)}$ ist definiert mit

$$\mathbf{m}_k^{(t)} = \begin{cases} \mathbf{m}_r^{(t)}, & r_6 < 0.5 \\ \mathbf{m}_{min} + r\left(\mathbf{m}_{max} - \mathbf{m}_{min}\right), & r_6 \geq 0.5\ , \end{cases} \tag{2.32}$$

wobei r_5 und r_6 zufällige reelle Werte aus der Uniformverteilung $\mathcal{U}(0,1)$ sind und $\mathbf{m}_r^{(t)}$ die Position eines zufällig ausgewählten Elefanten der Population ist.

2.5.3 Markov Random Fields

Ein sogenanntes Markov Random Field (MRF) ist ein ungerichteter Graph bestehend aus zufälligen Variablen (Knoten des Graphen), die über Kanten verbunden sind. Die einzelnen Knoten sind dabei basierend auf einem Nachbarschaftssystem mit anderen Knoten verbunden, die gemeinsam eine sogenannte Clique bilden. Dabei erfüllt das MRF die Markov Bedingungen, dass die Wahrscheinlichkeit für einen Wert eines Knotens nur von den dessen Nachbarschaft abhängt. [69–71] In dieser Arbeit werden verborgene Markov-Felder (HMRF, vom Englischen Hidden Markov Random Field) für die Segmentierung verwendet, daher werden Details zu HMRFs anhand dieses Anwendungsfalls erläutert.

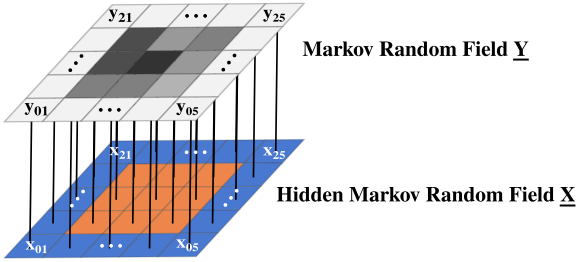


Abbildung 2.13: Veranschaulichung der Definition eines Bildes und dessen Segmentierung als Markov Random Field und Hidden Markov Random Field.

Hidden Markov Random Fields

Ein Bild ist definiert durch ein Bildraster $\mathcal{S}$ sowie ein Nachbarschaftssystem $\mathcal{N}$. Das Bildraster besteht aus einzelnen Voxeln $s \in \mathcal{S}$, die jeweils eine Nachbarschafts-Clique $\mathcal{N}_s \in \mathcal{N}$ besitzen. Das Bild $\mathbf{y}$ kann als eine mögliche Realisierung eines Markov Random Fields $\mathbf{Y}$ betrachtet werden, bei dem $y_s \in \mathbf{y}$ der Intensität des Voxels s entspricht. Eine Segmentierung $\mathbf{x}$ des Bildes, die jedem Voxel ein Label (z.B. eine Klasse) zuweist, kann als eine mögliche Realisierung eines zweiten Markov Random Fields $\mathbf{X}$ betrachtet werden. In diesem Fall wird $\mathbf{X}$ als ein Hidden Markov Random Field von $\mathbf{Y}$ bezeichnet. Die hierbei zugewiesene Klasse des Voxels s wird mit $x_s \in \mathcal{L}$ bezeichnet, wobei $\mathcal{L}$ die Menge der möglichen Klassen im Bild ist. [71, 72] In Abbildung 2.13 ist dargestellt, wie ein Bild und dessen dazugehörige Segmentierung mithilfe der MRF-Theorie modelliert werden können.

Die nachfolgende Beschreibung der Modellierung des Segmentierungsproblems mittels MRFs basiert auf dem Werk von Zhang et al.[71]. Ein Bild lässt sich über ein finites Gaußsches Mischmodell (GMM) darstellen, in dem jede Klasse $l \in \mathcal{L}$ im Bild durch eine Gaußverteilung mit dem Parameterset $\theta_l = \{\mu_l, \sigma_l\}$ repräsentiert wird. Hierbei sind μ_l und σ_l jeweils der Mittelwert und die Standardabweichung

der Klasse l. Die A-posteriori-Verteilung $P(y_s|\theta_l)$ für Klasse l lässt sich nun definieren als

$$P(y_s|\theta_l) = \frac{1}{\sqrt{2\pi\sigma_l^2}} \exp\left[-\frac{(y_s-\mu_l)^2}{2\sigma_l^2}\right] . \tag{2.33}$$

Die Likelihood Energiefunktion $\mathbf{U_{ll}}(\mathbf{y}|\mathbf{x})$ ist gegeben durch

$$\mathbf{U_{ll}}(\mathbf{y}|\mathbf{x}) = \sum_{s\in S}\left(\frac{(y_s-\mu_{x_s})^2}{2\sigma_{x_s}^2} + \ln\sigma_{x_s}\right). \tag{2.34}$$

und beschreibt die Energie der Segmentierung $\mathbf{x}$, falls das Bild $\mathbf{y}$ gemessen wurde. Für die Stabilisierung der Berechnung kann ein A-Priori Wahrscheinlichkeitsterm mithilfe einer Energiefunktion $\mathbf{U_n}(\mathbf{X}=\mathbf{x})$ definiert werden, die alle Nachbarschafts-Cliquen $\mathcal{N}_s$ im Bild $\mathbf{Y}$ berücksichtigt:

$$\mathbf{U_n}(\mathbf{X}=\mathbf{x}) = \sum_{s\in S} E(x_s) \tag{2.35}$$

In dieser Gleichung entspricht $E(x_s)$ dem Nachbarschafts-Strafterm der Nachbarschaft $\mathcal{N}_s$. Ein häufig verwendeter Nachbarschafts-Strafterm basiert auf der Potts-Nachbarschaft [71, 73], die nur die Klassen der benachbarten Voxel berücksichtigt:

$$E_{\text{potts}}(x_s) = \alpha \sum_{t\in\mathcal{N}_s}(1-2\delta_{x_s x_t}) . \tag{2.36}$$

Dabei entspricht $\delta_{x_s x_t}$ dem Kronecker Delta zwischen x_s und x_t und α entspricht dem Gewicht des Nachbarschaftsterms. Ein alternativer Nachbarschafts-Strafterm wurde von Banerjee und Maji vorgeschlagen, der neben der Klassen der Nachbarn, auch die Eigenschaften der Gaußverteilungen der Nachbarklassen verwendet [74]:

$$E_{\text{ban}}(x_s) = \frac{\alpha_s}{2|\mathcal{N}_s|}\sum_{t\in\mathcal{N}_s}\left((\mu_{x_s}-\mu_{x_t})^2\left(\frac{1}{\sigma_{x_s}^2}+\frac{1}{\sigma_{x_t}^2}\right)-1\right) \tag{2.37}$$

Hierbei entspricht α_s einem voxelspezifischen Gewichtungsfaktor der definiert ist durch

$$\alpha_s = \begin{cases} 2\alpha, & \left(r^{(2)}_{s(1)} - r^{(2)}_{s(2)}\right) > T \quad \text{und} \quad \left(r^{(1)}_{s(1)} - r^{(1)}_{s(2)}\right) > T \\ \alpha, & \left(r^{(2)}_{s(1)} - r^{(2)}_{s(2)}\right) > T \quad \text{und} \quad \left(r^{(1)}_{s(1)} - r^{(1)}_{s(2)}\right) \leq T \\ \alpha, & \left(r^{(2)}_{s(1)} - r^{(2)}_{s(2)}\right) \leq T \quad \text{und} \quad \left(r^{(1)}_{s(1)} - r^{(1)}_{s(2)}\right) > T \\ 0, & \left(r^{(2)}_{s(1)} - r^{(2)}_{s(2)}\right) \leq T \quad \text{und} \quad \left(r^{(1)}_{s(1)} - r^{(1)}_{s(2)}\right) \leq T \end{cases}, \tag{2.38}$$

mit dem Schwellwert T und der relativen Anzahl der i häufigsten Nachbarn $r^{(k)}_{s(i)}$ im Nachbarschaftssystem k-ter Ordnung. Basierend auf den zuvor genannten Gleichungen lässt sich nun die optimale Segmentierung $\hat{\mathbf{x}}$ finden, indem man die Summe beider Energieterme in den Gleichungen (2.34) und (2.35) minimiert:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \left(\sum_{s \in S} \left(\frac{(y - \mu_{x_s})^2}{2\sigma^2_{x_s}} + \ln \sigma_{x_s} + E(x_s) \right) \right). \tag{2.39}$$

3 Konzept der informationsgetriebenen Bildanalyse

Das übergeordnete Ziel der Bildanalyse ist immer, die Eigenschaften eines Objektes, eines Materials oder einer zu untersuchenden Person zu messen, um dann basierend darauf weitere Schritte zu ergreifen. Hierzu wird zunächst ein Bild des Zielobjekts angefertigt, wozu verschiedene Bildgebungstechniken wie das CT oder MRT zum Einsatz kommen. Im Anschluss werden die aufgenommenen Bilddaten mit Methoden aus dem Bereich der Bildverarbeitung aufbereitet. Bei der nachfolgenden quantitativen Auswertung kommen meist mathematische Modelle, Algorithmen, Modelle des maschinellen Lernens (ML) oder Simulationen zum Einsatz. Die gemessenen Eigenschaften des Objektes dienen dann als Grundlage für weitere Schritte wie eine Materialoptimierung oder das Ausstellen einer Behandlungsempfehlung.

Die datenzentrierte Forschung stellt die Generierung, Verwaltung und Weiternutzung von Forschungsdaten in den Mittelpunkt [75]. Eine Weiterentwicklung dieses Konzeptes ist die informationsgetriebene Forschung, bei der Schnittstellen zwischen den beiden Konzepten der datengetriebenen und der theoriegetriebenen Forschung erzeugt werden. [76] Ein zentraler Punkt der informationsgetriebenen Forschung ist die Verknüpfung von anwendungsspezifischem und methodischem Wissen. Hierbei helfen Konzepte, die eine Verallgemeinerung der Zusammenhänge von Methoden und Daten definieren und diese Konzepte damit anwendungsübergreifend verwendet werden können. In dieser Arbeit wurde ein Konzept zur informationsgetriebenen Bildanalyse entwickelt, welches in Abbildung 3.1

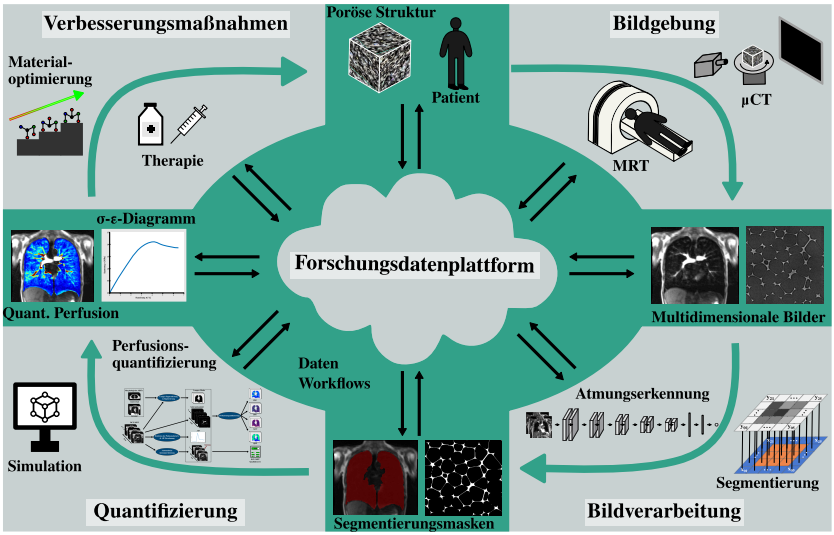


Abbildung 3.1: Darstellung des informationsgetriebenen Bildanalysezyklus. Bildgebungstechniken wie die Computertomographie oder Magnetresonanztomographie liefern mehrdimensionale Bilder von porösen Strukturen oder Patienten. Die Bilder lassen sich mit Bildverarbeitungsmethoden wie der Segmentierung oder Eventdetektion weiterverarbeiten, wodurch Segmentierungsmasken generiert oder besondere Ereignisse wie bspw. die Atmung detektiert werden können. Quantitative Auswerteverfahren wie die Simulation oder Algorithmen zur Perfusionsquantifizierung ermöglichen die Bestimmung von quantitativen Material- oder Perfusionswerten basierend auf den generierten Segmentierungen. Begründet durch die quantitativen Ergebnisse lassen sich anschließend Verbesserungsmaßnahmen wie eine Materialoptimierung oder die Verschreibung einer Therapie veranlassen. Alle Daten und Methoden werden mittels Daten-Workflows mit einer Forschungsdatenplattform verknüpft und synchronisiert.

dargestellt ist. Der dargestellte informationsgetriebene Bildanalysezyklus bildet alle Schritte der Bildanalyse ab. Dabei werden sowohl die erzeugten Daten als auch die genutzten Methoden innerhalb einer Forschungsdatenplattform geordnet abgespeichert und die Zusammenhänge der einzelnen Schritte festgehalten. Der Vorteil dieses Ansatzes ist, dass durch die strukturierte Ablage der Prozesse und Daten eine Weiterverwendung der Methoden und Daten ermöglicht und vereinfacht wird. Durch die Entwicklung von flexiblen und robusten Methoden

wird ein Werkzeugkasten für Aufgaben der Bildverarbeitung erzeugt, der sich auf unterschiedlichste heterogene Daten anwenden lässt. Nachfolgend wird detailliert auf die einzelnen Bestandteile dieses informationsgetriebenen Bildanalysezyklus eingegangen.

3.1 Bildgebung

Bei der Bildgebung werden unterschiedliche Bildgebungstechniken wie das CT oder MRT verwendet, um mehrdimensionale Bilder von Objekten oder Personen zu erstellen. Diese Bilder dienen als eine virtuelle Repräsentation des gemessenen Objektes und halten insbesondere auch den aktuellen Ist-Zustand fest. Der wesentliche Vorteil der bildgebenden Verfahren gegenüber anderen Messtechniken ist die zerstörungsfreie Messung des aktuellen Zustands. Neben den erzeugten Bildern sind auch die genauen Messparameter der verwendeten Methoden von zentraler Bedeutung. Diese Messparameter sollten als sogenannte Metadaten zusammen mit den Bildern dokumentiert werden, da diese häufig von Bedeutung für die nachfolgende Verarbeitung der Bilder sind. Typische Bildmetadaten sind die Bildauflösung, das Messdatum, aber auch die genauen Spezifikationen des Messgeräts.

3.2 Bildverarbeitung

Die Bildverarbeitung ist der erste wichtige Schritt in der Analyse von mehrdimensionalen Bilddaten. Hierbei kommen verschiedene Methoden wie Algorithmen und überwachte und unüberwachte ML-Modelle zum Einsatz. Die Ziele der Bildverarbeitung sind vielschichtig. Ein Anwendungsbereich ist die Aufbereitung der Daten, indem beispielsweise Normalisierungs- oder Filtertechniken auf die Bilder angewendet werden. Ein weiterer Anwendungsfall ist die Erhöhung der Bildauflösung, indem Methoden zur sogenannten Super-Resolution-Rekontruktion eingesetzt werden. Bei der Verarbeitung von zeitaufgelösten Daten kommen

Techniken zur Eventdetektion zum Einsatz, die besondere Ereignisse, wie bspw. Atmung, innerhalb einer Messung detektieren. Einer der wichtigsten Bereiche im Bereich der Bildverarbeitung ist allerdings die Segmentierung. Diese ermöglicht die Identifikation und Extraktion von Objekten innerhalb von Bildern. Hierdurch lassen sich unter anderem digitale Zwillinge von Materialien erstellen. Da die Segmentierung einen zentralen Aspekt dieser Arbeit darstellt, wird diese nachfolgend detailliert erläutert.

3.2.1 Segmentierung

Das Ziel der Segmentierung ist die Einteilung der Voxel eines Bildes in Gruppen mit gleicher Bedeutung. Eine gleiche Bedeutung kann hier die Zugehörigkeit zum gleichen Material, Organ oder Gewebe bedeuten, aber auch die Zugehörigkeit zum gleichen Objekt und ist daher immer abhängig vom jeweiligen Anwendungsfall. Im Wesentlichen lassen sich Segmentierungsmethoden in vier Hauptklassen einteilen: (i) schwellwertbasierte Verfahren, (ii) regionsbasierte Verfahren, (iii) modellbasierte Verfahren und (iv) Pixel-Klassifikations Techniken. [77]

Bei den schwellwertbasierten Verfahren wird ein Bild anhand seiner Intensitäten in einzelne Gruppen eingeteilt. Hierzu werden für n Gruppen $(n-1)$ Schwellwerte bestimmt, mithilfe derer das Bild dann in die einzelnen Gruppen eingeteilt werden kann. [78, 79] Die am weitesten verbreitete Methode ist das Thresholding mithilfe der Methode von Otsu. Hierbei wird ein Histogramm des Bildes angefertigt und anhand der Häufigkeit von Pixeln innerhalb jeder Klasse, sowie deren Mittelwert eine Optimierungsfunktion aufgestellt, die nur von den einzelnen Schwellwerten abhängt. Die Maximierung dieser Optimierungsfunktion liefert dann die optimalen $(n-1)$ Schwellwerte für eine Einteilung des Bildes in n Klassen. [79]

Zu der Gruppe der regionsbasierten Methoden gehört das Regionswachstumsverfahren (engl. region growing), bei dem zunächst ein oder mehrere Saatpunkte (engl. seed points) im Bild gewählt werden und ausgehend von diesen dann schrittweise benachbarte Voxel mit ähnlichen Eigenschaften gruppiert werden. Im Unterschied zum zuvor beschriebenen schwellwertbasierten Verfahren, werden

hier nur benachbarte Voxel der gleichen Gruppe zugeordnet. Daher muss für jedes erwartete Objekt einer Klasse ein eigener Saatpunkt gesetzt werden. [80, 81] Es gibt auch Methoden die schwellwertbasierte und regionsbasierte Methoden verbinden, um eine Segmentierung zu generieren [82].

Die dritte Gruppe der modellbasierten Verfahren nutzt meist Modelle des zu segmentierenden Objektes. Der Großteil dieser Verfahren verwendet dabei statistische Form-Modelle des Objektes. Das statistische Form-Modell wird aus einer Hauptkomponentenanalyse von der Menge einer großen Anzahl von Einzelsegmentierungen generiert, die zuvor in ein einheitliches Koordinatensystem transformiert werden. Für diese Transformation werden entweder charakteristische Punkte im Bild gesucht oder die Bilder werden mittels Registrierung, also dem Ausrichten der gleichen Objekte im Bild, in ein einheitliches Referenzsystem gebracht. Um eine Segmentierung zu generieren, wird zunächst das statistische Form-Modell mithilfe des Referenzsystems ins zu segmentierende Bild übertragen. Anschließend wird das Form-Modell elastisch verformt, bis eine optimale Ausrichtung, basierend auf einer Optimierungsfunktion, erreicht wurde. [83, 84]

Die letzte Kategorie der Pixel-Klassifikations Techniken enthält sowohl unüberwachte, als auch überwachte Lernverfahren. Der wesentliche Unterschied zwischen diesen beiden Kategorien besteht darin, dass für überwachte Lernverfahren eine große Menge an manuell annotierten Referenzdaten (GT, vom Englischen Groundtruth) notwendig ist, deren Erfassung sehr zeitaufwendig ist. Die unüberwachten Verfahren kommen dagegen ohne diese Referenzdaten aus. Zu den unüberwachten Segmentierungsmethoden gehören selbstorganisierende Karten (SOM, vom Englischen Self-Organizing Map) [85] und verborgene, zufällige Markov-Felder (HMRF, vom Englischen Hidden Markov Random Field) [71, 86]. Die überwachten Lernverfahren bestehen hauptsächlich aus faltenden neuronalen Netzwerken (CNN, vom Englischen Convolutional Neural Network) [87–89]. Durch die Verbreitung moderner Systeme der künstlichen Intelligenz (KI) sowie von leistungsstarken Grafikkarten werden Pixel-Klassifikations Techniken basierend auf ML-Modellen mittlerweile sehr häufig für die Segmentierung verwendet.

Zur Bewertung einer Segmentierung können verschiedene Metriken zum Einsatz kommen, die fast immer eine zweite Referenzsegmentierung als Groundtruth benötigen. Es gibt verschiedene Kategorien von Segmentierungsmetriken, wobei die volumenbasierten und randbasierten Metriken hier am häufigsten zum Einsatz kommen. Eine weitverbreitete volumenbasierte Metrik ist der DICE-Score (DSC) [90]. Dieser berechnet sich aus dem Quotienten der zweifachen Schnittmenge zweier Segmentierungen mit der Summe der Elemente beider Segmentierungen. Der DICE-Score zwischen zwei Segmentierungen A und B ergibt sich zu

$$\mathrm{DSC}(A,B) = \frac{2|A \cap B|}{|A| + |B|} \,. \tag{3.1}$$

Eine randbasierte Metrik ist der mittlere Oberflächenabstand (MASD, vom Englischen Mean Average Surface Distance) [91], der sich berechnet als

$$\mathrm{MASD}(A,B) = \frac{1}{2}\left(\frac{\sum_{a \in A} d(a,B)}{|A|} + \frac{\sum_{b \in B} d(b,A)}{|B|} \right) \tag{3.2}$$

mit

$$d(a,B) = \min_{b \in B} d(a,b) \,, \tag{3.3}$$

wobei a und b jeweils Punkte auf dem Rand von A bzw. B sind und $d()$ die euklidische Distanz repräsentiert. Eine zweite randbasierte Metrik ist die normalisierte Oberflächendistanz (NSD, vom Englischen Normalized Surface Distance) [92] die sich berechnet mit

$$\mathrm{NSD}(A,B)^{(\tau)} = \frac{|S_A \cap \mathcal{B}_B^{(\tau)}| + |S_B \cap \mathcal{B}_A^{(\tau)}|}{|S_A| + |S_B|} \,. \tag{3.4}$$

Hierbei sind S_A und S_B jeweils die Kanten von A bzw. B und $\mathcal{B}_A^{(\tau)}$ und $\mathcal{B}_B^{(\tau)}$ sind die Randregionen von A bzw. B mit einem Toleranzparameter τ, der die Toleranzregion definiert, innerhalb der abweichende Punkte nicht als Falsch definiert werden.

3.3 Quantifizierung

Nachdem die gemessenen Bilddaten mithilfe einer der beschriebenen Bildverarbeitungsmethoden aufbereitet oder verarbeitet wurden, können diese nun für eine quantitative Analyse verwendet werden. Für die Quantifizierung kommen hier meist mathematische und algorithmische Methoden, Simulationen oder ML-Modelle zum Einsatz. Ein häufiger Quantifizierungsschritt ist die Messung von geometrischen Eigenschaften. In den Materialwissenschaften werden Eigenschaften wie die Dichte, die Porengröße oder die Stegbreite bestimmt [93, 94], während in der Medizin beispielsweise die Größe und Form von Muskeln, Knochen oder dem Gehirn gemessen wird [95]. Außerdem lassen sich die digitalen Repräsentationen für Simulationen verwenden, in denen dann z. B. mechanische oder thermische Eigenschaften analysiert werden können [96, 97]. In Kombination mit funktionalen Messungen können innerhalb der segmentierten Objekte auch funktionale Eigenschaften wie die Diffusion oder Perfusion bestimmt werden [57, 98].

3.4 Verbesserungsmaßnahmen

Die Quantifizierungen aus dem vorangegangenen Schritt bilden die Grundlage für weitere Verbesserungsmaßnahmen. Hierbei spielen teilweise neben den aktuell gemessenen quantitativen Parametern auch Parameter von vergangenen Zeitpunkten oder von Vergleichsmaterialien eine Rolle. Im medizinischen Bereich werden basierend auf den gemessenen Parametern beispielsweise Medikamente verschrieben oder weitere Untersuchungen angeordnet. Außerdem können die gemessenen Parameter auch in Vergleichsstudien mit anderen PatientInnen verwendet werden und hiermit eine Erkrankung oder deren Behandlung näher erforscht werden. Im Bereich der Materialwissenschaften dienen die Ergebnisse aus quantitativen Messungen und Simulationen meist als Grundlage für Materialanpassungen oder die Erforschung von Materialklassen. Aus den gemessenen quantitativen Werten lassen sich beispielsweise Spannungs-Dehnungs-Diagramme

oder Struktur-Eigenschafts-Beziehung herleiten. Diese dienen dann als Grundlage für weitere Experimente und Simulationen, in denen Materialien und Strukturen mit angepassten Eigenschaften untersucht werden.

3.5 Forschungsdatenplattform und Daten-Workflows

Die strukturierte Ablage und Verwaltung von Forschungsdaten und Methoden spielt eine zentrale Rolle im informationsgetriebenen Bildanalysezyklus. Hierzu wird eine Forschungsdatenplattform wie Kadi4Mat [99] verwendet. Ein zentraler Aspekt des Konzeptes sind die automatisierten Daten-Workflows, die eine einheitliche und aufwandsarme Synchronisation der Daten ermöglichen. Dabei dokumentieren die Daten-Workflows automatisch alle Bestandteile des Bildanalysezyklus, beginnend bei den gemessenen Bilddaten. Die Bilder werden gemeinsam mit ihren Metadaten innerhalb der Forschungsdatenplattform abgespeichert und jeweils mit ihren Messgeräten verknüpft. Bei den Bildverarbeitungsmethoden werden nicht nur deren Ergebnisse, sondern auch die Methoden selbst in die Forschungsdatenplattform hochgeladen, was eine Weiterverwendung der Methode in der Zukunft ermöglicht. Hierzu werden ebenso etwaige Parameterstudien zusammen mit den Methoden automatisch dokumentiert, um redundante Entwicklungsschritte zu vermeiden. Das gleiche Vorgehen wird auch für die Dokumentation der Methoden und Ergebnisse der Quantifizierung verwendet. Die quantitativen Messwerte werden mit ihren jeweiligen Methoden verknüpft und automatisch in die Forschungsdatenplattform hochgeladen. Dort können die Ergebnisse dann auch von weiteren Projektmitgliedern betrachtet und heruntergeladen werden, um die Daten oder Methoden weiterzuverwenden oder aber Handlungsempfehlungen für weitere Schritte zu ergreifen. Auch diese Handlungsempfehlungen sollen als Teil des informationsgetriebenen Bildanalysezyklus innerhalb der Forschungsdatenplattform dokumentiert und mit den zugrundeliegenden quantitativen Messwerten verknüpft werden, da diese möglicherweise wichtige Informationen für weitere Untersuchungen sein könnten. Die Entwicklung von Daten-Workflows zur strukturierten

Erfassung und Speicherung der generierten Forschungsdaten ist das zentrale Thema von Kapitel 6 und bildet den inneren Teil des informationsgetriebenen Bildanalysezyklus ab.

3.6 Flexible und robuste Methoden

Die Verfügbarkeit von flexiblen und robusten Methoden ist ein wichtiger Katalysator, um große Mengen an heterogenen Daten zu erzeugen. Zusätzlich führen diese dazu, dass die Methoden für ein breites Spektrum von Anwendungen eingesetzt werden können. Somit sind diese ein wichtiger Baustein innerhalb des informationsgetriebenen Bildanalysezyklus.

3.6.1 Flexible Methoden

Eine flexible Methode besitzt die Eigenschaft, gegenüber dem Einsatzgebiet der Methode offen zu sein sowie mit einfachen Mitteln an neue Einsatzbereiche anpassbar zu sein. Ist eine Methode flexibel einsetzbar, so ist deren Anwendungsbereich im Vergleich zu unflexiblen Methoden größer. Schon während der Entwicklungsphase einer Methode kann Wert auf die Flexibilität gelegt werden, aber auch eine retrospektive Erhöhung der Flexibilität ist möglich. Nicht jede Methode ist jedoch gleichermaßen geeignet, um eine hohe Flexibilität zu erhalten. Insbesondere Methoden, die physikalische Modelle beinhalten, sind nur schwierig an andere Anwendungsfälle anpassbar, da das physikalische Modell oft eng mit den verwendeten Daten bzw. dem Anwendungsfall verknüpft ist. Auch überwachte ML-Methoden sind häufig nicht sehr flexibel, da diese immer auf problemspezifischen gelabelten Daten trainiert werden. Daher ist ein trainiertes ML-Modell nicht direkt für andere Daten, die sich von den Trainingsdaten unterscheiden, anwendbar. Um es in einem solchen Fall verwenden zu können, ist oft ein erneutes Training mit problemspezifischen gelabelten Daten des neuen Problems, die eventuell nicht verfügbar sind, nötig. Das Problem der Abhängigkeit von gelabelten Trainingsdaten besitzen unüberwachte ML-Modelle nicht und sind

daher flexibler als ihre überwachten Pendants. Auch sie müssen zwar auf neuen problemspezifischen Daten trainiert werden, jedoch werden hier keine gelabelten Groundtruth-Daten benötigt.

Die Flexibilität beschreibt jedoch nicht nur die Anpassbarkeit einer Methode an verschiedene Datenarten, sondern auch wie gut das Modell auf Sonderfälle innerhalb des eigenen Datensatzes angepasst werden kann. Diese Anpassbarkeit der Methode ist besonders für algorithmische Methoden stark gegeben, da hier ganz spezifische Regeln für Sonderfälle berücksichtigt werden können. Dies ist für überwachte ML-Methoden z. B. nicht möglich, da hier nur gehofft werden kann, dass das Modell diese Sonderfälle automatisch lernt, sollten diese im Trainingsdatensatz enthalten sein.

Kapitel 4 beschäftigt sich mit der Entwicklung von flexiblen Methoden zur Segmentierung von Objekten in unterschiedlichen Bildarten und heterogenen Datensätzen. Diese Segmentierungsmethoden gehören zur Kategorie der Bildverarbeitungsmethoden und können daher dem unteren rechten Quadranten des Bildanalysezyklus zugeordnet werden.

3.6.2 Robuste Methoden

Eine robuste Methode zeichnet aus, dass leichte Veränderungen oder Störungen innerhalb des Datensatzes kein Problem für die Methode darstellen und diese Daten trotzdem ausgewertet werden können. Die Methode sollte in solchen Fällen jedoch trotzdem vergleichbare Ergebnisse liefern. Veränderungen im Datensatz können beispielsweise durch veränderte Aufnahmebedingungen, wie die Verwendung eines unterschiedlichen Geräte-Modells, entstehen. Insbesondere in multizentrischen medizinischen Studien ist dies häufig der Fall. Eine robuste Methode sollte hier trotz der unterschiedlichen Aufnahmebedingungen vergleichbare Ergebnisse für alle Zentren liefern.

Die Robustheit von Methoden lässt sich auf unterschiedliche Weisen implementieren. Die wohl einfachste und auf alle Methoden anwendbare Möglichkeit stellt

die Normierung der Eingangsdaten auf einen einheitlichen Wertebereich dar. Für ML-Methoden wird eine Robustheit oft damit erreicht, dass das Modell mit einem möglichst heterogenen Datensatz trainiert wird, der alle zu erwarteten Fälle abdeckt. Die algorithmischen Verfahren sind besonders geeignet, um eine robuste Methode zu entwickeln, da man für unterschiedliche Messgeräte jeweils spezifische Auswerte-Einstellungen definieren kann. Die Robustheit einer Methode kann jedoch auch dadurch erhöht werden, indem gewisse Qualitätsmaßnahmen implementiert werden, die die Qualität der Eingangsdaten, aber auch die Qualität von Zwischenergebnissen bewerten. Diese Qualitätssicherungsmaßnahmen fördern zudem das Vertrauen in die Methode, insbesondere das von neuen AnwenderInnen.

In Kapitel 5 werden Methoden aus dem Bereich der Bildverarbeitung und der Quantifizierung entwickelt und angewendet. Schwerpunktmäßig geht es um die Entwicklung einer Methode zur robusten vollautomatischen Perfusionsquantifizierung, die eine Anwendung in heterogenen Kohorten wie bspw. multizentrischen Studien ermöglicht. Die Perfusionsquantifizierung lässt sich dem unteren linken Quadranten des Bildanalysezyklus zuordnen.

4 Bildverarbeitung: Flexible Methoden zur Segmentierung von multidimensionalen Bilddaten

Die Segmentierung von Objekten in Bildern ist für fast alle Anwendungen im Bereich der Bildverarbeitung ein zentraler Grundschritt, auf dem alle quantitativen Analysen basieren. Diese Abhängigkeit der nachfolgenden Analysen von der Segmentierung bedeutet, dass eine gute Qualität der Segmentierung essenziell ist. Aufgrund der Relevanz und häufigen Verwendung der Segmentierung haben sich über die Jahre zahlreiche Forschende mit diesem Thema beschäftigt. Der Großteil der bislang entwickelten Methoden ist jedoch, insbesondere im Bereich der ML-Modelle, sehr problemspezifisch und lässt sich daher häufig nicht auf andere Anwendungsgebiete transferieren. Dadurch werden meist ähnliche Entwicklungsschritte mehrfach für verschiedene Anwendungen wiederholt. Daher gibt es einen Bedarf an Segmentierungsmethoden, die einfach und ohne großen zusätzlichen Entwicklungsaufwand auf andere Bilddaten flexibel anwendbar sind.

Dieses Kapitel beschäftigt sich mit der Entwicklung von flexiblen Segmentierungsmethoden, die auf verschiedene Bildarten oder heterogene Datensätze angewendet werden können. Die Segmentierung befindet sich, wie in Abbildung 4.1 dargestellt, im unteren rechten Quadranten des in Kapitel 3 vorgestellten informationsgetriebenen Bildanalysezyklus und ist Teil der großen Kategorie der Bildverarbeitungs-Methoden. Für die Validierung der entwickelten Segmentierungsmethoden werden sowohl Daten aus dem Bereich der Materialwissenschaften, als auch medizinische MRT-Bilder verwendet. Zunächst wird ein kurzer Überblick über Segmentierungsmethoden in der Literatur gegeben. Anschließend



Abbildung 4.1: Einordnung der Segmentierung in den informationsgetriebenen Bildanalysezyklus.

werden die verwendeten Bilddatensätze genauer erläutert. Im Folgenden werden dann zwei verschiedene flexible Segmentierungsmethoden vorgestellt, die beide auf verborgenen, zufälligen Markov-Feldern beruhen. Diese werden dann jeweils auf die materialwissenschaftlichen und die medizinischen Bilddaten angewendet und damit validiert. Zusätzlich wird eine flexible, maßgeschneiderte Methode für die Lungensegmentierung entwickelt und getestet.

4.1 Stand der Forschung

Wie in Abschnitt 3.2.1 beschrieben, lassen sich Segmentierungsmethoden in vier Klassen einteilen, wobei sich die in dieser Arbeit untersuchten Methoden in die Klassen der Pixel-Klassifikations Techniken, sowie der regionsbasierten Verfahren einordnen lassen.

Eine flexible Segmentierungsmethode lässt sich im Wesentlichen auf zwei Arten erreichen. Entweder wurde die Methode so entwickelt, dass sie unabhängig von den Eingangsdaten Objekte verschiedener Art segmentieren kann. Oder aber die Methode besitzt Einstellungsmöglichkeiten, die an alternative Eingangsdaten leicht angepasst werden können. Überwachte CNN-Modelle sind die am häufigsten verwendeten Segmentierungsmodelle, da sie erprobt und leistungsstark sind

und über viele Anwendungsbereiche hinweg Segmentierungen auf dem höchsten technischen Stand produzieren. Diese Modelle basieren meist entweder auf einem ResNet-Rückgrat [89] oder verwenden eine U-Net Architektur [87, 88]. Insbesondere mit dem U-Net konnten in der Vergangenheit sehr gute Segmentierungsergebnisse für materialwissenschaftliche Bilder [100–104] und medizinische Bilder [87, 88, 105–108] erzielt werden. Der wesentliche Nachteil dieser überwachten CNN-Modelle ist jedoch, dass eine große Menge an annotierten Groundtruth-Daten benötigt wird, deren Generierung insbesondere bei dreidimensionalen Bildern sehr zeitaufwendig sein kann und daher nicht immer verfügbar ist. Ein weiterer Nachteil ist, dass die Modelle meist für den spezifischen Anwendungsfall trainiert wurden und daher auf unterschiedliche Bilddaten nur anwendbar sind, wenn die Modelle zuvor erneut mit angepassten Daten trainiert werden. Daher sind diese Modelle grundsätzlich nicht sehr flexibel, wenn man das nötige erneute Training in die Bewertung mit einbezieht.

Einen Sonderfall stellen sogenannte Foundation-Modelle wie das „Segment Anything“-Modell dar. Diese Modelle werden mit riesigen heterogenen Bildern verschiedenster Bereiche trainiert und sind daher in der Lage, ein breites Spektrum an Objekten in Bildern zu segmentieren. [109] Es existieren auch Spezial-Varianten dieses Modells, die mit disziplinspezifischen Daten nachtrainiert wurden und damit bessere Segmentierungen für Bilder dieser Disziplin erzeugen. Ein Beispiel ist „Segment Anything Medical“, das auf medizinischen Bildern nachtrainiert und damit für medizinische Anwendungen optimiert wurde. [110]. Daher sind Foundation-Modelle grundsätzlich relativ flexibel, besitzen jedoch den Nachteil, dass selbst das Nachtrainieren aufgrund der enormen Größe des ML-Modells viele Ressourcen beansprucht, die nicht immer verfügbar sind. Ebenso werden hier zwangsläufig erneut annotierte Bilddaten benötigt, was die Flexibilität dieser Modelle wiederum einschränkt.

Um das Problem von fehlenden annotierten Trainingsdaten zu kompensieren, können automatisch erzeugte Trainingsdaten verwendet werden. Bo et al. [103] nutzten hierfür einen KMeans-Clustering Algorithmus, um Segmentierungsmasken für die Bilder zu generieren. In der Arbeit von Griem et al. [94] wurden

künstliche Bilddaten mithilfe eines generativen ML-Modells basierend auf binären Bildern generiert, wodurch Bild-Groundtruth-Paare fürs Training verfügbar waren. Eine alternative Möglichkeit, die keinerlei annotierte Daten benötigt, sind unüberwachte Segmentierungsmethoden. Klassische Vertreter sind neben gängigen Cluster-Verfahren wie KMeans oder Otsu [79], selbstorganisierende Karten [85] oder verborgene zufällige Markov-Felder [71]. Zusätzlich zählen auch algorithmische Segmentierungen [111, 112] zu unüberwachten Segmentierungsverfahren.

Ortiz et al. [113] verwendeten SOMs für die unüberwachte Segmentierung des Gehirns in Gehirn-MRT Bildern. Aganj et al. berechneten jeweils die lokalen Massenzentren für einzelne Voxel und kombinierten diese dann iterativ, um daraus eine Segmentierung zu generieren [114]. In einer anderen Arbeit wurde Transferlearning und KMeans-Clustering für die Segmentierung von Strukturen und Defekten in Bildern von Stahloberflächen verwendet [115]. Die Kombination eines CNN mit einem Superpixel-Algorithmus wurde in der Arbeit von Kim et al. [116] zur Segmentierung von Strukturen in Bildern verschiedener Stahllegierungen vorgeschlagen.

Aber insbesondere Verfahren basierend auf HMRFs wurden vermehrt für die unüberwachte Segmentierung eingesetzt. Zhang et al. waren die Ersten, die HMRF für die Segmentierung in Gehirn-MRT Bildern einsetzten. Sie verwendeten für die Lösung des NP-harten Problems eine Kombination eines Expectation Maximization (EM) und Iterated Conditional Modes (ICM) Algorithmus, der jedoch sehr zeitintensiv ist. [71] Auch weitere Autoren verwendeten die HMRF-Methode in Kombination mit EM-ICM, um Zellen [117] oder ebenso das Gehirn [74, 118] zu segmentieren. Banerjee und Maji schlugen verschiedene Verteilungsfunktionen für die Klassen [118], sowie verschiedene Nachbarschaftsfunktionen [74] vor. In letzter Zeit wurde die alternative Verwendung von populationsbasierten Algorithmen für die Lösung des HMRF-Optimierungsproblems vorgeschlagen. Aufgrund des explorativen Vorgehens sind solche Algorithmen besonders geeignet, um das NP-harte, nicht lineare Problem mit vielen lokalen Minima zu lösen. [72, 119] In Guerrout et al. [72] wurde hierzu ein Kuckuck-Suchalgorithmus verwendet,

während Mahiou et al. [119] einen Elefantenherden-Optimierungs Algorithmus für die Lösung des HMRF-Problems verwendeten.

4.2 Datensätze

In dieser Arbeit wurden drei verschiedene Datensätze für die Validierung der entwickelten Segmentierungsmethoden verwendet. Die ersten zwei Datensätze sind Datensätze aus den Materialwissenschaften, während der dritte Datensatz aus dem medizinischen Bereich stammt.

Polyurethan Datensätze

Der erste Datensatz (*RealPUFoam*) bestand aus drei µCT-Aufnahmen einer Polyurethan (PU) Schaumstruktur. Die Bilder hatten eine Größe von $1\,300 \times 951 \times 960$ Voxeln mit einer räumlichen Auflösung von $2{,}75\,\mu m \times 2{,}75\,\mu m \times 2{,}75\,\mu m$. Der zweite Datensatz (*ArtPUFoam*) bestand aus 20 000 künstlich erzeugten 2D µCT-Bildern von PU-Schaumstrukturen mit einer Größe von 256×256 Pixeln und dazugehörigen Groundtruth Segmentierungen. Die Bilder wurden mithilfe eines mehrstufigen Workflows, wie in der Arbeit von Griem et al. [94] beschrieben, erzeugt. In diesem Workflow wurden digitale Zwillinge von einem initialen Datensatz von echten PU µCT-Aufnahmen generiert. Mithilfe des Pace3D Simulationsframeworks [120] wurden strukturelle Merkmale (engl. Features) extrahiert, die anschließend zur Generierung von ähnlichen, binären Schaumstrukturen verwendet wurden. Diese binären Schaumstrukturen wurden in Grauwertbilder umgewandelt und dann gemeinsam mit dem ursprünglichen binären Bild, welches der Segmentierung entspricht, für das Training eines Segmentierungs U-Net verwendet. Das trainierte U-Net wurde anschließend für die Segmentierung der Original µCT-Bilder verwendet, die dann wiederum für das Training eines generativen neuronalen Netzes (GAN) für die Generierung von binären Schaumstrukturen verwendet wurden. Abschließend wurde ein CycleGAN (Cycle-Consistent Adversarial Network) verwendet, das die vom GAN generierten binären Schaumstrukturen

in Grauwertbilder umwandelt. Die gemessenen strukturellen Eigenschaften der Original µCT-Bilder steuerten hierbei den Generierungsprozess. [94]

Lungen-MRT Datensatz

Der dritte Datensatz war ein Datensatz mit 524 T1-gewichteten Gradientenecho MRT-Bildern der Lunge, die in koronaler Ausrichtung aufgenommen wurden. Die Bilder entstammen der COSYCONET-Studie, die in Abschnitt 5.12.1 genau beschrieben wird. Es handelt sich um eine multizentrische, deutschlandweite Studie mit COPD-PatientInnen, die mit verschiedenen MRT-Scannern, aber mit einem einheitlichen MRT-Protokoll gemessen wurden. Aufgrund seiner Heterogenität ist dieser Datensatz besonders gut geeignet, um eine Segmentierungsmethode auf verschiedenen MRT-Scannern zu testen. Für drei der MRT-Messungen wurden jeweils manuelle Lungensegmentierungen von einer erfahrenen Radiologin erstellt, die als Goldstandard für die abschließende Validierung dienten.

4.3 EHOHMRF

Die erste entwickelte Methode kombiniert verborgene zufällige Markovfelder aus Abschnitt 2.5.3 mit der Elefantenherden-Optimierung (EHO) aus Abschnitt 2.5.2, um Objekte in Bildern zu segmentieren. Die Methode wird im Folgenden mit EHOHMRF abgekürzt.

4.3.1 Aufbau des EHOHMRF-Algorithmus

Beim EHOHMRF-Algorithmus werden mögliche Lösungen des HMRF-Problems durch die Position einzelner Elefanten repräsentiert. Hierbei besitzt jeder Elefant je ein Parameterset θ_l pro erwarteter Klasse im Bild, das von der zu modellierenden Verteilungsfunktion der Klasse abhängt. In dieser Arbeit wurden zwei

verschiedene Verteilungen betrachtet, eine Gaußsche Normalverteilung, beschrieben durch den Mittelwert μ_l und die Standardabweichung σ_l, und eine Stomped Normalverteilung, die durch den Mittelwert μ_l, die Standardabweichung σ_l und eine Plateaubreite k_l repräsentiert wird. Die Stomped Normalverteilung wurde von Banerjee und Maji [118] vorgeschlagen und stellt eine Normalverteilung dar, die in der Mitte ein Plateau besitzt, welches durch das Abschneiden der Spitze der Normalverteilung entsteht. Die Wahrscheinlichkeitsdichtefunktion der Stomped Normalverteilung wird beschrieben durch

$$f_l(x;\mu_l,\sigma_l,k_l) = \frac{1}{2(1-\Phi(k_l)+k_l\phi(k_l))}\frac{1}{\sigma_l}\phi(z) \tag{4.1}$$

mit

$$z = \begin{cases} k_l, & |\frac{x-\mu_l}{\sigma_l}| < k_l \\ \frac{x-\mu_l}{\sigma_l}, & |\frac{x-\mu_l}{\sigma_l}| \geq k_l \end{cases}, \tag{4.2}$$

wobei $\phi()$ und $\Phi()$ jeweils die Funktionen der Wahrscheinlichkeitsdichte und der Wahrscheinlichkeitsverteilung sind. [118]

Da sich die Verwendung von chaotischen Karten für die Berechnung der zufälligen Zahlen im EHO-Algorithmus als hilfreich herausgestellt hat [121], wurden diese auch in dieser Arbeit verwendet. Es wurde eine Gauss-Chaoskarte verwendet, die neue Werte R_{n+1} in Abhängigkeit des vorherigen Wertes R_n berechnet:

$$R_{n+1} = \frac{1}{R_n} - \left\lfloor \frac{1}{R_n} \right\rfloor . \tag{4.3}$$

Hierbei stellt $\lfloor\ \rfloor$ die Abrundungsfunktion dar und es gilt $R_n \in [0,1)$. Die chaotische Karte wurde ebenso für die Initialisierung der gesamten Elefantenpopulation verwendet, also der Generierung der Ausgangslösungen des HMRF-Problems. Für die Initialisierung musste die Herdenanzahl und die Größe der Elefantenherden festgelegt werden. Anschließend fand eine zufällige Einteilung in Herden statt, wobei jede Herde von einer Matriarchin angeführt wird, die jeweils einen der fittesten Elefanten in der Gesamtpopulation darstellt. Hierzu wurde zuvor für jeden Elefanten der Population die Fitness berechnet. Für die Fitnessberechnung

wurde die HMRF-Optimierungsfunktion, die in Gleichung (2.39) minimiert wird, verwendet, die je nach Untersuchung mit verschiedenen Energiefunktionen für die Verteilung und Nachbarschaftsarten berechnet wurde. Da die Fitness von der Definition her maximiert werden sollte und die HMRF-Optimierungsfunktion eine Minimierungsfunktion darstellt, wurde als Fitness der negative Wert der HMRF-Optimierungsfunktion gewählt.

Im Anschluss an die Fitnessberechnung wurden die Positionen der Elefanten jeweils geupdated. Hierzu wurden die Gleichungen (2.13) bis (2.15) der optimierten Herdenwanderung von Duan et al. [68] verwendet. Das ebenso von Duan et al. vorgeschlagene gradientenbasierte Positionsupdate wurde, je nach Anwendungsfall, an das Positionsupdate angeschlossen. Nach dem Positionsupdate und dem Austausch der Elefantenbullen durch Elefantenkälber, wurde erneut die Fitness jedes Elefanten innerhalb der Population berechnet. Falls das Abbruchkriterium (Anzahl der maximalen Epochen) nicht erreicht wurde, wurde nun erneut die Herdenwanderung simuliert und die Schritte so lange wiederholt, bis die maximale Anzahl an Epochen erreicht wurde. Die Position des Elefanten mit der höchsten Fitness wurde als optimale Lösung des Segmentierungsproblems gewählt und damit dann die Segmentierung berechnet. Für die Segmentierung wurde die Klassenzugehörigkeit mittels der normalen HMRF-Optimierungsfunktion berechnet, indem ein Voxel der Klasse zugeordnet wurde, deren HMRF-Funktionswert am geringsten ausfiel.

Da für den *ArtPUFoam*-Datensatz mit künstlich erzeugten PU-Schäumen auch Groundtruth-Segmentierungen vorlagen, wurden nachfolgende Parameterstudien alle auf diesem Datensatz durchgeführt, um eine Entscheidung für die optimalen Parameter zu treffen. Hierzu wurden jeweils die DICE-Scores (DSC) zwischen der wahren und der generierten Segmentierung berechnet. Zum Vergleich wurde zusätzlich eine Segmentierung mithilfe eines Otsu Schwellwertverfahrens generiert, die als Standardreferenzmethode diente. Aus allen berechneten DSC-Werten

Tabelle 4.1: Parameter des EHOHMRF-Algorithmus für alle durchgeführten Parameterstudien.

	Parameter	Wert
EHO	Herdengröße	10
	Herdenanzahl	10
	γ	0.5
	PSR	0,6
HMRF	α	1,0

wurde jeweils der Mittelwert und die Standardabweichung einer Methode berechnet. Zusätzlich wurden Signifikanztests mithilfe eines Wilcoxon-Vorzeichen-Rangtest durchgeführt. Für alle durchgeführten Parameterstudien des EHOHMRF-Algorithmus wurden die in Tabelle 4.1 aufgelisteten Parameter verwendet, sofern nicht anders erwähnt.

4.3.2 EHO mit und ohne Gradientenupdate

Zunächst wurde untersucht, ob die gradientenbasierte EHO-Variante besser oder schlechter für den Anwendungsfall der Segmentierung geeignet ist. Hierzu wurden für 100 zufällig ausgewählte Bilder des Datensatzes jeweils die Segmentierungen mit dem normalen EHO und dem gradientenbasierten EHO (GBEHO) berechnet und mit der Groundtruth-Segmentierung verglichen. Für beide Varianten wurde für das HMRF-Minimierungsproblem eine Gaußsche Normalverteilung für die Klassen und eine normale Potts-Nachbarschaft verwendet. Um die vermutlich erhöhten Konvergenzzeiten des normalen EHO zu kompensieren, wurden für den regulären EHO 100 Epochen gewählt, während für den GBEHO nur 25 Epochen verwendet wurden.

In Tabelle 4.2 sind die berechneten Mittelwerte und Standardabweichungen der DSC-Werte der 100 berechneten Segmentierungen jeweils für die Variante mit normaler und gradientenbasierter EHO dargestellt. Die Segmentierungen, die

Tabelle 4.2: Vergleich der DICE-Scores (Mittelwert ± Standardabweichung) für EHOHMRF mit normaler und gradientenbasierter EHO.

	EHO	**GBEHO**
DSC	0,9398 ± 0,0291	0,9244 ± 0,0461***

mit der normalen EHO-Variante generiert wurden, besaßen signifikant höhere mittlere DSC-Werte. Ein Grund für die besseren Segmentierungen der normalen EHO-Variante könnte in der besseren Exploration dieser Variante liegen, welche die Chance erhöht, das globale Optimum zu finden. Zwar wurde in der gradientenbasierten Variante der LEO-Term in Gleichung (2.28) eingeführt, um aus lokalen Minima auszubrechen, jedoch scheint dieser hier möglicherweise nicht ausreichend gut zu funktionieren.

In Abbildung 4.2 sind jeweils exemplarisch gute, mittlere und schlechte Segmentierungsbeispiele für die beiden EHO-Varianten dargestellt. Während für die gute und mittlere Segmentierung nur minimale Unterschiede zwischen beiden Varianten erkennbar sind, unterscheiden sich die Segmentierungsergebnisse für den schlechten Fall deutlich. Die GBEHO-Variante segmentierte die sehr dunklen Bereiche, die zwischen den Stegen und den Poren im Bild lagen, als Teil der PU-Klasse. Möglicherweise hätte die Einführung einer dritten Klasse die Segmentierungsergebnisse für dieses Bild verbessern können. Es fällt zusätzlich auf, dass in dem rechten Bild, die nicht richtig segmentierten feinen PU-Stege sehr dunkel sind und sich nicht so gut vom Porenbereich abheben. Diese Abgrenzung vom Porenbereich ist in den beiden Bildern der ersten zwei Spalten deutlich ersichtlicher.

4.3.3 Untersuchung verschiedener Verteilungsfunktionen

Die Energiefunktion der Verteilungsfunktion ist ein weiterer möglicher variierbarer Parameter. Die Verteilungsfunktion beschreibt die Verteilung, die für die

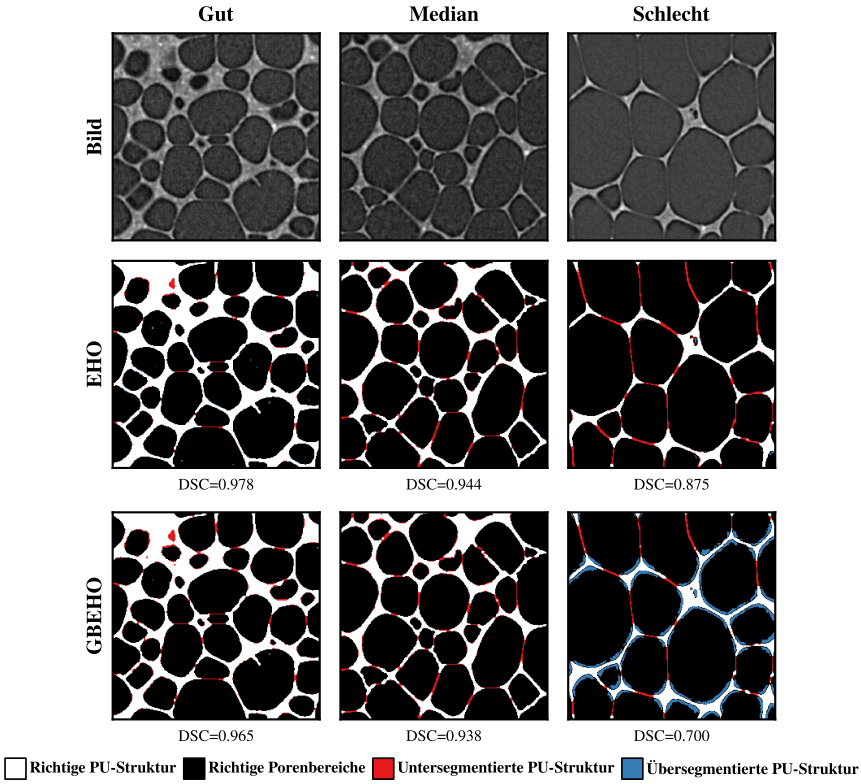


Abbildung 4.2: Beispielhafte Segmentierungsergebnisse für die normale und gradientenbasierte EHOHMRF Variante. Es ist jeweils eine gemeinsame gute Segmentierung (erste Spalte), die gemeinsame Median Segmentierung (zweite Spalte) und eine gemeinsame schlechte Segmentierung (dritte Spalte) dargestellt. Korrekt segmentierte PU-Strukuren sind in Weiß, korrekt segmentierte Porenbereiche in Schwarz, untersegmentierte PU-Strukturen in Rot und übersegmentierte PU-Strukturen in Blau dargestellt.

Modellierung der Klassen im Bild verwendet wird. Hierbei wurden drei verschiedene Varianten der Energiefunktion betrachtet. Die ersten beiden Varianten modellieren die Klassen entweder mit einer Gaußschen Normalverteilung oder einer Stompnorm Verteilung. In Mahiou et al. [119] wurde eine dritte Variante vorgeschlagen, die den klassischen Energieterm der Gaußschen Normalverteilung

Tabelle 4.3: Vergleich der DICE-Scores (Mittelwert ± Standardabweichung) für EHOHMRF mit verschiedenen Verteilungsfunktionen für die Klassenmodellierung.

	LogStd	**GMM**	**Stompnorm**
DSC	0,9326 ± 0,0331	0,9398 ± 0,0291	0,9450 ± 0,0308

auf den natürlichen Logarithmus der Standardabweichung einer Klasse vereinfacht. Die drei Varianten werden nachfolgend als GMM, Stompnorm und LogStd bezeichnet. Für alle drei Varianten wurden jeweils Modelle mit normaler Potts-Nachbarschaft und dem normalen EHO-Algorithmus für je 100 Epochen trainiert.

Die berechneten Mittelwerte und Standardabweichungen der DSC-Werte der 100 berechneten Segmentierungen sind für alle drei Varianten jeweils in Tabelle 4.3 dargestellt. Die erzeugten Segmentierungen aller drei Varianten besitzen ähnlich hohe mittlere DICE-Scores. Die vereinfachte Modellierung mit LogStd erzeugte dabei im Vergleich die schlechtesten Segmentierungen, während die Klassenmodellierung mittels der Stompnorm-Verteilung die besten Segmentierungsergebnisse erzeugte. In Abbildung 4.3 sind jeweils exemplarisch gute, mittlere und schlechte Segmentierungsbeispiele für die drei EHOHMRF Varianten dargestellt. Das Modell mit Stompnorm-Verteilung scheint die feinen Stege meist etwas besser zu segmentieren, als die anderen beiden Modelle. Jedoch zeigt sich, dass für die Stompnorm-Verteilung teilweise auch Voxel in der Mitte einer Pore als PU-Struktur klassifiziert wurden (blaue Punkte in schwarzen Bereichen). Solch eine fehlerhafte Segmentierung könnte jedoch sehr einfach durch ein Postprocessing behoben werden. Auch hier erkennt man erneut, dass alle Modelle vor allem Probleme mit Stegen hatten, die sich nicht gut vom Signal der Porenbereiche abhoben.

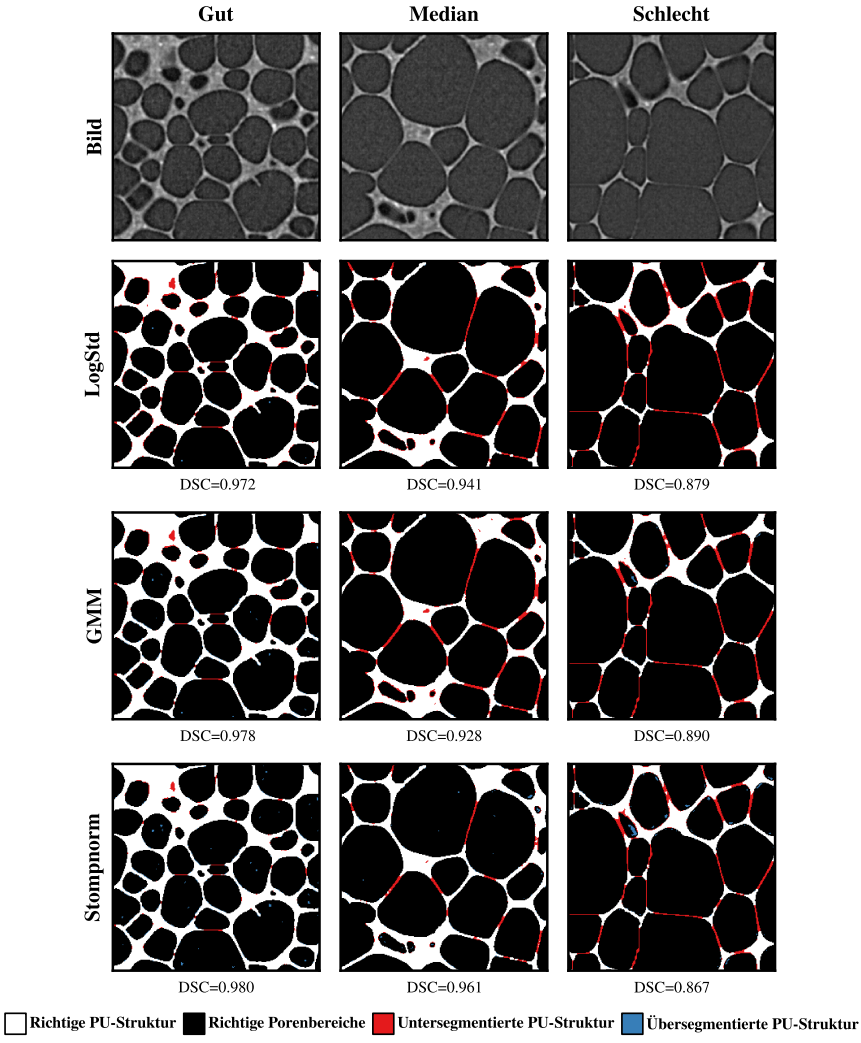


Abbildung 4.3: Beispielhafte Segmentierungsergebnisse für drei verschiedene HMRF Verteilungsfunktionen. Es ist jeweils eine gemeinsame gute Segmentierung (erste Spalte), die gemeinsame Median Segmentierung (zweite Spalte) und eine gemeinsame schlechte Segmentierung (dritte Spalte) dargestellt. Korrekt segmentierte PU-Strukuren sind in Weiß, korrekt segmentierte Porenbereiche in Schwarz, untersegmentierte PU-Strukturen in Rot und übersegmentierte PU-Strukturen in Blau dargestellt.

Tabelle 4.4: Vergleich der DICE-Scores (Mittelwert ± Standardabweichung) für EHOHMRF mit verschiedenen Nachbarschaftsarten mit und ohne voxelweise Berechnung des Gewichtungsfaktors.

	DSC	
	Potts	**Banerjee**
Einheitliches Gewicht	0,9395 ± 0,0211	0,2108 ± 0,3495
Voxelspezifisches Gewicht	0,7959 ± 0,0858	0,9487 ± 0,0168

4.3.4 Einfluss des Nachbarschaftsterms

Der zweite wesentliche Einflussfaktor in der HMRF-Optimierungsfunktion ist die gewählte Nachbarschaftsvariante. Zur Untersuchung dieses Einflusses wurden vier verschiedene Nachbarschaftsvarianten betrachtet: Die normale Potts-Nachbarschaft und die normale Banerjee-Nachbarschaft, die jeweils ein konstantes Nachbarschaftsgewicht besitzen und die gewichtete Potts-Nachbarschaft und die gewichtete Banerjee-Nachbarschaft, die jeweils ein voxelspezifisches Nachbarschaftsgewicht besitzen. Die Potts-Nachbarschaft wurde mit Gleichung (2.36) und die Banerjee-Nachbarschaft mit Gleichung (2.37) berechnet. Die voxelweise Gewichtung wurde für beide Varianten mit Gleichung (2.38) berechnet. Da die Berechnung der voxelweisen Nachbarschaftsgewichte sehr zeitintensiv ist und die Segmentierungszeit um einen Faktor von mehr als 50 erhöht, wurden in diesem Experiment nur 25 zufällige Bilder segmentiert. Für alle vier Varianten wurden jeweils Modelle mit dem Gaußschen Energieterm mit dem normalen EHO-Algorithmus verwendet.

Die berechneten DICE-Scores in Tabelle 4.4 zeigen, dass die voxelweise Berechnung des Nachbarschaftsgewichtes nur für die Banerjee-Nachbarschaft hilfreich war. Hier konnten die Segmentierungsergebnisse durch die spezifische Gewichtung deutlich verbessert werden. Dagegen scheint die spezifische Gewichtung bei der Potts-Nachbarschaft kontraproduktiv zu sein und führte hier zu deutlich schlechteren Segmentierungsergebnissen. Insgesamt waren die Segmentierungsergebnisse für die gewichtete Banerjee-Nachbarschaft etwas besser als für die

Tabelle 4.5: Vergleich der DICE-Scores (Mittelwert ± Standardabweichung) für EHOHMRF mit verschiedenen Nachbarschaftsarten und Gewichtungsfaktoren für den Nachbarschaftsterm.

	DSC		
α	**0,1**	**1,0**	**10,0**
Normal Potts	0,8911 ± 0,0880	0,9398 ± 0,0291	0,2846 ± 0,0887
Normal Banerjee	0,8867 ± 0,0955	0,8101 ± 0,0867	0,8402 ± 0,0949

normale Potts-Nachbarschaft. Die sehr hohen Standardabweichungen für die normale Banerjee-Nachbarschaft lassen auf eine deutliche Abweichung zwischen den Segmentierungsergebnissen schließen.

4.3.5 Gewichtung des Nachbarschaftsterms

Neben dem Nachbarschaftsterm selbst, ist auch die Gewichtung zwischen den beiden Energietermen für die Verteilung und die Nachbarschaft ein zentraler Faktor. Daher wurde eine Parameterstudie für die drei Gewichtungsfaktoren $\alpha = 0,1$, $\alpha = 1,0$ und $\alpha = 10,0$ durchgeführt. Es wurden sowohl Modelle mit normaler Potts Nachbarschaft und mit normaler Banerjee Nachbarschaft untersucht. Alle Modelle wurden mit dem Standard EHO für je 100 Epochen trainiert und es wurden so insgesamt pro Modell 100 Bilder segmentiert, wovon erneut der mittlere DICE-Score pro Modell und Nachbarschaftsgewicht berechnet wurde.

In Tabelle 4.5 sind die mittleren DICE-Scores für die drei verschiedenen Gewichte des Nachbarschaftsterms für die normale Potts- und normale Banerjee-Nachbarschaft dargestellt. Für die Potts-Nachbarschaft wurden die besten Segmentierungen für eine Gewichtung von $\alpha = 1,0$ erzielt, während für die Banerjee Nachbarschaft eine Gewichtung von $\alpha = 1,0$ zu dem schlechtesten Ergebnis führte. Für die Banerjee Nachbarschaft konnten die besten Ergebnisse für $\alpha = 0,1$ erzielt werden. Jedoch war der mittlere DSC-Wert des besten Banerjee-Modells deutlich geringer, als der mittlere DSC-Wert des besten Potts-Modells.

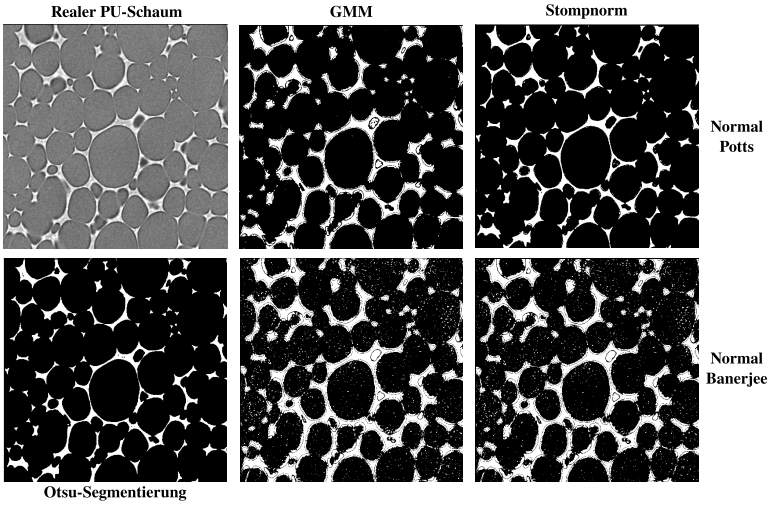


Abbildung 4.4: Darstellung einer beispielhaften Schicht der realen PU-Schaumstruktur sowie den zugehörigen Segmentierungsergebnissen der Otsu Referenzmethode und von verschiedenen EHOHMRF-Modellen.

4.3.6 Anwendung auf *RealPUFoam*-Datensatz

Nachdem zuvor die wichtigsten Parameter der EHOHMRF-Methode untersucht worden sind, sollte das EHOHMRF-Modell auf den *RealPUFoam*-Datensatz angewandt werden. Hierzu wurden jeweils Modelle mit normaler Potts- bzw. Banerjee-Nachbarschaft und mit GMM- bzw. Stompnorm-Verteilungsfunktion mit einem Nachbarschaftsgewicht von $\alpha = 1,0$ trainiert. Für die EHO wurden erneut zehn Herden mit je zehn Elefanten gewählt und es wurden 100 Epochen als maximale Epochenanzahl ausgewählt. In Abbildung 4.4 sind die Segmentierungsergebnisse für eine beispielhafte Schicht des realen PU-Schaums dargestellt. Man erkennt, dass sich die Segmentierungsergebnisse teilweise deutlich unterscheiden. Die Modelle mit Banerjee-Nachbarschaft erzeugten beide sehr viele kleine weiße Punkte innerhalb der Poren, die für die Modelle mit Potts-Nachbarschaft nicht erzeugt wurden. Der Grund hierfür liegt vermutlich darin, dass das natürliche Rauschen, welches in den µCT-Bildern vorhanden ist, sehr helle Voxel innerhalb

der Poren erzeugt, die dann vom Modell der PU-Klasse zugeordnet werden. Der Fehler tritt vor allem im Banerjee-Modell auf, da hier die Mittelwerte und Standardabweichungen der Klassen in der Nachbarschaft berücksichtigt werden. In diesem Fall wäre es für das Modell ungünstig, die hellen Voxel vom Rauschen innerhalb der Poren dem Hintergrund zuzuordnen, wodurch diese Voxel dann stattdessen der PU-Klasse zugeordnet werden. Die Segmentierung mit Stompnorm Verteilung und Potts-Nachbarschaft erzeugte eine bessere Segmentierung, als die Referenzmethode der Otsu-Schwellwertbestimmung und Thresholding. Manche der feinen Stege wurden in der Otsu-Segmentierung nicht korrekt segmentiert, während diese vom EHOHMRF-Modell segmentiert wurden.

In allen vier Fällen wurden die feinen Stege, deren Signale sich im µCT-Bild kaum von den Luftporen unterscheiden, fast immer der Luft-Klasse zugeordnet. Nur die feinen Stege mit deutlich hellerem Signal wurden der PU-Klasse zugeordnet. Ob jedoch alle dieser nicht segmentierten Stege einen Fehler darstellen muss hinterfragt werden, da es sein könnte, dass der gemessene Voxel ein Mischsignal aus PU-Klasse und Luft-Klasse besitzt, da der Steg in dieser Schicht tatsächlich noch nicht (komplett) sichtbar ist. Des Weiteren fällt auf, dass für alle Modelle, außer dem Modell mit Stompnorm-Verteilung und Potts-Nachbarschaft, ein zusätzlicher Rand rund um die PU-Stege segmentiert wurde. Dieser Rand erscheint in den Bildern sehr dunkel und unterscheidet sich daher vom typischen Signal der Luftporen und des PU-Materials. Das Problem entsteht möglicherweise, da dem Modell nur eine Einteilung in zwei Klassen erlaubt war, es aber eigentlich drei Klassen definieren müsste, um das Bild korrekt zu segmentieren. Um dies zu überprüfen, wurden alle Modelle erneut trainiert, mit dem Unterschied, dass diesmal nach drei Klassen im Bild gesucht wurde. Mit einem Postprocessing-Schritt wurden dann die beiden Klassen mit dem niedrigsten Mittelwerten zu einer Klasse zusammengefasst.

Die Ergebnisse dieses Versuchs sind in Abbildung 4.5 dargestellt. Die oberste Reihe zeigt die direkten Ergebnisse der EHOHMRF-Segmentierung in drei Klassen. Man erkennt, dass die dunklen Ränder entlang der PU-Stege nun einer eigenen Klasse (Dunkelblau) zugewiesen wurden und sich damit von der Steg-Klasse

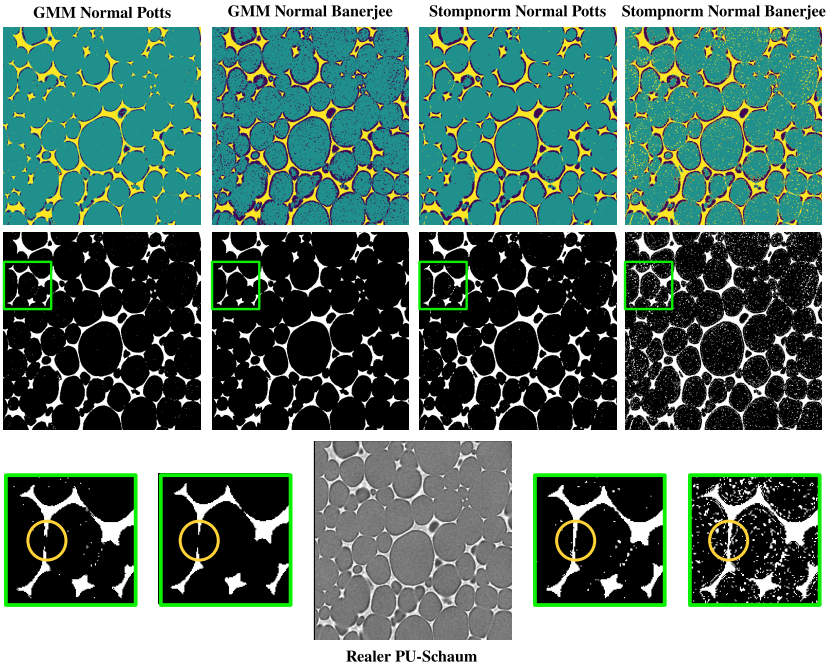


Abbildung 4.5: Darstellung einer beispielhaften Schicht der realen PU-Schaumstruktur, sowie den zugehörigen Segmentierungsergebnissen verschiedener EHOHMRF-Modelle mit Clustering in drei Klassen und anschließendem Postprocessing zur Einteilung in zwei Klassen. Die erste Reihe zeigt die vom EHOHMRF-Modell vorhergesagte Einteilung in drei Klassen, die zweite Reihe zeigt die finale Segmentierung in zwei Klassen und die dritte Reihe zeigt das Grauwertbild sowie vergrößerte Ausschnitte der Segmentierung.

(Gelb) und der Luft-Klasse (Türkis) unterscheiden. Die zweite Reihe zeigt die Bilder nach dem Postprocessing, bei dem die türkise und dunkelblaue Klasse zusammengeführt wurden. Keines der vier verschiedenen Modelle erzeugte mehr zu dicke Stege, was zeigt, dass der Grund für die fehlerhaften Segmentierungen zuvor wirklich eine zu geringe Anzahl an Klassen war. Die vier Segmentierungsergebnisse ähneln sich alle sehr stark, aber die Segmentierung des Modells mit Stompnorm-Verteilung und Banerjee-Nachbarschaft zeigte erneut viele kleine weiße Punkte innerhalb der Poren. Betrachtet man die Detailausschnitte in der dritten Reihe, so erkennt man jedoch leichte Unterschiede in den Segmentierungen. Während für die beiden Modelle mit GMM-Verteilung der mit einem gelben Kreis markierte Steg nicht komplett durchgängig war, wurde dieser in den beiden Stompnorm-Segmentierungen geschlossen.

Insgesamt scheint die EHOHMRF-Methode bei passender Wahl der Parameter gute Segmentierungen von µCT-Bildern von PU-Schäumen generieren zu können. Jedoch ist es wichtig, die passende Nachbarschaft und Verteilungsfunktion zu wählen und insbesondere die Anzahl der Klassen lieber etwas höher zu wählen und die Klassen dann per Postprocessing auf die gewünschte Klassenanzahl zu verkleinern.

4.3.7 Anwendung auf Lungen-MRT-Datensatz

Um die Flexibilität der EHOHMRF-Segmentierung zu zeigen, wurde die Methode neben den PU-Schaum-Datensätzen ebenso auf dem Lungen-MRT Datensatz getestet. Hierzu wurden Modelle mit verschiedenen Nachbarschaftsarten und Nachbarschaftsgewichten trainiert. Da die Anzahl an Klassen im Lungen-MRT Bild höher ist, wurden in diesem Fall fünf statt zwei Klassen definiert. Die Modelle wurden jeweils mit dem GBEHO-Algorithmus für je 25 Epochen trainiert, da aufgrund der größeren 3D-Bilder die schnellere Konvergenz des GBEHO gegenüber den leicht verbesserten Ergebnissen des normalen EHO präferiert wurde. Im Anschluss wurde die Segmentierungsmaske mit fünf Klassen nachbearbeitet, um die Lungenmaske zu extrahieren. Hierzu wurde zunächst für jede der fünf

Klassen jeweils nur das größte Objekt extrahiert, um kleine Artefakte zu entfernen. Anschließend wurde ein Schwellwert τ für die Lungenwerte bestimmt. Hierzu wurde ein Histogramm des mediangefilterten Bildes erstellt und in diesem die Position des ersten Tales bestimmt, das die Lungen und den Hintergrund jeweils vom umliegenden Gewebe trennt. Für jede Klasse im Bild wurde jeweils der Mittelwert bestimmt und dann wurden alle Klassen, deren Mittelwert $\mu_l < 1,5\tau$ betrug, als Lungenklassen definiert. Die dadurch entstandene Lungenmaske wurde final noch mit morphologischen Öffnungs- und Schließungsoperationen nachbearbeitet und es wurden zusätzlich kleine Löcher in der Maske automatisch gefüllt.

In Abbildung 4.6 sind für eine vordere, eine mittlere und eine hintere Schicht jeweils das MRT-Bild sowie die beiden Segmentierungsmasken dargestellt. Die zweite Reihe zeigt das direkte Segmentierungsergebnis des EHOHMRF-Modells mit fünf vorhergesagten Klassen. Die dritte Reihe zeigt das Segmentierungsergebnis nach dem Postprocessing. Man erkennt, dass die Lunge für alle Schichten gut segmentiert wurde. Ein paar besonders helle Gefäße innerhalb der Lunge wurden nicht mitsegmentiert. Dies dürfte, je nach Anwendungsfall, eventuell jedoch gewünscht sein. In diesem Fall müsste man mit weiteren Postprocessing Schritten Gefäße innerhalb der Lunge wieder einschließen. Insgesamt wurden drei Beispielbilder des Lungen-MRT-Datensatzes untersucht, für die jeweils Groundtruth-Segmentierungen einer Radiologin vorlagen. Die Evaluation der Segmentierungsergebnisse ist in Tabelle 4.6 dargestellt. Es zeigt sich, dass das Modell mit der normalen Potts-Nachbarschaft und einem Gewicht von $\alpha = 0,1$ für das erste und dritte Beispielbild jeweils die beste Segmentierung erzeugte. Auch für das zweite Bild erzeugte das Modell mit normaler Potts-Nachbarschaft die beste Segmentierung, jedoch diesmal für $\alpha = 1,0$. Das Modell mit normaler Potts-Nachbarschaft und $\alpha = 1,0$ lieferte im Mittel die konstantesten und besten Segmentierungsergebnisse. Die Ergebnisse lassen zusätzlich darauf schließen, dass die Modelle mit normaler Potts-Nachbarschaft konstantere Segmentierungsergebnisse als die Modelle mit normaler Banerjee-Nachbarschaft liefern.

In Abbildung 4.7 sind die Ergebnisse für alle drei Beispielbilder und drei verschiedene Schichten für das Modell mit Potts-Nachbarschaft und $\alpha = 1,0$ dargestellt.

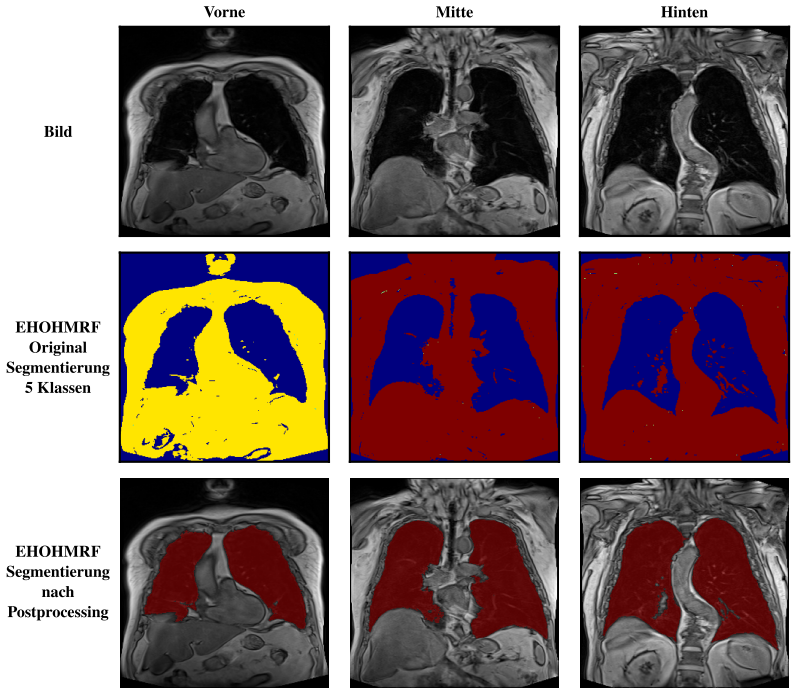


Abbildung 4.6: Darstellung von verschiedenen Schichten eines Beispielbildes sowie den zugehörigen Segmentierungsergebnissen des EHOHMRF-Modells. In der ersten Reihe ist das Bild, in der zweiten Reihe das ursprüngliche Segmentierungsergebnis des EHOHMRF-Modells für fünf Klassen und in der dritten Reihe die generierte Lungenmaske nach dem Postprocessing des Segmentierungsergebnisses abgebildet.

Dabei sind Bereiche, die mit der Groundtruth-Segmentierung übereinstimmen, Grün eingefärbt, Bereiche, die nur von der Radiologin segmentiert wurden, in Blau eingefärbt und Bereiche, die nur von der EHOHMRF-Segmentierung segmentiert wurden, sind in Gelb dargestellt. Man erkennt für alle drei Beispiele eine gute Übereinstimmung zwischen der EHOHMRF-Segmentierung und der Groundtruth-Segmentierung. Die fehlerhaften Bereiche entsprachen meist entweder der Lungenrändern oder Gefäßen innerhalb der Lunge, welche von der Radiologin mitsegmentiert wurden und von der EHOHMRF-Segmentierung ausgeschlossen

Tabelle 4.6: Vergleich der DICE-Scores von EHOHMRF-Modellen mit verschiedenen Nachbarschaftsarten und -gewichten für drei beispielhafte Lungen-MRT Bilder.

		DSC			
	α	**Beispiel 1**	**Beispiel 2**	**Beispiel 3**	$\emptyset$
Normal Potts	0,1	**0,9744**	0,9283	**0,9668**	0,9565
	1,0	0,9598	**0,9598**	0,9627	**0,9608**
	10,0	0,0000	0,0000	0,0000	0,0000
Normal Banerjee	0,1	0,8302	0,8449	0,9648	0,8800
	1,0	0,9235	0,9192	0,9645	0,9357
	10,0	0,0000	0,7044	0,7395	0,4813

wurden. Gerade leichte Abweichungen der Lungenränder sind jedoch normale erwartete Abweichungen. Im dritten Beispiel zeigt sich jedoch, dass sowohl am unteren Rand der rechten Lunge, als auch im Bereich der Trachea größere Segmentierungsfehler entstanden sind. Allerdings sind Segmentierungsfehler im Bereich der Trachea ein bekanntes Problem für Lungen-MRT Segmentierungen, da Lunge und Trachea ein sehr ähnliches Signal besitzen und oft keine klare Trennung zwischen der Trachea und dem Lungenparenchym erkennbar ist. Dies führt zwangsläufig dazu, dass diese in einer Schicht des 3D-Bildes verbunden sind und daher die Trachea nicht beim Postprocessing herausgefiltert wird. Jedoch könnte man die Trachea in einer separaten Segmentierung selbst segmentieren und anschließend von der Segmentierung ausschließen, wie es in einer Arbeit von Kohlmann et al. [111] vorgeschlagen wurde.

4.3.8 Bewertung der EHOHMRF-Methode

Die hier entwickelte EHOHMRF-Methode zeigt sich in ihrer Flexibilität als vielversprechend. Die guten Segmentierungsergebnisse sowohl für die materialwissenschaftlichen Datensätze als auch für den medizinischen Datensatz zeigen, dass diese Methode für die Segmentierung verschiedenster Segmentierungsprobleme

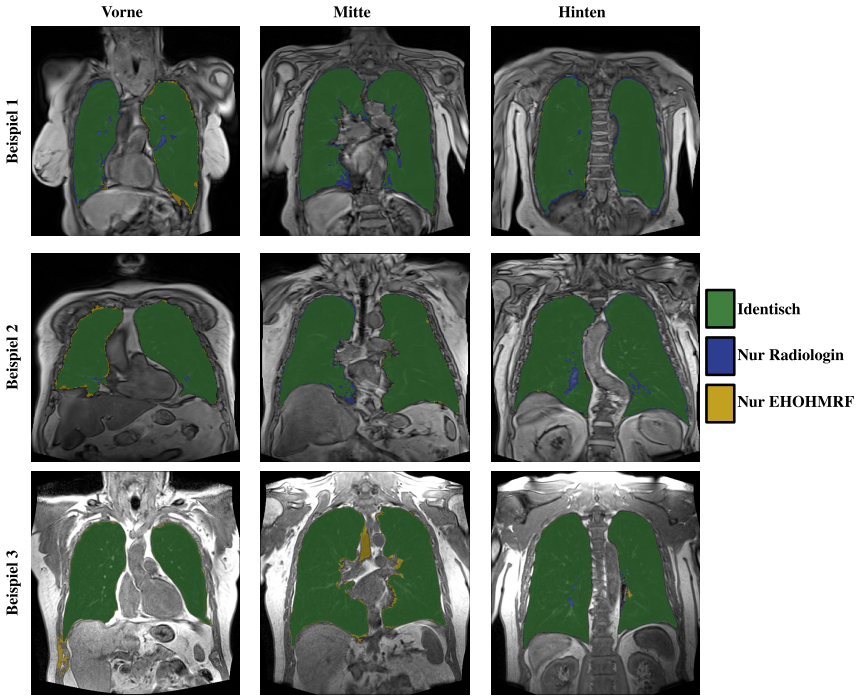


Abbildung 4.7: Vergleich der EHOHMRF Lungensegmentierung mit Groundtruth-Segmentierungen. Übereinstimmende Segmentierungsbereiche sind in Grün dargestellt, Bereiche, die nur von der Radiologin segmentiert wurden, in Blau und Bereiche, die nur vom EHOHMRF segmentiert wurden, in Gelb dargestellt.

einsetzbar ist. Während die Segmentierung der PU-Schaumdatensätze dabei eine relativ geringe Komplexität aufweist, da es nur zwei Klassen gibt, sind die Lungen-MRT Bilder deutlich komplexer, da hier neben der Lunge auch weitere Organe und umliegendes Gewebe vorhanden sind. Eine wesentliche Limitation des EHOHMRF ist, dass die Optimierung für jedes Bild erneut durchgeführt werden muss und daher jedes Mal große Rechenkapazitäten benötigt werden. In dieser Arbeit wurde der EHOHMRF-Algorithmus auf einer Nvidia A100 mit 80 GB GPU-RAM und 128 AMD EPYC 7513 32-Core Prozessoren mit einem Gesamt RAM von 2 100 GB durchgeführt. Hierbei dauerte die Segmentierung eines 256×256 Bildes

je nach verwendeten Nachbarschaftstermen und Verteilungsfunktionen zwischen 1,5–80 min, wobei die voxelweise gewichteten Nachbarschaftsfunktionen hier zu den höchsten Segmentierungszeiten führten. Bei der Anwendung der Methode auf ein ganzes 3D-Volumen, wie die Lungen-MRT Bilder, erhöhte sich selbst die Rechenzeit für die schnellsten Modelle mit normaler Potts Nachbarschaft um einen Faktor 20, während Modelle mit gewichteter Nachbarschaft fast 30 h für eines der 3D-Lungenbilder benötigten.

Ein weiterer Vorteil der EHOHMRF-Methode ist, dass diese ohne große Anpassungen auf 3D-Daten angewendet werden kann. Dabei muss nur zur passenden Berechnungsmethode für die HMRF-Funktionen gewechselt werden. Wie sich insbesondere für die Anwendung auf den *RealPUFoam*-Datensatz gezeigt hat, spielt die richtige Wahl an erwarteten Klassen eine zentrale Rolle. Dies macht die Methode zwar flexibel, da spontan auf eine andere Anzahl an Klassen gewechselt werden kann, jedoch kann die falsche Wahl auch zu schlechteren Segmentierungen führen. Die Wahl von mehr Klassen, als eigentlich erwartet werden, erhöht hierbei die Chance für eine finale richtige Segmentierung. Jedoch erfordert dies anschließend ein Postprocessing, das je nach verwendeter Methode, möglicherweise die vom HMRF berücksichtigte Nachbarschaft für die Segmentierung wieder zerstören kann.

Bei der Anwendung auf den Lungen-MRT Datensatz hat sich ein weiteres Problem der EHOHMRF-Methode gezeigt. Diese ordnet die Voxel rein anhand ihrer Signalwerte und der Klassenzugehörigkeit der Nachbarschaft einer Klasse zu. Dies führt dazu, dass möglicherweise mehrere Objekte mit ähnlichen Signalen der gleichen Klasse zugeordnet werden können. Ein Beispiel war hier die Zuordnung der Trachea zur gleichen Klasse wie die Lunge, da beide Objekte ein sehr ähnliches MRT-Signal besaßen. Um solche Probleme zu beheben, könnten entweder schon während der Segmentierung Eigenschaften wie die direkte Verbundenheit von Clustern vorausgesetzt werden, wie sie z. B. auch für SOMs, bei der GeoSOM-Implementierung eingesetzt wurden [122]. Die zweite Möglichkeit ist die Korrektur solcher Effekte während des Postprocessing, so wie es auch in dieser Arbeit geschehen ist.

4.4 HMRF-UNet

Die Verwendung der verborgenen zufälligen Markov-Felder im vorigen Unterkapitel hat das Potenzial dieses Konzeptes deutlich gezeigt. Jedoch erschweren die langen Segmentierungszeiten der EHOHMRF-Methode die breite Anwendung dieser Technik, insbesondere für große Bilder. Außerdem erfordert die EHOHMRF-Methode für jedes Bild ein erneutes Training, da das Modell nicht von vorherigen Segmentierungen lernt. Hierdurch ist die Berechnung von neuen Segmentierungen sehr rechenintensiv. Aus diesem Grund kam die Idee eine hybride Methode zu entwickeln, die die Stärken der HMRFs beinhaltet, aber von vergangenen Segmentierungen lernt und somit eine schnellere Segmentierung erzeugen kann. Hierzu wurde ein faltendes neuronales Netzwerk, in Form des weit verbreiteten U-Net, mit einer unüberwachten Loss-Funktion basierend auf der HMRF-Theorie kombiniert, um daraus ein unüberwachtes ML-Modell, das HMRF-UNet zu bauen. Das HMRF-UNet benötigt keine annotierten Groundtruth-Daten, lernt aber von den eigenen generierten Segmentierungen. In der Anwendungsphase generiert das HMRF-UNet schnelle ressourcenschonende Segmentierungen. Große Bestandteile dieses Abschnitts basieren auf einer eigenen Erstautor-Veröffentlichung mit dem Titel „Unsupervised Segmentation of Micro-CT Scans of Polyurethane Structures by Combining Hidden Markov Random Fields and a U-Net“ [123].

4.4.1 Neurartige Fuzzy HMRF-Lossfunktion für neuronale Netze

Die HMRF-Segmentierungsmethode ist normalerweise für den diskreten Fall beschrieben, in dem jeder Voxel eindeutig einer Klasse zugeordnet wird. Die Vorhersage eines neuronalen Netzes ist jedoch meistens eine unscharfe (engl. fuzzy) Wahrscheinlichkeitskarte $\mathbf{c}$, die die Wahrscheinlichkeit eines Voxels repräsentiert, dass dieser zu einer Klasse gehört. Dabei wird für jeden Voxel je ein Wahrscheinlichkeitswert für jede Klasse im Bild ausgegeben, der die Zuordnungsstärke zu der jeweiligen Klasse darstellt. Der Begriff „fuzzy“ steht hier für die

unscharfe Zuordnung eines Voxels zu allen Klassen eines Bildes. Aus diesem Grund wird eine Reformulierung der diskreten Energiefunktionen des HMRF in Gleichungen (2.34) und (2.35) vorgeschlagen, bei der die Zugehörigkeitswahrscheinlichkeiten in **c** jeweils den Beitrag der Energiefunktionen steuern. Der Mittelwert und die Standardabweichung für jede Klasse wurden mit dem Ansatz des gewichteten Mittelwertes berechnet, bei dem der Wert eines jeden Voxels zum fuzzy Mittelwert $\widetilde{\mu}_l$ und zur fuzzy Standardabweichung $\widetilde{\sigma}_l$ einer Klasse basierend auf seiner Klassenwahrscheinlichkeit $c_{s,l}$ beiträgt.

$$\widetilde{\mu}_l = \frac{\sum_{s \in S} (c_{s,l} \odot y_s)}{|c_{s,l}|} \tag{4.4}$$

$$\widetilde{\sigma}_l = \sqrt{\frac{\sum_{s \in S} (c_{s,l} \odot (y_s - \widetilde{\mu}_l)^2)}{|c_{s,l}|}} \tag{4.5}$$

Hierbei ist $\odot$ das Hadamard-Produkt und $|c_{s,l}|$ ist die Summe aller Wahrscheinlichkeiten einer Klasse l. Mit dem fuzzy Mittelwert und der fuzzy Standardabweichung einer Klasse lässt sich ein fuzzy Mittelwert $\widetilde{\mu}_{x_s}$ und eine fuzzy Standardabweichung $\widetilde{\sigma}_{x_s}$ für einen Voxel s berechnen:

$$\widetilde{\mu}_{x_s} = \sum_{l \in \mathcal{L}} (c_{s,l} \cdot \widetilde{\mu}_l) \tag{4.6}$$

$$\widetilde{\sigma}_{x_s} = \sum_{l \in \mathcal{L}} (c_{s,l} \cdot \widetilde{\sigma}_l). \tag{4.7}$$

Durch die Verwendung der fuzzy Mittelwerte und Standardabweichungen lässt sich Gleichung (2.34) umformulieren zu:

$$\widetilde{\mathbf{U}}_{\mathbf{II}}(\mathbf{y}|\mathbf{x}) = \sum_{s \in S} \left(\frac{(y_s - \widetilde{\mu}_{x_s})^2}{2\widetilde{\sigma}_{x_s}^2} + \ln \widetilde{\sigma}_{x_s} \right) . \tag{4.8}$$

Neben der Likelihood-Energiefunktion wurden auch die Energiefunktionen der beiden Nachbarschaftsterme umformuliert. Für die Umformulierung des Potts-Nachbarschaftsterms in Gleichung (2.36) wurde statt der Bestimmung der Anzahl

an ungleichen Klassen innerhalb der Nachbarschaft, die durchschnittliche euklidische Distanz zwischen den Wahrscheinlichkeitsvektoren $\mathbf{c}_s$ und $\mathbf{c}_t$ der Voxel x_s und x_t berechnet:

$$\widetilde{E}_{\text{potts}}(x_s) = \frac{\alpha}{|\mathcal{N}_s|} \sum_{t \in \mathcal{N}_s} \|\mathbf{c}_s - \mathbf{c}_t\|^2 \,. \tag{4.9}$$

Die Energiefunktion der Banerjee-Nachbarschaft aus Gleichung (2.37) wurde mithilfe der fuzzy Mittelwerte und Standardabweichung ebenso umformuliert zu:

$$\widetilde{E}_{\text{ban}}(x_s) = \frac{\alpha}{2|\mathcal{N}_s|} \sum_{t \in \mathcal{N}_s} \left((\widetilde{\mu}_{x_s} - \widetilde{\mu}_{x_t})^2 \left(\frac{1}{\widetilde{\sigma}^2_{x_s}} + \frac{1}{\widetilde{\sigma}^2_{x_t}} \right) + 1 \right). \tag{4.10}$$

Auch die voxelspezifische Nachbarschaftsgewichtung α_s aus Gleichung (2.38) wurde angepasst, um innerhalb von neuronalen Netzen verwendet werden zu können. Hierzu wurde statt der relativen Klassenhäufigkeit die Standardabweichung innerhalb der Nachbarschaft verwendet:

$$\widetilde{\alpha}_s = \begin{cases} 2\alpha, & \sigma^{(1)}_{\mathcal{N}_s} < \sigma_{\text{thresh}} \quad \text{und} \quad \sigma^{(2)}_{\mathcal{N}_s} < \sigma_{\text{thresh}} \\ \alpha, & \sigma^{(1)}_{\mathcal{N}_s} \geq \sigma_{\text{thresh}} \quad \text{und} \quad \sigma^{(2)}_{\mathcal{N}_s} < \sigma_{\text{thresh}} \\ \alpha, & \sigma^{(1)}_{\mathcal{N}_s} < \sigma_{\text{thresh}} \quad \text{und} \quad \sigma^{(2)}_{\mathcal{N}_s} \geq \sigma_{\text{thresh}} \\ 0, & \sigma^{(1)}_{\mathcal{N}_s} \geq \sigma_{\text{thresh}} \quad \text{und} \quad \sigma^{(2)}_{\mathcal{N}_s} \geq \sigma_{\text{thresh}} \end{cases}, \tag{4.11}$$

wobei $\sigma^{(k)}_{\mathcal{N}_s}$ die Standardabweichung innerhalb der Nachbarschaft k-ten Ordnung eines Voxels s und σ_{thresh} ein Schwellwert für die Standardabweichung innerhalb der Nachbarschaft ist. Mit diesen Anpassungen ergibt sich die fuzzy Energiefunktion der Nachbarschaft $\widetilde{\mathbf{U}}_{\mathbf{n}}$ zu

$$\widetilde{\mathbf{U}}_{\mathbf{n}}(\mathbf{X} = \mathbf{x}) = \sum_{s \in S} \left(\alpha \widetilde{E}_*(x_s) \right) \,. \tag{4.12}$$

Hier kann α entweder als konstanter Wert für alle Voxel (normaler Nachbarschaftsterm) oder als das voxelspezifische Gewicht (gewichteter Nachbarschaftsterm)

aus Gleichung (4.11) definiert werden. In dieser Gleichung wird $\widetilde{E}_*$ entweder mit Gleichung (4.9) oder Gleichung (4.10) berechnet.

4.4.2 Aufbau des HMRF-UNet

Für das HMRF-UNet wurde ein standardmäßiges U-Net verwendet. Das U-Net besteht aus einem Encoder- und einem Decoder-Zweig, bei dem die Encoder- und Decoder-Stufen der gleichen Höhe jeweils mit Überbrückungsverbindungen (engl. skip-connections) verknüpft sind. Auf jeder Stufe wurden mehrere Convolution-Blöcke verwendet, wobei jeder Convolution-Block aus einer 3×3 Conv2D-Schicht, einer Batch-Normalisierungs-Schicht und einer ReLu-Schicht bestand. Zwischen den einzelnen Stufen des Encoding-Zweiges wurde eine Maximum-Pooling-Schicht mit Stride=2 verwendet, während für die Stufen des Decoding-Zweiges eine UpConv2D-Schicht verwendet wurde. Die Anzahl der Kernel pro Stufe wurde für den Encoding-Zweig jeweils mit jeder Stufe verdoppelt und für den Decoding-Zweig jeweils mit jeder Stufe halbiert. Die abschließende Schicht war eine 1×1 Conv2D-Schicht mit Softmax-Aktivierung und zwei Ausgabekanälen. Eine repräsentative Darstellung des verwendeten U-Net ist in Abbildung 4.8 dargestellt.

Die Loss-Funktion des Modells bestand aus zwei Komponenten, dem Verteilungs-Loss L_d und einem Nachbarschafts-Loss L_n, die jeweils mit den beiden Gewichtungsfaktoren λ_d und λ_n gewichtet wurden:

$$L = \lambda_d L_d + \lambda_n L_n \tag{4.13}$$

Der Verteilungs-Loss L_d wurde hierbei mit Gleichung (4.8) berechnet, während der Nachbarschafts-Loss mit Gleichung (4.12) und einer Nachbarschaft erster Ordnung mit je acht Nachbarn berechnet wurde. Für alle Studien wurde der Gewichtungsfaktor des Verteilungs-Loss als $\lambda_d = 1 - \lambda_n$ definiert.

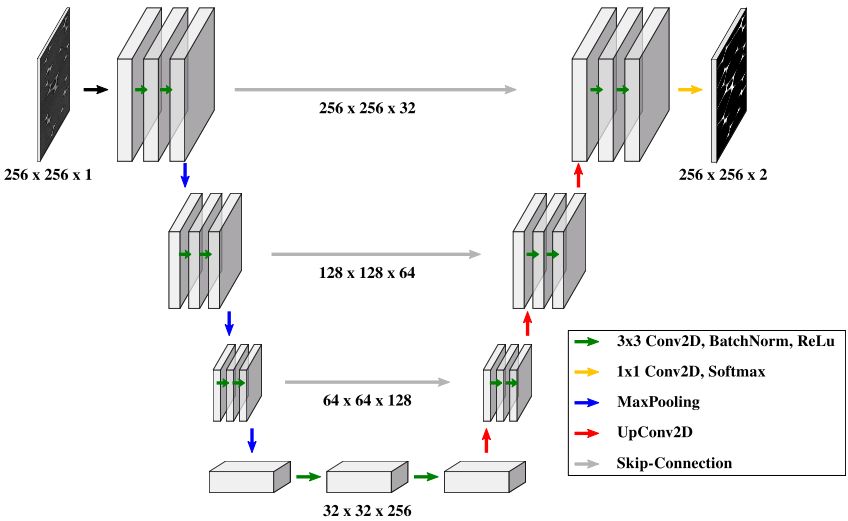


Abbildung 4.8: Überblick über die verwendete U-Net Architektur für das HMRF-UNet

4.4.3 Daten Vorverarbeitung

Alle drei Datensätze wurden vor der Verwendung mit dem HMRF-UNet vorverarbeitet. Alle Grauwertbilder wurden hierzu auf den Bereich $[0, 1]$ normalisiert. Der *ArtPUFoam*-Datensatz beinhaltet bereits eine große Menge an heterogenen und unabhängigen 2D Bildern und erforderte daher keinerlei zusätzliche Schritte, weshalb er direkt mit einer Aufteilung von $(70\%/15\%/15\%)$ in je einen Trainings-, einen Validierungs- und einen Testdatensatz aufgeteilt werden konnte.

Der *RealPUFoam*-Datensatz besteht aus einem 3D Volumen, das zunächst in einzelne 2D Bilder aufgespalten werden musste. Eine Besonderheit ist hierbei, dass sich die benachbarten Schichten sehr ähneln und sehr viele gemeinsame Informationen enthalten. Daher konnte keine normale, zufällige Aufspaltung in Trainings-, Validierungs- und Testdatensatz geschehen, sondern es wurde ein spezielles Vorgehen zur Datenaufspaltung entwickelt. Dies garantiert, das benachbarte Schichten nicht in zwei oder mehr der verschiedenen ML-Datensätze vorhanden sind, um

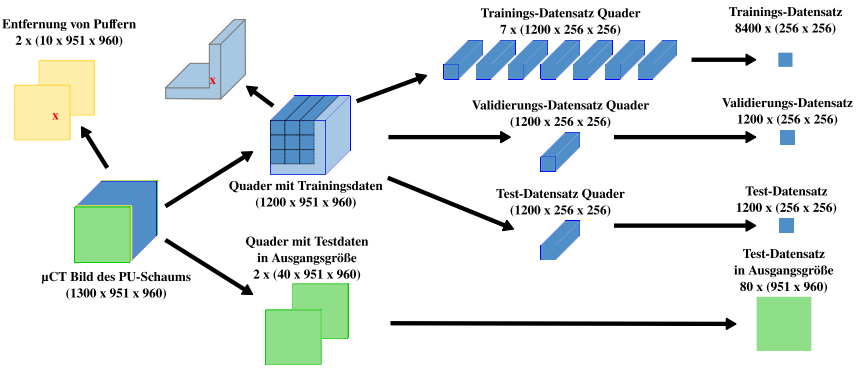


Abbildung 4.9: Visualisierung der Quader-basierten Datenaufspaltung, die für den realen PU-Schaum Datensatz verwendet wurde.

ein Datenleck zu vermeiden. Ein Datenleck zwischen den drei ML-Datensätzen ist problematisch, da hierdurch die Datensätze nicht mehr unabhängig voneinander sind und das Modell beispielsweise die Daten des Testdatensatzes bereits während des Trainings sehen könnte.

Die neu entwickelte Datenaufspaltung ist eine Quader-basierte Aufspaltung, die in Abbildung 4.9 dargestellt ist. Angelehnt an die klassische gruppenbasierte Aufspaltung, verhindert die Einteilung in Quader ein Leck zwischen den Datensätzen. Zunächst wurden je 40 Schichten von vorne und hinten entfernt, um einen unabhängigen Testdatensatz in Ausgangsgröße (grüne Quader) zu erhalten. Die zehn jeweils dahinter bzw. davor liegenden Schichten wurden als Puffer (gelbe Quader) komplett entfernt, um ein Leck zwischen dem extrahierten Testdatensatz und den restlichen Daten zu verhindern. Von dem übriggebliebenen blauen Quader mit 1 200 Schichten wurden neun kleinere Quader mithilfe eines wandernden Fensters der Größe $1200 \times 256 \times 256$ extrahiert. Sieben dieser Quader bildeten anschließend den Trainingsdatensatz, wobei nun jeweils alle Schichten der kleinen Quader extrahiert wurden, um daraus jeweils 2D Bilder der Größe 256×256 zu generieren. Jeweils ein weiterer Quader bildete den Validierungs- und Testdatensatz, die ebenso aus den einzelnen Schichten der Quader bestanden. Um zusätzlich die Datenmenge und Heterogenität der Daten zu erhöhen, wurden pro 2D Bild

in einem Datensatz jeweils fünf alternative augmentierte Versionen generiert. Hierzu wurde eine zufällige Kombination einer Kontrastanpassung im Bereich $[0.8, 1.2]$ mit einer von fünf verschiedenen Transformationen kombiniert. Die Transformationen bestanden aus einer 90°-Drehung, einer 180°-Drehung, einer 270°-Drehung, einer Spiegelung entlang der x-Achse und einer Spiegelung entlang der y-Achse. Insgesamt resultierten daraus 42 000 Trainings-, 6 000 Validierungs- und 6 000 Testbilder.

Die Bilder des Lungen-MRT Datensatzes wurden alle auf eine einheitliche Größe von 256×256 innerhalb einer jeden Schicht interpoliert. Hierzu fand die Interpolation im Frequenzbereich statt. Dazu wurde auf das Bild eine Fourier-Transformation angewendet, dann der Rand im Frequenzbereich entweder mit Nullen gefüllt, oder aber hohe Frequenzen entfernt und anschließend das Bild wieder in den Bildbereich rücktransformiert. Vor der Normalisierung wurden hohe Signale im Bild entfernt, indem alle Werte oberhalb des 99. Perzentils auf den Wert des 99. Perzentils gesetzt wurden. Nach der Normalisierung wurden jeweils 2D Bilder aus den 3D-Bildern generiert, indem die Schichten jeweils extrahiert wurden. Für die Aufteilung in die drei Trainingsdatensätze wurde eine Patienten-weise Aufspaltung verwendet, die garantiert, dass alle Daten einer Patientin oder eines Patienten im gleichen Datensatz landen, um ein Datenleck zu verhindern. Auch hier wurde eine Aufteilung basierend von einer Verteilung von $(70\%/15\%/15\%)$ in je einen Trainings-, einen Validierungs- und einen Testdatensatz durchgeführt.

4.4.4 Optimierung & Training des HMRF-UNet

In einem ersten Schritt wurde der *ArtPUFoam*-Datensatz mit den künstlich generierten µCT-Bildern verwendet, um die optimalen Parameter für das U-Net zu ermitteln. Zunächst wurde eine Hyperparametersuche durchgeführt, in der nur die Netzwerkparameter berücksichtigt wurden. Hierzu wurde das Bayesian Optimierungsverfahren für die Hyperparameteroptimierung verwendet. Als Metrik diente der DICE-Score zwischen generierter Segmentierung und Groundtruth-Segmentierung. Es wurden verschiedene Werte für die Größe und Anzahl der

Kernel, die Anzahl der Stufen, die Anzahl der Faltungs-Schichten pro Stufe und die Art der Pooling-Operation variiert. Die ermittelte optimale Architektur war ein U-Net mit drei Stufen mit je drei Faltungs-Schichten, einer Max-Pooling-Schicht, einer Kernel-Größe von 3×3 und einer Ausgangskernelanzahl von 64 Kerneln. Mit diesem Modell wurde in einem zweiten Optimierungsverfahren der Gewichtungsfaktor λ_n des Nachbarschaftsterms variiert, um hier die optimalen Werte zu bestimmen. Die fünf besten ermittelten Gewichtungsfaktoren wurden anschließend für das Training aller weiteren Modelle verwendet. Die Lernrate wurde basierend auf einer empirischen Untersuchung auf einen Wert von $1e-5$ festgelegt und die Batchgröße wurde auf 128 gesetzt, um den GPU-Knoten optimal auszulasten. Falls nicht anders erwähnt, wurden alle Modelle für je 200 Epochen trainiert. Die Hyperparametersuche und alle Trainings wurden mithilfe von Tensorflow 2.11 auf einer Nvidia A100 Grafikkarte mit 80 GB GPU-RAM und 128 AMD EPYC 7513 32-Core Prozessoren mit einem Gesamt-RAM von 2 100 GB durchgeführt.

4.4.5 Einfluss des Nachbarschaftsterms

Der Nachbarschafts-Loss kann durch vier verschiedene Loss-Varianten, die normale und gewichtete Potts- sowie die normale und gewichtete Banerjee-Nachbarschaft berechnet werden. In Tabelle 4.7 sind die mittleren DICE-Scores für die Segmentierungen des Testdatensatzes für verschiedene Modelle mit unterschiedlichen Nachbarschaftstermen und Nachbarschaftsgewichten dargestellt. Alle Modelle, die mit dem normalen Banerjee-Loss trainiert wurden, generierten im Vergleich zu allen anderen Modellen signifikant schlechtere Segmentierungen. Die beiden Modelle mit Potts-Nachbarschaft besaßen minimal höhere DSC-Werte im Vergleich zum Modell mit gewichteter Banerjee-Nachbarschaft. Für die beiden gewichteten Nachbarschaftsformulierungen führte ein geringerer Gewichtungsfaktor λ_n zu besseren Segmentierungen, während für die Modelle ohne voxelspezifische Gewichtung kein einheitlicher Trend erkennbar war.

Abbildung 4.10 zeigt einen Vergleich der vorhergesagten Segmentierungen von Modellen, die mit verschiedenen Nachbarschaftstermen und $\lambda_n = 0,31$ trainiert

Tabelle 4.7: Vergleich der mittleren DICE-Scores (Mittelwert ± Standardabweichung) des HMRF-UNet für verschiedene Nachbarschaftsarten und Nachbarschaftsgewichte.

	DSC				
Ohne Nachbarschaft	0.950 ± 0.015				
λ_n	**0.09**	**0.18**	**0.31**	**0.35**	**0.56**
Normal Potts	0,927 ± 0,077	**0,956 ± 0,016**	**0,957 ± 0,017**	**0,957 ± 0,017**	**0,951 ± 0,019**
Weighted Potts	**0,956 ± 0,015**	**0,956 ± 0,017**	0,954 ± 0,018	0,952 ± 0,019	0,950 ± 0,016
Normal Banerjee	0,952 ± 0,015	0,914 ± 0,091	0,882 ± 0,103	0,882 ± 0,103	0,771 ± 0,088
Weighted Banerjee	0,955 ± 0,016	0,953 ± 0,018	0,947 ± 0,019	0,946 ± 0,019	0,943 ± 0,013

worden sind. Abgesehen vom schlechten Bild in der dritten Spalte erzeugten die Modelle mit normaler und gewichteter Potts-Nachbarschaft sowie mit gewichteter Banerjee-Nachbarschaft bessere Segmentierungen als das Modell, welches ohne Nachbarschafts-Loss trainiert wurde. Das Modell ohne Nachbarschafts-Loss HMRF-UNet$_{\text{none}}$ erzeugte eine Übersegmentierung der PU-Struktur innerhalb der Luftporen. Man erkennt außerdem, dass das Modell mit normaler Banerjee-Nachbarschaft in allen Fällen die schlechtesten Segmentierungen generiert hat. Für das schlechteste Bild in der dritten Spalte war hierfür eine Übersegmentierung der Stege (blaue Bereiche) verantwortlich. Die voxelspezifische Gewichtung führte bei der Banerjee-Nachbarschaft zu deutlich besseren Segmentierungen, während sie bei der Potts-Nachbarschaft zu einer Verschlechterung führte. Für alle Modelle war die Hauptfehlerursache eine Untersegmentierung der feinen PU-Strukturen (rote Bereiche).

Zusätzlich wurde der Einfluss des Schwellwerts σ_{thresh} untersucht, der verwendet wird, um die voxelspezifischen Gewichtungsfaktoren α_s zu berechnen. Abbildung 4.11 zeigt den Einfluss des Schwellwertes für zwei beispielhafte Bilder des Testdatensatzes. Für beide Bilder führte eine Reduktion des Schwellwerts zu einer Verbesserung der Segmentierungsergebnisse. In den mit roten Kreisen markierten Bereichen erkennt man, dass für den kleinsten Schwellwert von 0,05 sogar dünne PU-Stege segmentiert wurden, die beispielsweise für einen Schwellwert von $0,20$ nicht detektiert wurden.

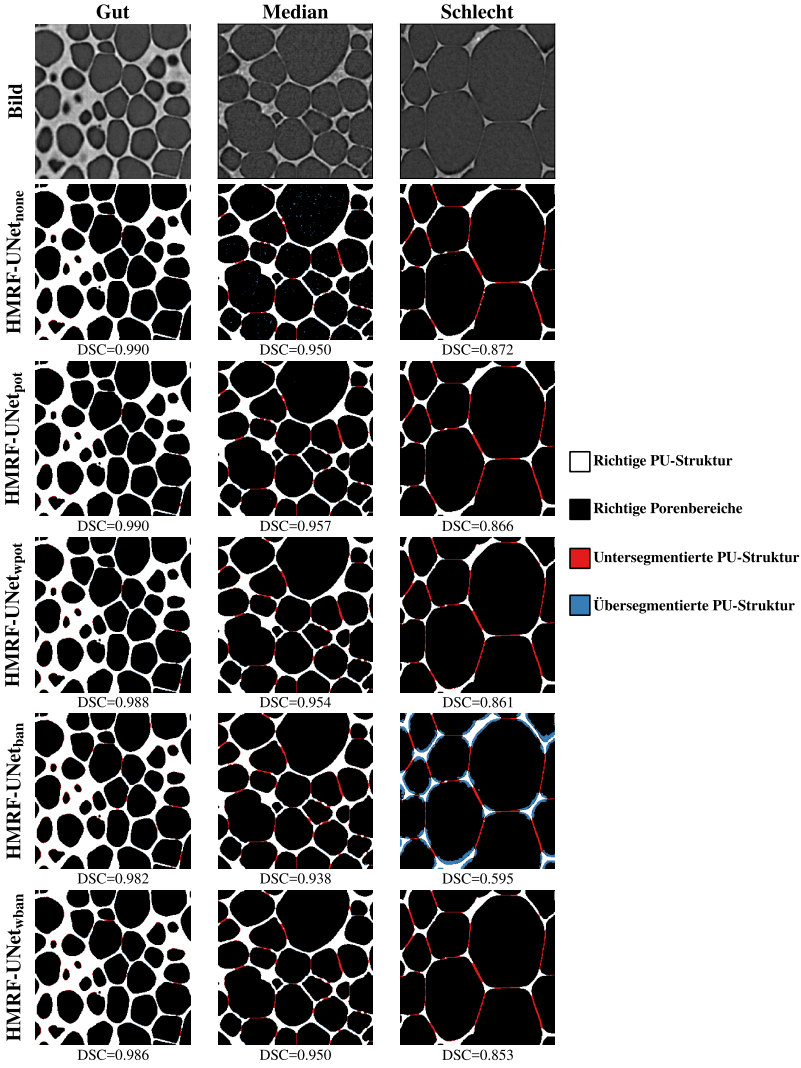


Abbildung 4.10: Vergleich der Segmentierungsergebnisse verschiedener HMRF-UNet Modelle für jeweils ein gutes, ein mittleres und ein schlechtes Ergebnis für Bilder des Testdatensatzes. Korrekt segmentierte Porenbereiche sind in Schwarz, korrekt segmentierte PU-Strukuren in Weiß, untersegmentierte PU-Strukturen in Rot und übersegmentierte PU-Strukturen in Blau dargestellt.

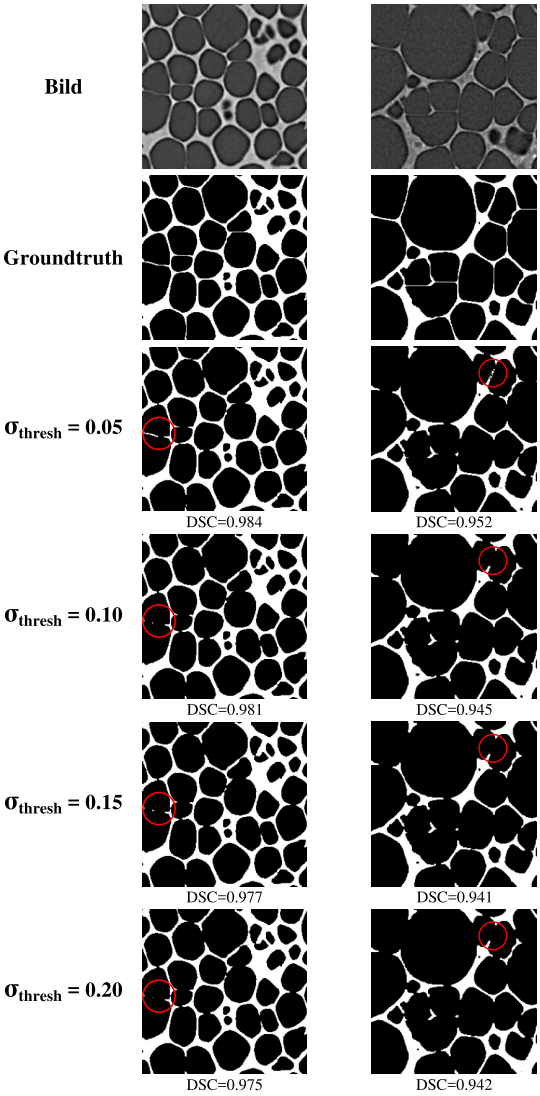


Abbildung 4.11: Vergleich der Segmentierungsergebnisse von Modellen mit gewichteter Banerjee-Nachbarschaft und verschiedenen Schwellwerten σ_{thresh} für zwei beispielhafte Bilder des Testdatensatzes.

Die passende Auswahl des Nachbarschaftsterms scheint sowohl für das HMRF-UNet als auch für die EHOHMRF-Methode eine zentrale Rolle zu spielen. Die Ergebnisse lassen darauf schließen, dass beide Potts-Nachbarschaftsvarianten beide Banerjee-Nachbarschaftsvarianten übertreffen. Ein Grund hierfür könnte sein, dass für eine binäre Segmentierung die Banerjee-Formulierung keinen wirklichen Mehrwert gegenüber der Potts-Methode besitzt, da die Berechnung der Differenz zweier Mittelwerte sich nicht zu stark von der Berechnung der Distanz der Wahrscheinlichkeitswerte unterscheidet. Ein weiterer Grund könnte sein, dass in der originalen Implementierung des Banerjee-Terms die berechneten Mittelwerte und Standardabweichungen auf genauen Klassenzuordnungen basieren. Dadurch stellen diese direkt die Eigenschaften der Klassenverteilung dar, was für die fuzzy Berechnung der Werte nicht der Fall ist, da hier eine Gewichtung anhand der Wahrscheinlichkeitszugehörigkeiten erfolgt.

Eine weitere Beobachtung war, dass die voxelweise Gewichtung insbesondere für den Banerjee-Loss wichtig ist. Ohne die voxelweise Gewichtung waren die Segmentierungsergebnisse deutlich schlechter als für das Modell mit dieser voxelweisen Gewichtung. Die beispielhaften Segmentierungen in Abbildung 4.10 deuten darauf hin, dass die niedrigen DICE-Scores vor allem an Übersegmentierungen von Bildern mit geringem Kontrast liegen. Der Bildkontrast schien die Segmentierungsqualität aller Modelle zu beeinflussen, da alle vier Modelle für das Bild mit niedrigem Kontrast in der letzten Spalte einen deutlich niedrigeren DSC-Wert besaßen. Um diese Hypothese zu prüfen, wurde die Standardabweichung innerhalb der Bilder als ein Maß des Bildkontrastes berechnet. Es wurden jeweils die mittlere Standardabweichung der je 50 schlechtesten und besten Segmentierungen für das HMRF-UNet$_{\text{ban}}$-Modell berechnet. Dabei ergab sich eine mittlere Standardabweichung von 0.187 ± 0.020 für die besten und 0.097 ± 0.011 für die schlechtesten Segmentierungen. Dies stützt die Hypothese, dass ein kleinerer Bildkontrast zu schlechteren Segmentierungen führt. Außerdem wurde festgestellt, dass die richtige Wahl des Schwellwerts σ_{thresh} ein ebenso wichtiger Faktor ist, um die Segmentierungen insbesondere im Bereich der feinen Stege zu verbessern.

Tabelle 4.8: Analyse der Verwendung des HMRF-UNet zum Vortrainieren eines UNet als Ausgangspunkt für ein überwachtes Training. Vergleich der mittleren DICE-Scores (Mittelwert ± Standardabweichung) für Modelle mit und ohne HMRF-Vortraining für unterschiedlich große Trainingsdatensätze.

	DSC				
# Trainingsbilder	**5**	**10**	**100**	**1000**	**14000 (alle)**
Unüberwachtes HMRF-UNet	-	-	-	-	$0,957 \pm 0,017$
Überwachtes U-Net ohne Vortraining	$0,848 \pm 0,040$	$0,859 \pm 0,043$	$0,854 \pm 0,038$	$0,999 \pm 0,001$	$1,000 \pm 0,000$
Überwachtes U-Net HMRF Vortraining	$0,977 \pm 0,015$	$0,983 \pm 0,010$	$0,992 \pm 0,005$	$0,999 \pm 0,000$	-

4.4.6 Verwendung des HMRF-UNet zum Vortrainieren

Ein typischer Anwendungsfall für unüberwachte ML-Methoden ist die Verwendung zum Vortrainieren eines Netzes, um dieses damit gut zu initialisieren. Anschließend kann das Netz mit einem überwachten Training feinjustiert werden. Um die Eignung des HMRF-UNet für das Vortrainieren zu testen, wurde ein Experiment erstellt, das ein U-Net mit überwachtem Loss auf verschiedenen Anteilen des *ArtPUFoam*-Datensatzes trainiert. Hierbei wurden die Gewichte des U-Net entweder zufällig initialisiert oder die Gewichte wurden aus einem zuvor trainierten HMRF-UNet extrahiert. Für die Extraktion der Gewichte wurde hier ein $\text{HMRF-UNet}_{\text{pot}}$-Modell mit $\lambda_n = 0.31$ gewählt, das mit dem kompletten Trainingsdatensatz für 200 Epochen trainiert wurde. Der gewählte Loss für das überwachte Training war der für die Segmentierung weit verbreitete DICE-Loss [124].

In Tabelle 4.8 werden verschiedene U-Net Modelle verglichen, die alle mit einem überwachten DICE-Loss mit und ohne vortrainierten Modellgewichten trainiert wurden. Dabei wurden unterschiedlich große Trainingsdatensätze für das Training verwendet. Man erkennt, dass die Modelle, die ohne vortrainierte Gewichte trainiert wurden, erst ab 1 000 Bildern eine gute Segmentierungsqualität

erreichten. Für die Modelle mit den vortrainierten Gewichten aus dem HMRF-UNet wurden gute Segmentierungsergebnisse bereits für die kleinste Anzahl von fünf Trainingsbildern erreicht. Dabei wurde durch das überwachte Feinjustieren die Segmentierungsqualität schrittweise von DSC $= 0,957$ auf DSC $= 0,977$, DSC $= 0,983$ und DSC $= 0,992$ verbessert, wenn jeweils mit fünf, zehn oder 100 Bildern nachtrainiert wurde.

Der Effekt des Vortrainierens kann ebenso in Abbildung 4.12 beobachtet werden. Hier werden jeweils die Segmentierungsergebnisse mit den größten (erste Spalte), mittleren (zweite Spalte) und kleinsten Veränderungen (dritte Spalte) durch das Vortrainieren dargestellt. Die überlagerten Segmentierungskarten in den letzten beiden Reihen zeigen mittels einer Farbcodierung, welchen Effekt das Vortrainieren auf Modelle hat, die mit unterschiedlichen Größen des Trainingsdatensatzes trainiert wurden. Man erkennt, dass bereits das unüberwachte HMRF-UNet gute Segmentierungen in allen drei Fällen generiert hat (weiße Bereiche im untersten Bild). Fast alle der noch fehlenden feinen PU-Stege konnten durch Feinjustieren mittels überwachtem Training mit zehn Bildern ebenso segmentiert werden. Für ein Feinjustieren mit 100 Bildern konnten dann bereits nahezu perfekte Segmentierungsergebnisse erzeugt werden. Für das überwachte U-Net ohne vortrainierte Gewichte (dritte Reihe) war dagegen ein Training mit 1 000 Bildern notwendig, um die feinen PU-Stege zu segmentieren. Diese Beobachtungen werden durch die Boxplots der DICE-Scores beider Modellvarianten in Abbildung 4.13 bestätigt. Für alle Mengen an Feinjustierungs-/Trainingsdaten war das Modell, das mit dem HMRF-UNet vortrainiert wurde, dem Modell ohne Vortraining überlegen. Die Signifikanz dieser Beobachtung wurde mithilfe eines Wilcoxon-Vorzeichen-Rangtest überprüft, der einen p-Wert von $p < 0,001$ für alle fünf Fälle ergab.

Der positive Effekt eines unüberwachten Vortrainierens aufgrund eines Regularisierungseffektes konnte bereits in diversen Arbeiten gezeigt werden [125–127]. Die Ergebnisse der HMRF-UNet Studien unterstützen die Vermutung, dass durch ein Vortrainieren mit dem HMRF-UNet die benötigte Anzahl an annotierten Groundtruth-Daten für das Trainieren von Segmentierungsmodellen deutlich reduziert werden kann. Schon kleine Mengen von fünf bis zehn Bildern reichten

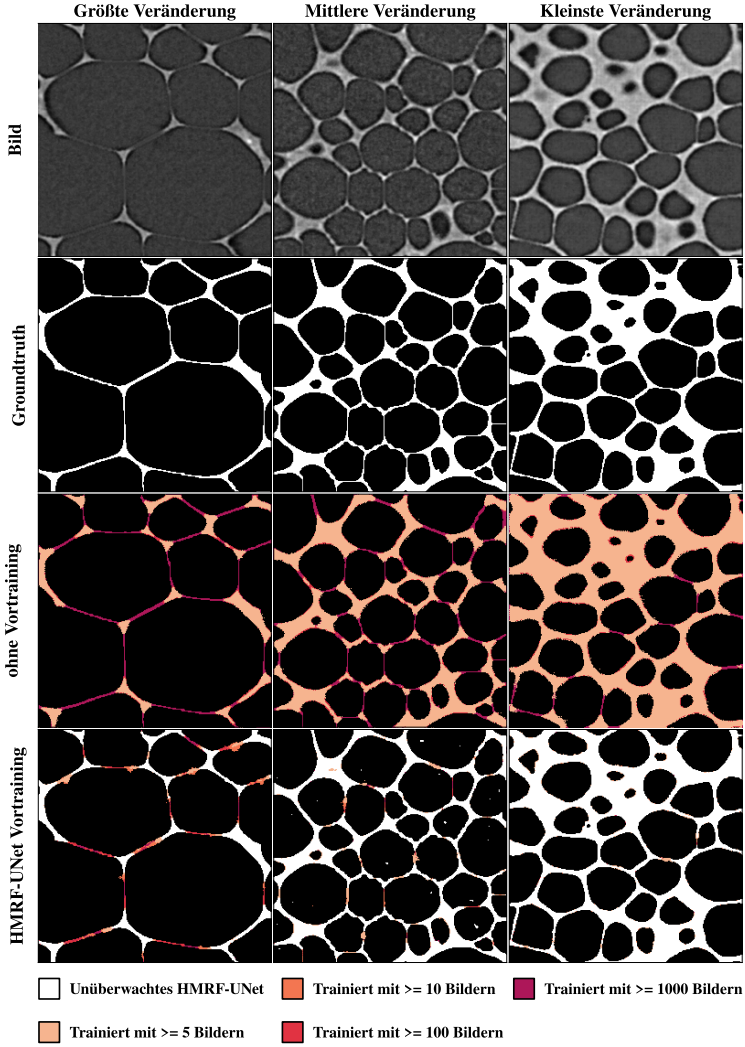


Abbildung 4.12: Visualisierung der Effekte des Vortrainierens mit dem HMRF-UNet auf ein überwachtes Training für verschiedene Mengen an Trainingsdaten. Dargestellt sind die Bilder (erste Reihe), die Groundtruth-Segmentierungen (zweite Reihe), die Segmentierungsergebnisse für Modelle ohne (dritte Reihe) und mit HMRF-UNet Vortraining (vierte Reihe). Die Einfärbung in den unteren beiden Reihen zeigt den Effekt der Verwendung von mehr Trainingsdaten auf das Segmentierungsergebnis.

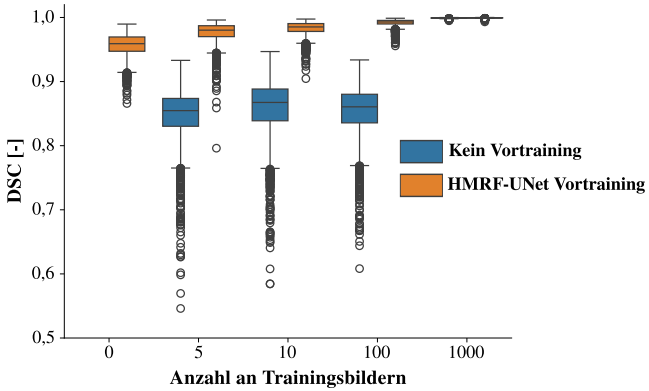


Abbildung 4.13: Boxplots der DICE-Scores für Modell mit (Orange) und ohne (Blau) Vortraining durch das HMRF-UNet für verschiedene Größen des Trainingsdatensatzes. Für alle Mengen an Trainingsdaten erzielte das Modell mit Vortraining signifikant ($p<0{,}001$) höhere DSC-Werte.

in den Versuchen aus, um die Segmentierungsqualität signifikant zu verbessern. Bei einem Nachtrainieren mit 100 Bildern wurde sogar ein nahezu perfektes Segmentierungsergebnis erzeugt. Im Vergleich dazu benötigte ein Modell ohne Vortraining zehnmal mehr Daten für eine vergleichbare Segmentierungsgüte. Diese Beobachtungen decken sich mit den Ergebnissen von Kalapos und Gyires-Tóth [127], die zeigen konnten, dass nahezu perfekte Segmentierungsergebnisse erzeugt werden können, wenn ein vortrainiertes Modell ebenso mit nur ca. 100 Bildern nachtrainiert wird. Diese Ergebnisse zeigen somit das große Potenzial des hier entwickelten unüberwachten HMRF-Loss für jede Segmentierungsaufgabe. Der HMRF-Loss könnte daher eine gute Ergänzung zu bereits bekannten Vortrainings-Methoden wie Clustering, kontrastivem Lernen und generativen Modellen sein.

4.4.7 Anwendung auf *RealPUFoam*-Datensatz

Nachdem die Parameter des HMRF-UNet in den Abschnitten zuvor intensiv erforscht wurden, kann nun eine Validierung des Modells auf dem realen µCT-Datensatz geschehen. Hierbei sollte untersucht werden, ob ein Training anhand künstlich generierter Daten oder realer Daten zu unterschiedlichen Resultaten führt. Zunächst wurden ein HMRF-UNet mit normaler Potts-Nachbarschaft und $\lambda_n = 0,18$ sowie ein HMRF-UNet mit gewichteter Banerjee-Nachbarschaft und $\lambda_n = 0,18$ für je 200 Epochen auf dem *RealPUFoam*-Datensatz trainiert. Anschließend wurde das trainierte Modell auf den Testdatensatz angewendet. Ebenso wurden die zuvor bereits auf dem *ArtPUFoam*-Datensatz trainierten Modelle mit normaler Potts-Nachbarschaft mit und ohne überwachtem Finetuning mit 100 Bildern auf die Testbilder des *RealPUFoam*-Datensatzes angewendet.

Die Ergebnisse für ein beispielhaftes Bild sind in Abbildung 4.14 dargestellt. Es zeigt sich, dass in allen vier Segmentierungen manche dünnen PU-Stege nicht segmentiert wurden (rote Kreise), während andere feine Stege zumindest von dem feingetunten Modell, das mit dem künstlichen Datensatz trainiert wurde, segmentiert wurden (gelb-gestrichelte Kreise). Das gewichtete Banerjee-Modell erzeugte hier übersegmentierte PU-Stege sowie insgesamt die schlechtesten Segmentierungsergebnisse. Zusätzlich schien es keine wesentlichen Unterschiede im Segmentierungsergebnis für die beiden normalen Potts-Modelle zu geben, die nicht nachtrainiert wurden. Dies zeigt, dass auch ein komplett unüberwachtes Training auf den realen Daten genauso gut funktioniert, wie ein Training auf künstlich erzeugten Daten. Der Grund für die übersegmentierten Stege für das gewichtete Banerjee-Modell könnte erneut an einer zu geringen Anzahl an möglichen Klassen im Bild liegen.

Um dies zu überprüfen, wurde ein weiteres Experiment durchgeführt, indem die gleichen Modelle erneut trainiert wurden, aber diesmal nach drei Klassen im Bild suchen sollten. Die Ergebnisse dieses Experiments sind in Abbildung 4.15 dargestellt. Man erkennt, dass tatsächlich die Ränder rund um die Stege in eine eigene Klasse (dunkelblau) eingeteilt werden und von der PU-Klasse (gelb) getrennt sind.

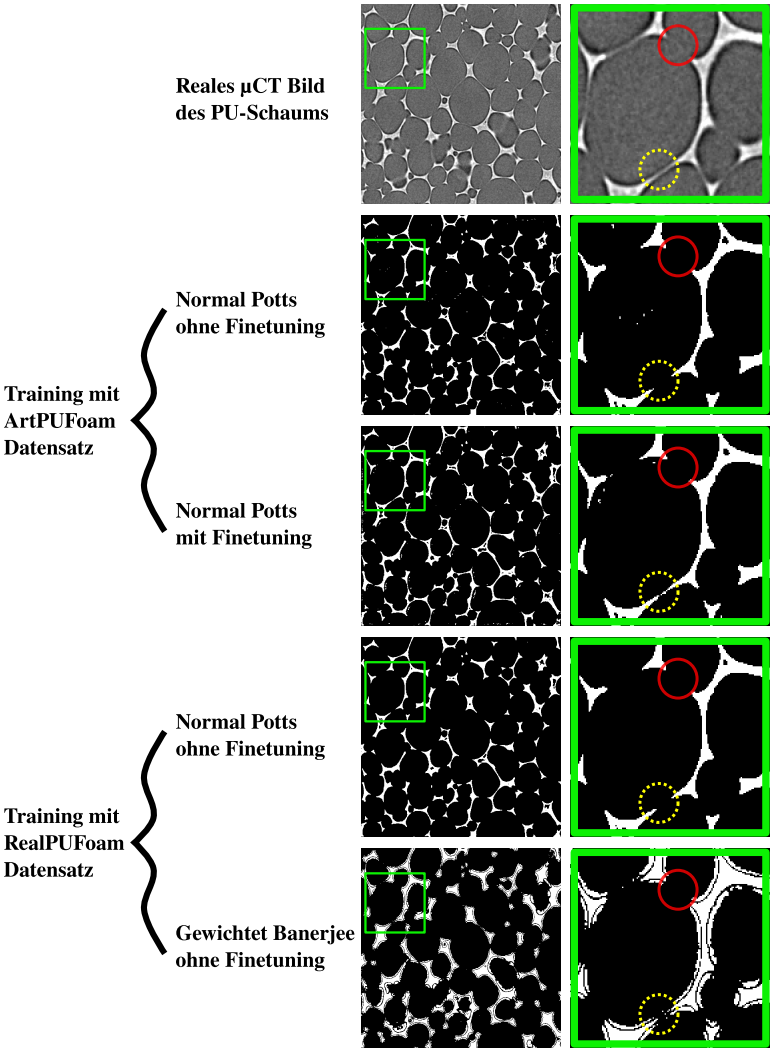


Abbildung 4.14: Vergleich der Segmentierungsergebnisse verschiedener HMRF-UNet Modelle für ein beispielhaftes Bild des realen µCT-Datensatzes. In der zweiten und dritten Reihe sind zwei Modelle gezeigt, die nur auf dem *ArtPUFoam*-Datensatz trainiert wurden. Die Ergebnisse in der vierten und fünften Reihe stammen von Modellen, die mit dem *RealPUFoam*-Datensatz trainiert wurden.

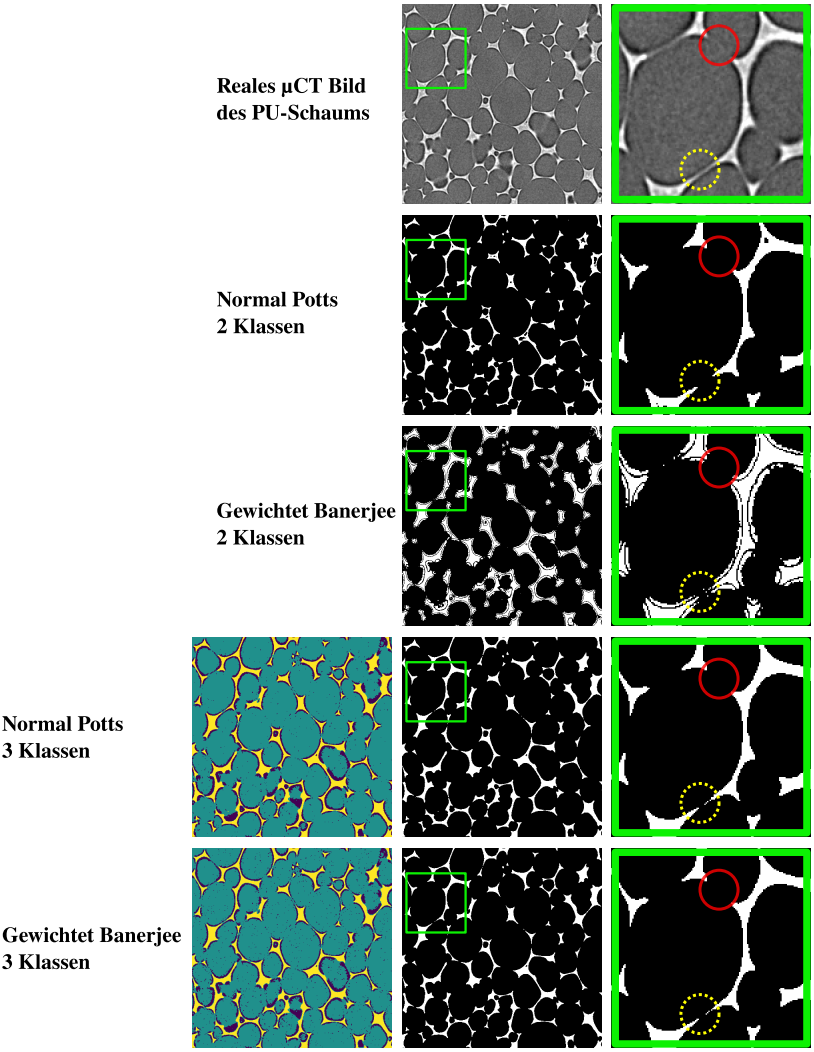


Abbildung 4.15: Vergleich der Segmentierungsergebnisse von HMRF-UNet Modellen, die zwei oder drei Klassen im Bild suchen. Die Modelle, die drei Klassen suchen (vierte und fünfte Reihe) erzeugen hierbei bessere Segmentierungen und verhindern eine Übersegmentierung der PU-Stege.

Mit einem Postprocessing-Schritt wurden die Segmentierungsmasken mit drei Klassen in eine Segmentierung für zwei Klassen umgewandelt. Die generierten binären Segmentierungsmasken zeigen, dass das Problem der übersegmentierten Stege für das gewichtete Banerjee-Modell behoben werden konnte (dritte Reihe vs. fünfte Reihe). Außerdem scheint die Einteilung in drei Klassen sowohl für das normale Potts-Modell als auch für das gewichtete Banerjee-Modell zusätzlich positive Auswirkungen auf die Segmentierung zu haben. Im Bereich des gelb-gestrichelten Kreises erkennt man, dass mehr Teile des feinen Steges segmentiert wurden und die Ergebnisse somit fast so gut waren wie die Ergebnisse des nachtrainierten Modells aus Abbildung 4.14.

Der Vergleich der Segmentierungsergebnisse des HMRF-UNet auf künstlichen und realen µCT-Daten zeigt, dass die Segmentierung auf den künstlich generierten Daten besser funktionierte. Jedoch scheint sich die Segmentierungsgüte auf den realen Daten ähnlich zu verhalten wie für die künstlich generierten Daten mit geringem Kontrast. In diesen Bildern sind manche Stege genauso wie für den realen Datensatz nur schwer vom Porenbereich zu unterscheiden. Daher könnte es sein, dass manche der künstlich erzeugten Daten die realen Daten nicht optimal abbilden. Da sich die Voxelintensitäten in diesem Fall nicht groß vom Porenbereich unterscheiden, werden diese vermutlich in vielen Fällen der Luft-Klasse zugewiesen, da der Wert besser zur Verteilung der Luft-Klasse passt. Ein solches Problem kann von überwachten Lernverfahren behoben werden, was sich auch in den Experimenten mit dem Finetuning in Abschnitt 4.4.6 bestätigt hat. Für die schlechte Segmentierung der feinen Stege könnte jedoch auch der Nachbarschafts-Loss verantwortlich sein. Da die dünnen Stege oft nur einen Voxel breit sind, besitzen diese meist lediglich zwei Nachbarn, die ebenso der PU-Klasse angehören. Da in diesem Fall der Großteil, nämlich sechs Nachbarn, der Luft-Klasse angehört, bestraft der Nachbarschafts-Loss in diesem Fall die Zuweisung in die PU-Klasse. Auch das von Nan et al. [128] beschriebene Problem der redundanten Klassen könnte die Qualität der Segmentierung beeinflussen. Manche der Voxel liegen vermutlich genau zwischen den zwei Verteilungen und springen daher potenziell während des Trainings ständig zwischen den beiden Klassen hin und her.

Ein weiterer Grund für die schlechteren Ergebnisse auf dem realen Datensatz könnten partielle Volumeneffekte sein, die möglicherweise auch nur im realen µCT-Datensatz vorhanden sind. Diese Effekte könnten besonders stark im Bereich rund um die dünnen PU-Stege ausfallen und damit die teilweise schlechten Segmentierungsergebnisse in diesen Bereichen erklären. Zusätzlich befindet sich in den realen Bildern sehr viel Rauschen, das künstlich hohe Signale innerhalb der Porenklasse erzeugt. Durch die hohen Signale könnten die berechneten Verteilungswerte möglicherweise verfälscht werden. Nichtsdestotrotz sind die Segmentierungsergebnisse auf dem realen *RealPUFoam*-Datensatz visuell sehr gut und die meisten Bereiche scheinen korrekt segmentiert zu sein. Möglicherweise könnte auch hier ein Nachtrainieren mittels annotierter Groundtruth-Segmentierungen dieser realen Strukturen helfen.

4.4.8 Anwendung auf Lungen-MRT-Datensatz

Nachdem das HMRF-UNet gute Segmentierungsergebnisse für den realen PU-Schaum geliefert hat, soll nun als Nächstes das entwickelte Modell auch mit dem Lungen-MRT-Datensatz getestet werden. Hierzu wurden HMRF-UNet-Modelle mit verschiedenen Nachbarschaftsfunktionen und -gewichten sowie σ_{thresh}-Werten trainiert und anschließend mit den drei vorhandenen Groundtruth-Segmentierungen getestet. Für das Training des HMRF-UNet wurden die Bilder zunächst jeweils auf den Bereich [0, 1] normiert. Anschließend wurden jeweils alle Schichten der 3D-MRT-Bilder des Datensatzes extrahiert und diese auf eine Größe von 256×256 skaliert, indem im Frequenzbereich interpoliert wurde. Die extrahierten 2D Bilder wurden dann mit einer patientenweisen Aufspaltung und einer Aufteilung von $(70\%/15\%/15\%)$ in einen Trainings-, Validierungs- und Testdatensatz aufgeteilt. Alle Modelle wurden für 200 Epochen mit einer Lernrate von 1×10^{-5} trainiert. Wie schon für das EHOHMRF-Modell wurde hier ebenso nach fünf Klassen im Bild gesucht, da das Lungen-MRT-Bild eine höhere Komplexität im im Vergleich zum µCT-Bild aufweist.

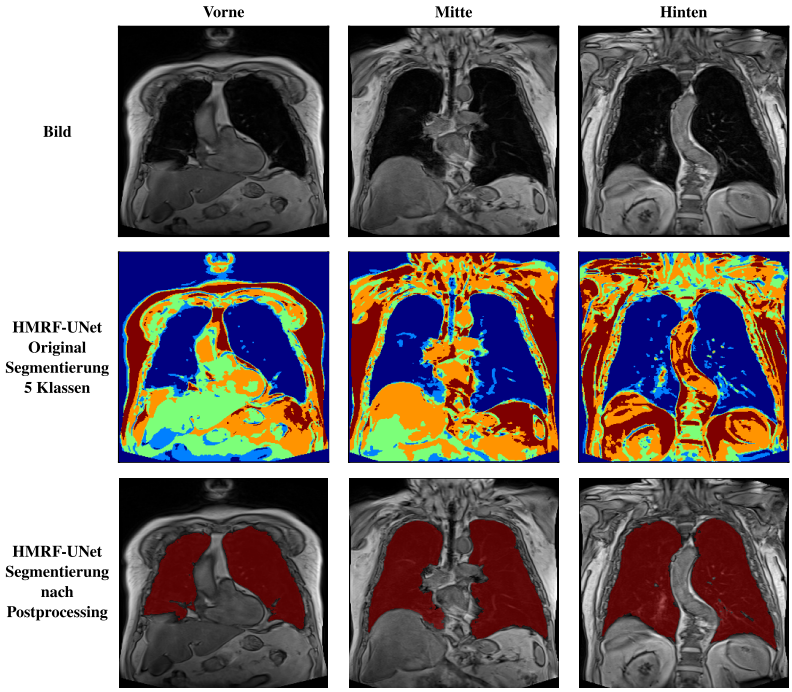


Abbildung 4.16: Darstellung von verschiedenen Schichten eines Beispielbildes des Lungen-MRT Datensatzes sowie den zugehörigen Segmentierungsergebnissen des HMRF-UNet. In der ersten Reihe ist das Bild, in der zweiten Reihe das ursprüngliche Segmentierungsergebnis des HMRF-UNets für fünf Klassen und in der dritten Reihe die generierte Lungenmaske nach dem Postprocessing des Segmentierungsergebnisses abgebildet.

In Abbildung 4.16 ist die Segmentierung eines HMRF-UNet mit normaler Potts-Nachbarschaft mit $\lambda_n = 0,3$ und $\sigma_{\text{thresh}} = 0,1$ für ein beispielhaftes Bild dargestellt. Die zweite Reihe zeigt die Vorhersage des HMRF-UNet mit fünf Klassen. Man erkennt, dass die fünf Klassen besser ausgenutzt werden, als beim EHOHMRF-Modell in Abbildung 4.6. Die finale Segmentierung nach einer Weiterverarbeitung der Segmentierungsmaske mit fünf Klassen ist in der dritten Reihe zu sehen. Für alle drei Schichten wurde ein gutes Segmentierungsergebnis erzielt. Helle Gefäße

Tabelle 4.9: Vergleich der DICE-Scores von HMRF-UNet-Modellen mit verschiedenen Nachbarschaftsarten und -gewichten sowie Schwellwerten σ_{thresh} für drei beispielhafte Lungen-MRT Bilder.

			DSC			
	λ_n	σ_{thresh}	**Beispiel 1**	**Beispiel 2**	**Beispiel 3**	**Ø**
Normal Potts	0,3	0,10	0,9309	0,8288	0,9570	0,9055
	0,5	0,10	0,9381	0,9502	0,9619	0,9501
	0,7	0,10	0,9410	0,9521	0,9617	**0,9516**
	0,5	0,05	0,9381	0,9507	**0,9622**	0,9503
	0,5	0,15	0,9714	0,9546	0,8608	0,9290
Weighted Banerjee	0,1	0,10	0,9623	0,9099	0,7107	0,8609
	0,5	0,10	**0,9730**	**0,9594**	0,8756	0,9360
	0,9	0,10	0,9445	0,9247	0,9023	0,9238
	0,5	0,05	0,9640	0,9131	0,7230	0,8667
	0,9	0,05	0,9701	0,9519	0,8720	0,9313

innerhalb der Lunge wurden zwar nicht segmentiert, jedoch könnte dies durch ein angepasstes Nachbearbeiten korrigiert werden.

In Tabelle 4.9 sind die berechneten DICE-Scores für verschiedene HMRF-UNet Modelle dargestellt. Es gab kein Modell, das in allen drei Fällen konstant am besten abschnitt. Während für die ersten beiden Beispiele das Modell mit gewichteter Banerjee-Nachbarschaft mit $\lambda_n = 0,5$ und $\sigma_{\text{thresh}} = 0,1$ die besten Segmentierungsergebnisse erzielte, schnitt es für das dritte Bild deutlich schlechter ab als fast alle Potts-Modelle. Hier erzeugte das Modell mit normaler Potts-Nachbarschaft und $\lambda_n = 0,5$ und $\sigma_{\text{thresh}} = 0,05$ die besten Segmentierungsergebnisse. Im Durchschnitt erzeugte das Modell mit normaler Potts-Nachbarschaft mit $\lambda_n = 0,7$ und $\sigma_{\text{thresh}} = 0,1$ die besten Segmentierungsergebnisse über alle drei Beispiele hinweg.

Zusätzlich wurde untersucht, ob sich der berechnete DICE-Score zwischen den einzelnen Schichten deutlich unterscheidet. Hierzu wurde jeweils der DICE-Score

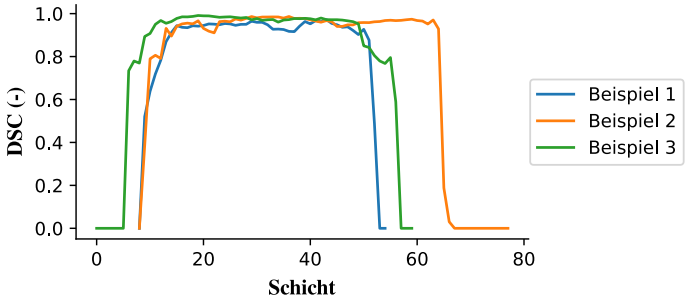


Abbildung 4.17: Darstellung des DICE-Scores pro Schicht für die Segmentierungen der drei Beispielbilder mit einem HMRF-UNet mit Potts Nachbarschaft und $\lambda_n = 0,5$ und $\sigma_{\text{thresh}} = 0,1$.

pro Schicht für jedes der drei Bilder für das HMRF-UNet mit normaler Potts-Nachbarschaft berechnet. Die Ergebnisse sind in Abbildung 4.17 dargestellt. Man erkennt, dass für alle drei Beispiele der DSC-Wert in den Randschichten geringer war, als in den mittleren Schichten. Ein Grund hierfür könnte einerseits sein, dass es deutlich mehr Bilder im Trainingsdatensatz gab, die Lunge beinhalteten. In den Randschichten ist jedoch meist keine Lunge vorhanden. Außerdem kommt es in den Randschichten zu stärkeren partiellen Volumeneffekten mit dem umliegenden Gewebe, wodurch hier eine eindeutige Zuordnung erschwert sein kann.

Bei der Betrachtung der Unterschiede zwischen der Groundtruth-Segmentierung und der Segmentierung des HMRF-UNet in Abbildung 4.18 fällt auf, dass der Großteil der Segmentierungsfehler am Rand der Lunge zu finden war. Ebenso wurden helle Gefäße von der Radiologin als Teil der Lunge segmentiert, während diese vom HMRF-UNet, aufgrund des stärkeren Signals dieser Gefäße, nicht als Teil der Lunge gezählt wurden. Ebenso verhält es sich für Bereiche mit Konsolidierungen, die eine pathologische Flüssigkeits- oder Zellvermehrung im Lungengewebe darstellen. Dies ist bspw. für das zweite MRT-Bild in der mittleren Schicht sichtbar. Hier wurde bei der rechten Lungenhälfte am unteren linken Ende ein großer Teil nicht segmentiert. Betrachtet man das MRT-Bild in Abbildung 4.16, so zeigt sich, dass sich an dieser Stelle eine Konsolidierung mit einem stärkeren Signal befindet. Außerdem fällt für das dritte MRT-Bild, genauso wie

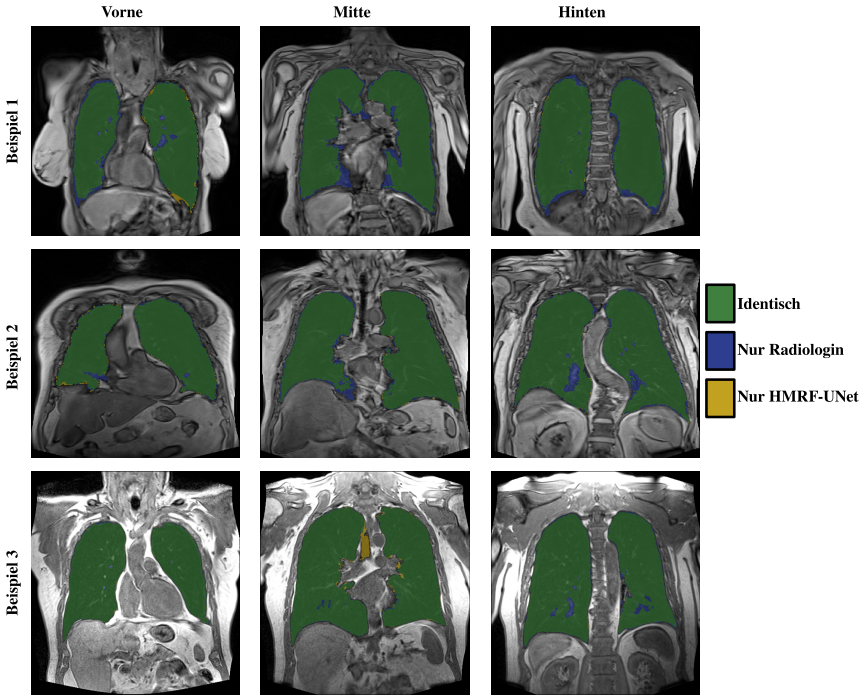


Abbildung 4.18: Vergleich der HMRF-UNet Lungensegmentierung mit Groundtruth Segmentierungen. Übereinstimmende Segmentierungsbereiche sind in Grün, Bereiche, die nur von der Radiologin segmentiert wurden, in Blau und Bereiche die nur vom HMRF-UNet segmentiert wurden in Gelb dargestellt.

für die EHOHMRF-Methode auf, dass die Trachea hier in der mittleren Schicht fälschlicherweise mit segmentiert wurde.

4.4.9 Bewertung des HMRF-UNet

Das entwickelte HMRF-UNet ist ebenso wie die EHOHMRF-Methode sehr flexibel einsetzbar. Durch den unüberwachten Loss werden keinerlei annotierte Groundtruth-Daten für das Training benötigt, was die Anpassbarkeit an neue Bilddaten vereinfacht. Wie im Abschnitt zuvor gezeigt wurde, funktioniert das

HMRF-UNet selbst für die deutlich komplexeren MRT-Bilder ähnlich gut, wie für die binären PU-Schaumbilder. Für beide Anwendungen konnten sehr gute Segmentierungsergebnisse erzielt werden. Ein großer Vorteil des HMRF-UNet gegenüber des EHOHMRF-Modells ist, dass die Segmentierungszeit deutlich kürzer ist. Während das EHOHMRF-Modell im Mittel zwischen 20 min mit der normalen Potts-Nachbarschaft und sogar fast 30 h mit der gewichteten Banerjee-Nachbarschaft für ein 3D-Lungenbild benötigt hat, segmentierte das HMRF-UNet ein Bild, unabhängig von der Nachbarschaftsfunktion in durchschnittlich 6 s. Dies bietet eine enorme Zeitersparnis, die insbesondere bei der Auswertung von großen Studien relevant sein könnte. Aktuell wurde das HMRF-UNet lediglich auf 2D-Daten trainiert, wodurch eine Segmentierung der Lungen-MRT Bilder nur schichtweise möglich ist. Die Anpassung des U-Net, sodass es dreidimensionale Bilder verarbeiten kann, könnte sehr hilfreich sein, da somit auch die dreidimensionale Nachbarschaft eines Voxels berücksichtigt werden kann. Die Nachbarschaftsbeziehung von benachbarten Voxeln zweier Schichten wird aktuell durch die schichtweise Anwendung nicht bei der Klassenzuweisung eines Voxels berücksichtigt.

Eine weitere Stärke des HMRF-UNet ist, dass es, wie für den PU-Schaum-Datensatz gezeigt wurde, für das Vortrainieren von ML-Modellen verwendet werden kann, was die benötigte Anzahl an annotierten Groundtruth-Daten deutlich reduzieren kann. Dies ist insbesondere für 3D-Bilddaten wie die meisten CT- und MRT-Bilder sehr relevant, da hier die manuelle Segmentierung sehr zeitintensiv ist.

Eine Limitation der Methode ist, dass diese zur Segmentierung des Bildes nur dessen Grauwerte verwendet. Dadurch werden häufig Objekte, die ähnliche Grauwerte besitzen, der gleichen Klasse zugeordnet. Dies wurde insbesondere für die Segmentierung der Lunge beobachtet, wo der Hintergrund meist der selben Klasse, wie der Lunge zugeordnet wurde. Möglicherweise könnte es hierzu hilfreich sein, dem Modell noch weitere Daten zur Verfügung zu stellen. Ebenso wäre es möglich, hier mit einer Art Strafterm zu arbeiten, der die Zuweisung von zwei

räumlich weit entfernten Objekten in die gleiche Klasse bestraft. Das Nachtrainieren mit annotierten Daten könnte möglicherweise das Problem der gleichen Klassenzuweisung verschiedener Objekte lösen. Zusätzlich könnte es aber auch grundsätzlich zu einer verbesserten Segmentierung führen, die dann auch Gefäße und Konsolidierungen in der Lunge mit in die Lungenmaske einschließt.

Die anhand der beiden Datensätze des realen PU-Schaums und der Lungen-MRT-Bilder durchgeführten Validierungsstudien waren aufgrund der fehlenden Groundtruth-Daten für den PU-Schaum und aufgrund geringer Mengen an Groundtruth-Daten für die MRT-Bilder nur eingeschränkt repräsentativ. Daher wären weitere Untersuchungen sinnvoll, die das HMRF-UNet anhand von größeren Mengen an Groundtruth-Daten validieren. Ebenso könnte die Validierung an bekannten, öffentlich verfügbaren Testdatensätzen eine interessante Benchmarking-Option darstellen.

4.5 Algorithmische Lungensegmentierung

Die algorithmische Lungensegmentierung wurde von der Arbeit von Kohlmann et al. [111] motiviert und basiert auf einer implementierten Adaption dieser Methode von Konietzke [129].

4.5.1 Segmentierungs-Pipeline

Einen Überblick über die einzelnen Schritte der algorithmischen Lungensegmentierung liefert Abbildung 4.19. Die Segmentierung der Lunge wurde schwerpunktmäßig im koronalen VIBE-Bild durchgeführt. In einem ersten Schritt wurde mithilfe eines Schwellwertverfahrens eine grobe Maske des Thorax erstellt. Der Schwellwert entspricht der Position des ersten Tals des Histogramms aller Signale im Bild und damit der Trennung zwischen Vorder- und Hintergrund im MRT-Bild. Anschließend wurde die mittlere Schicht des Volumens gewählt und innerhalb der Thoraxmaske in dieser Schicht ein zweiter Schwellwert berechnet,

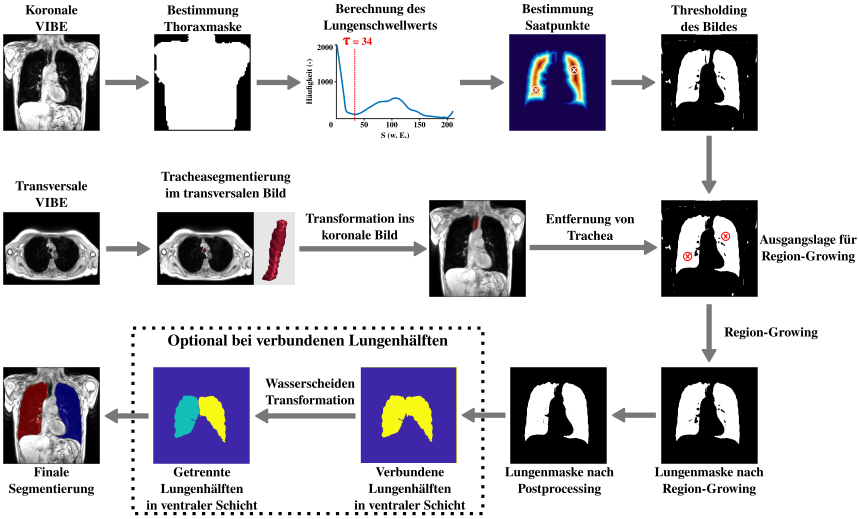


Abbildung 4.19: Darstellung der einzelnen Schritte der algorithmischen Lungensegmentierung. Zur Vereinfachung sind die Schritte jeweils nur für eine 2D-Schicht dargestellt, obwohl die Schritte immer auf dem kompletten 3D-Bild durchgeführt werden.

der wieder der Position des ersten Tales entspricht. Der ermittelte Schwellwert wurde auf den ermittelten Thoraxbereich in der mittleren Schicht angewendet und alle Voxel unterhalb dieses Schwellwerts als mögliche Lungenvoxel gewählt. Die zwei größten zusammenhängenden Objekte wurden dann als die Lunge gewählt. Falls das Thresholding nicht zwei ungefähr gleich große Objekte lieferte, wurde angenommen, dass beide Lungenhälften zusammengeschmolzen sind und die Lunge wurde mithilfe einer Wasserscheidentransformation in zwei Hälften getrennt. Bestehende Löcher in dieser groben Maske wurden aufgefüllt. Für den nächsten Schritt wurde zunächst eine euklidische Distanzkarte für je beide Lungenhälften berechnet, indem für jeden Voxel innerhalb der Maske jeweils die kürzeste Distanz zum Rand der Lungenhälfte berechnet wurde. Basierend auf dieser Karte wurde dann jeweils der Punkt mit der maximalen Distanz zum Rand gewählt, um dann als Saatpunkt für die anschließende Flutfüllung (engl. Floodfill) zu dienen.

Das MRT-Signal der Trachea ist ebenso wie das Lungensignal sehr dunkel. Daher besteht die Gefahr, dass bei einer Floodfill Operation die Segmentierung in die Trachea überläuft. Um dies zu verhindern, wurde die Trachea in einem Vorschritt segmentiert und anschließend von der Floodfill Operation ausgeschlossen. Da die Kanten der Trachea in x-Richtung im koronalen Bild schlecht zu erkennen sind, wurde die Trachea stattdessen im transversalen VIBE-Bild segmentiert. Zunächst wurde auch hier ein Schwellwert zur groben Segmentierung der Lunge gewählt, der im ersten Tal des Histogramms liegt. Alle Voxel unterhalb dieses Schwellwerts wurden als mögliche Lungen- und Tracheavoxel definiert. Die zwei größten Objekte wurden als Lunge definiert und die Flächenveränderung über die Schichten der beiden Lungenhälften berechnet. Die Schicht, in der sich die relative Flächenänderung am stärksten änderte, entspricht dem Start der Lungen (oberes Ende) und diente als Startschicht für die Tracheasegmentierung, da hier die Trachea sehr gut vom umliegenden Gewebe unterschieden werden kann. In dieser Startschicht wurde die Trachea als ein rundes Objekt in der Mitte des Bildes mit einer Fläche zwischen 30–300 Voxeln definiert. Da sich die Fläche und die Position der Trachea zwischen zwei benachbarten Schichten kaum verändert, wurden zunächst alle verbundenen Objekte in Kopfrichtung gewählt, die $\pm 40\,\%$ der Tracheafläche aus der nachfolgenden Schicht besaßen. Dies wurde so lange wiederholt, bis kein tracheaähnliches Objekt mehr gefunden werden konnte. Anschließend wurde mit dem gleichen Vorgehen in entgegengesetzter Richtung (fußwärts) jedes verbundene Objekt gewählt, das keine Flächenveränderung von mehr als $40\,\%$ aufwies und zwischen 30–300 Voxeln groß war. Dies wurde so lange durchgeführt, bis die Bifurkation, also die Aufspaltung der Trachea in die zwei Hauptbronchien, stattfand.

Die damit erhaltene Trachea-Maske wurde in das koronale Bild transformiert und dann wie oben erwähnt von der Floodfill Operation ausgeschlossen. Bei der Floodfill Operation wurden nun ausgehend von den Saatpunkten nach und nach benachbarte Voxel hinzugefügt, sofern diese unterhalb des zuvor bestimmten

Schwellwerts lagen. Abschließend fand noch eine Nachbearbeitung der Lungenmaske statt, bei der, wenn nötig, die beiden Lungenhälften per Wasserscheidentransformation getrennt und bestehende kleine Löcher in der Lungenmaske gefüllt wurden.

4.5.2 Optimierte Tracheasegmentierung

Die ursprüngliche Tracheasegmentierung aus Abschnitt 4.5.1 besitzt zwei wesentliche Schwachstellen. Einerseits ist der Schwellwert für die Trachea nach der Berechnung konstant und führt bei falscher Wahl zu schlechten Segmentierungsergebnissen. Andererseits wird die Trachea nur bis zur Bifurkation segmentiert, sodass die beiden Hauptbronchien nicht mit segmentiert werden. Dies führt häufig dazu, dass die Segmentierung während der Floodfill Operation in die Bronchien läuft und dadurch eine fehlerhafte Lungenmaske erzeugt wird. Um diese Limitationen zu beheben, wurde eine optimierte Tracheasegmentierung entwickelt.

Die Startschicht für die Tracheasegmentierung wurde wie in der alten Methode bestimmt, also als die Schicht der größten Volumenveränderung der Lungen. Der erste Schwellwert für die Trachea wurde bestimmt, indem das transversale VIBE-Bild zunächst mit einem Medianfilter gefiltert und anschließend mit der kMeans-Methode drei Cluster in diesem Bild gesucht wurden. Der Schwellwert wurde basierend auf Experimenten festgelegt als 1/3 des Zentrums des mittleren Clusters. Mithilfe dieses Schwellwertes wurden für jede Schicht mögliche Trachea-Objekte bestimmt, deren Signale unterhalb des Schwellwerts lagen. Zunächst wurde nach der Trachea in der Startschicht gesucht. Die zuvor bestimmten Trachea-Objekte wurden anhand ihrer geometrischen Eigenschaften bewertet, um zu schauen, ob diese der Trachea entsprechen. Hierzu wurde die Rundheit und die Fläche der Trachea-Objekte berechnet. Trachea-Objekte mit einer Rundheit geringer als 0,6 oder einer Fläche mit weniger als 10 Voxeln bzw. mehr als 300 Voxeln wurden von der Liste der möglichen Trachea-Objekte entfernt.

Nach der Festlegung auf ein Trachea-Objekt in der Startschicht wurden zunächst verbundene Objekte in den vorherigen Schichten gesucht. Als mögliche Trachea-Objekte wurden hier nur Objekte gewählt, die mindestens eine Rundheit von 0,6 aufwiesen, zwischen 10–300 Voxel groß waren und bei denen das Flächenverhältnis zum Trachea-Objekt der nachfolgenden Schicht zwischen $1/3$–3 betrug. Gab es nach dem Filtern nur ein Trachea-Objekt, so wurde dieses als Teil der Trachea festgelegt und es wurde zur nächsten Schicht gewechselt, um dort nach einem passenden Trachea-Objekt zu suchen. Gab es nach dem Filtern dagegen kein oder mehr als ein Trachea-Objekt, so wurde der Schwellwert um eins erhöht und die Suche in der Schicht mit dem neuen Schwellwert erneut durchgeführt. Die Erhöhung des Schwellwerts wurde für eine maximale Anzahl an 50 Iterationen wiederholt. Sobald keine Trachea-Objekte mehr oberhalb der Startschicht gefunden wurden, wurde das gleiche Vorgehen für Schichten unterhalb der Startschicht wiederholt. Hier stoppte die Suche nach Trachea-Objekten jedoch zunächst in der Schicht der Bifurkation der Trachea. Nach der Bifurkation wurde die Suche nach Trachea-Objekten für beide Bronchien separat durchgeführt und endete, sobald keine weiteren Trachea-/Bronchien-Objekte mehr identifiziert werden konnten. Abschließend wurde die Trachea aus den einzelnen Trachea-Objekten zusammengesetzt und es wurden Löcher in der Trachea geschlossen.

Ein Vergleich zwischen der alten und der optimierten Tracheasegmentierung ist in Abbildung 4.20 dargestellt. Die optimierte Segmentierung umfasst dabei sowohl die blauen, als auch die grünen Bereiche, während die alte Segmentierung nur aus den blau markierten Bereichen besteht. Man erkennt, dass die optimierte Tracheasegmentierung eine verbesserte Tracheasegmentierung erzeugte, da insbesondere auch die beiden Hauptbronchien mit segmentiert wurden.

4.5.3 Anwendung auf Lungen-MRT-Datensatz

Die entwickelte algorithmische Lungensegmentierung wurde auf die drei Beispielbilder des Lungen-MRT-Datensatzes angewendet. Hierbei wurde einmal die

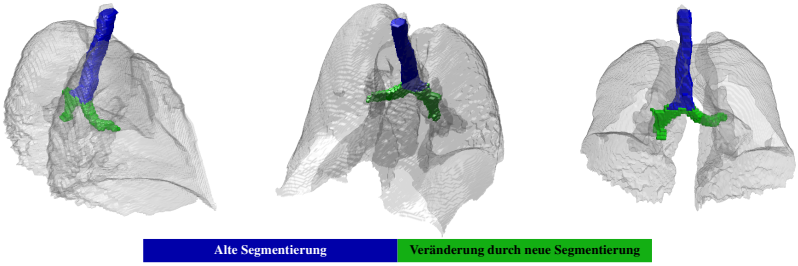


Abbildung 4.20: Darstellung der Unterschiede zwischen alter Tracheasegmentierung und optimierter Tracheasegmentierung. Die ursprüngliche Segmentierung ist in Dunkelblau dargestellt, die zusätzlich erzeugten Bereiche durch die optimierte Segmentierung sind in Grün dargestellt.

Version mit der alten Tracheasegmentierung und einmal die Version mit optimierter Tracheasegmentierung verwendet. In Abbildung 4.21 ist die Segmentierung eines Beispielbildes für eine vordere, eine mittlere und eine hintere Schicht dargestellt. Man erkennt, dass die Lunge in allen drei Schichten sehr gut segmentiert wurde. Die berechneten DICE-Scores für alle drei Beispielbilder sind in Tabelle 4.10 dargestellt. Hier zeigt sich, dass beide Varianten für alle drei Bilder sehr hohe DICE-Scores erzielten. Für beide Varianten waren insbesondere die mittleren DICE-Scores höher, als für die beiden entwickelten unüberwachten HMRF-basierten Modelle. Bezüglich des Einflusses der verwendeten Tracheasegmentierung ließ sich kein wirklicher Unterschied feststellen. Nur in einem der drei Fälle führte die optimierte Tracheasegmentierung im Vergleich zur alten Variante zu einem minimal besseren Segmentierungsergebnis. Ursache hierfür könnte sein, dass für diese drei beispielhaften Bilder die Hauptbronchien eine deutliche Abtrennung zum Lungenparenchym aufweisen, sodass auch bei fehlender Segmentierung der beiden Hauptbronchien keine Probleme während der Floodfill-Operation auftreten.

Betrachtet man die Visualisierung der Segmentierungsfehler in Abbildung 4.22, so zeigt sich, dass die Segmentierungsfehler je nach Bild unterschiedlich ausgefallen sind. Während für das Beispiel 1 die Segmentierungsfehler sehr diffus und über die ganze Lunge verteilt waren, beschränkten sich die Segmentierungsfehler

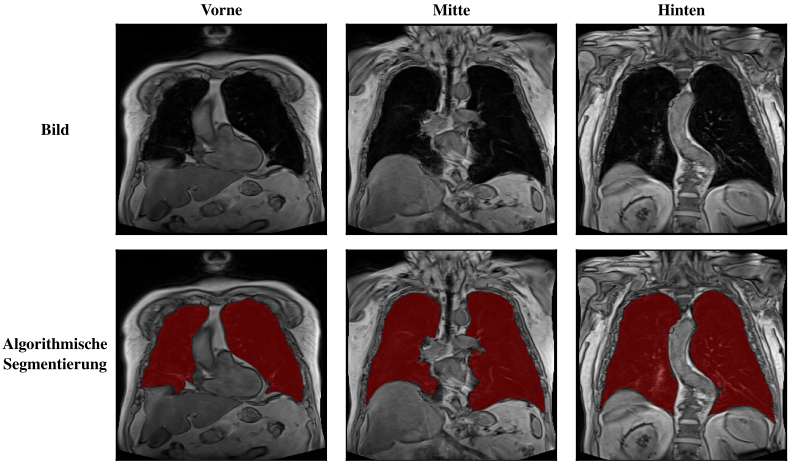


Abbildung 4.21: Darstellung von verschiedenen Schichten eines Beispielbildes sowie den zugehörigen Segmentierungsergebnissen der algorithmischen Segmentierungsmethode. In der ersten Reihe ist das Bild und in der zweiten Reihe die generierte Lungenmaske abgebildet.

Tabelle 4.10: Vergleich der DICE-Scores der algorithmischen Lungensegmentierung mit alter und neuer Tracheasegmentierung für drei Beispielbilder.

	DSC			
	Beispiel 1	**Beispiel 2**	**Beispiel 3**	**∅**
Alte Tracheasegmentierung	**0,9730**	0,9649	**0,9679**	**0,9686**
Neue Tracheasegmentierung	**0,9730**	**0,9650**	**0,9679**	**0,9686**

beim zweiten Bild auf den unteren Lungenrand der rechten Lunge. Beim dritten Beispielbild wurden fälschlicherweise insbesondere Teile der Trachea und des Herzens als Teil der Lunge segmentiert. Ähnliches konnte für die mittlere Schicht des ersten Beispielbildes beobachtet werden. Auch die algorithmische Segmentierung scheint Schwierigkeiten mit der Segmentierung von Konsolidierungen zu besitzen, da die Konsolidierung in der mittleren Schicht des zweiten Bildes ebenso wie bereits beim HMRF-UNet nicht segmentiert wurden.

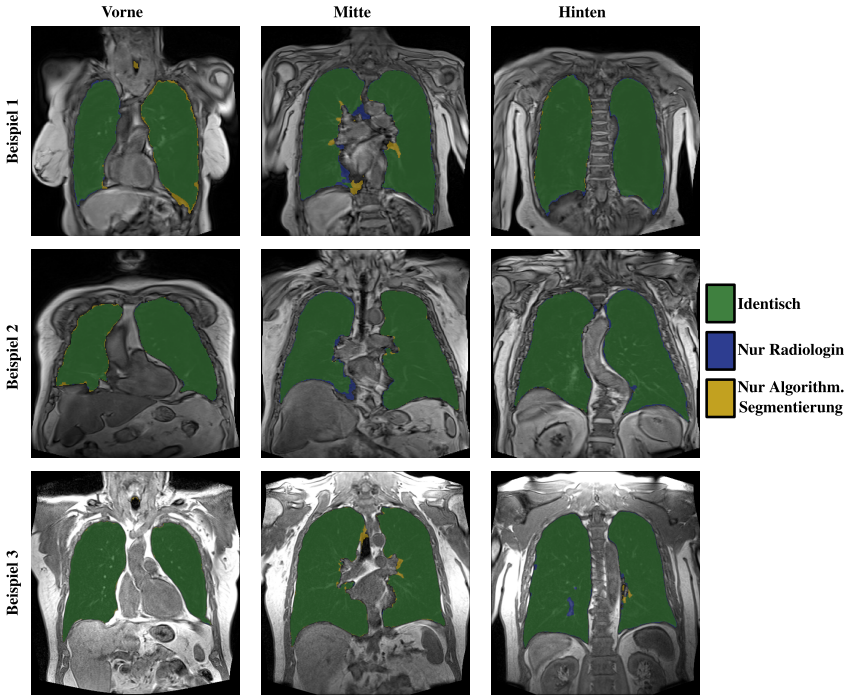


Abbildung 4.22: Vergleich der algorithmischen Lungensegmentierung mit Groundtruth Segmentierungen. Übereinstimmende Segmentierungsbereiche sind in Grün, Bereiche, die nur von der Radiologin segmentiert wurden, in Blau und Bereiche die nur von algorithmischen Segmentierungsmethode segmentiert wurden in Gelb dargestellt.

4.5.4 Bewertung der algorithmischen Lungensegmentierung

Die verwendete algorithmische Lungensegmentierung zeigte eine sehr gute Performance für alle drei betrachteten Beispielbilder. Während bei beiden HMRF-basierten Modellen meist mindestens eines der Beispielbilder nicht so gut segmentiert werden konnte, erreichte die algorithmische Segmentierung konstant hohe DICE-Scores oberhalb von 0,9650. Dabei schien die Verwendung der optimierten

Tracheasegmentierung zu keiner Verbesserung, aber auch zu keiner Verschlechterung des Segmentierungsergebnisses zu führen. Bezüglich der Segmentierungszeit liegt die algorithmische Methode mit einer durchschnittlichen Segmentierungszeit von ca. 60 s zwischen den beiden HMRF-basierten Methoden.

Die hier beobachteten Probleme bei der Segmentierung ähneln den Problemen der HMRF-basierten Methoden. Insbesondere Konsolidierungen und allgemein helle Bereiche innerhalb der Lunge schienen zu fehlerhaften Segmentierungen zu führen. Ebenso fiel auf, dass selbst mit der optimierten Tracheasegmentierung weiterhin Teile der Trachea nicht richtig ausgeschlossen wurden. Dies spricht dafür, dass entweder die Trachea nicht richtig segmentiert wurde oder dass es bei der Transformation der Tracheamaske in das koronale Bild zu Fehlern kam. Daher sollte an einer weiteren Optimierung der Tracheasegmentierung gearbeitet werden. Insbesondere Techniken, die die Auflösung des Bildes erhöhen, wie die Super-Resolution-Rekonstruktion [130] könnten hier zu Verbesserung der Bildqualität und damit zu besseren Segmentierungen führen.

Die algorithmische Lungensegmentierung zeigt, dass eine spezialisierte Methode, die genau auf das Problem zugeschnitten ist, Vorteile gegenüber allgemein anwendbaren Methoden besitzen kann. Durch die Berücksichtigung der Besonderheiten, wie die Exklusion der Trachea aufgrund deren Ähnlichkeit zur Lunge, können bessere Ergebnisse erzielt werden. Aufgrund der konstanteren und teilweise besseren Segmentierungsergebnisse der algorithmischen Lungensegmentierung wurde in Kapitel 5, falls nicht anders erwähnt, ausschließlich die algorithmische Segmentierung verwendet.

5 Quantifizierung: Robuste automatische Perfusionsquantifizierung

Die quantitative Auswertung von Perfusions-MRT Messungen der Lunge ist ein wichtiger Baustein für die Bewertung der Gesundheit von PatientInnen. Hierfür spielt die hypoxische pulmonale Vasokonstriktion eine wesentliche Rolle, da sie Lungen-Perfusion und Lungen-Ventilation verbindet. Die Perfusionsquantifizierung ermöglicht insbesondere auch ein Monitoring des Gesundheitszustands der PatientInnen über mehrere zeitlich verteilte Messungen hinweg. Hiermit lässt sich der Krankheitsverlauf, aber auch der Therapieerfolg anhand von vergleichbaren Zahlen messen. Das DCE-MRT ist eine der wichtigsten Perfusions-MRT Techniken, da es mit weit verbreiteten MRT-Sequenzen auskommt und als zusätzliches Gerät lediglich eine Kontrastmittelinjektionspumpe notwendig ist, die jedoch ebenso breit verfügbar ist. Zusätzlich existiert bereits eine Menge Literatur für die erfolgreiche Anwendung des DCE-MRT in anderen Organen. Jedoch ist die Anzahl an RadiologInnen, die DCE-MRT Aufnahmen der Lunge auswerten können relativ gering, da das Lungen-MRT allgemein noch nicht breit angewendet wird [40]. Die fehlende Expertise führt dazu, dass an Standorten, an denen eine DCE-MRT Messung durchgeführt werden könnte, oft keine Auswertungen möglich sind. Um einen breiteren Einsatz des DCE-MRT zu ermöglichen, ist die Entwicklung von vollautomatischen Perfusionsquantifizierungssystemen sehr interessant. Dieses Kapitel beschäftigt sich mit der Entwicklung eines solchen vollautomatischen Systems zur Perfusionsquantifizierung. Die Entwicklung dieser Perfusionsquantifizierungs-Pipeline lässt sich, wie in Abbildung 5.1 visualisiert, in

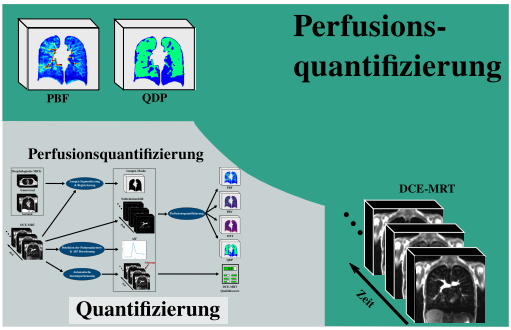


Abbildung 5.1: Einordnung der Perfusionsquantifizierung in den informationsgetriebenen Bildanalysezyklus.

den unteren linken Quadranten der Quantifizierung des in Kapitel 3 vorgestellten informationsgetriebenen Bildanalysezyklus einordnen.

Nachfolgend wird zunächst der aktuelle Forschungsstand im Bereich Perfusionsquantifizierung von DCE-MRTs vorgestellt sowie ein Überblick über die dieser Arbeit zugrundeliegenden Vorarbeiten gegeben. Anschließend werden die einzelnen Bestandteile der Perfusionsquantifizierungs-Pipeline im Detail vorgestellt. Das Kapitel schließt mit einer Validierung der entwickelten Pipeline in einer monozentrischen und einer multizentrischen Kohorte. Dabei wird ein besonderer Fokus auf die Generierung von vergleichbaren und konsistenten Ergebnissen in beiden Kohorten gelegt.

5.1 Stand der Forschung

Die Quantifizierung der Perfusion basiert auf der von Meier und Zierler [51] definierten Indikator-Verdünnungs-Theorie. Hierbei wird zunächst ein geschlossenes Einkammersystem mit einem einzigen Eingang und einem einzigen Ausgang definiert, das ein Volumen enthält, welches mit einer konstanten Fließgeschwindigkeit durch das System fließt. Aufgrund von Verzweigungen im System besitzen verschiedene Partikel eine unterschiedliche Zeit, um vom Ein- zum Ausgang zu

gelangen. Zusätzlich werden vier Annahmen getroffen: (i) Die Verteilung der Fließgeschwindigkeiten aller Partikel ist zeitlich unveränderlich, (ii) der Fluss von Indikator-Partikeln ist repräsentativ für den Fluss des kompletten Flüssigkeitsvolumens, (iii) das System besitzt keine Stauseen und (iv) es existiert keine Rezirkulation des Indikators. Die Methode benötigt in einem solch definierten System einen Indikator, der am Eingang injiziert wird und dessen Konzentration am Ausgang des Systems als Funktion der Zeit gemessen wird. [51]
Østergaard et al. [131] verwendeten diese Vorarbeiten und definierten die Konzentration des arteriellen Inputs als $\mathbf{C_a}(t)$ (im Folgenden meist auch als AIF bezeichnet) und die Konzentration des venösen Outputs als $\mathbf{C_v}(t)$. Der venöse Output lässt sich dann berechnen durch

$$\mathbf{C_v}(t) = \mathbf{C_a}(t) * \mathbf{h}(t) = \int_0^t \mathbf{C_a}(\tau)\mathbf{h}(t-\tau)\,dt, \tag{5.1}$$

wobei $*$ der Faltungsoperation und $\mathbf{h}(t)$ der Wahrscheinlichkeitsdichtefunktion der Transitzeiten entspricht. Der Anteil des injizierten Indikators (engl. Tracer), der zum Zeitpunkt t noch in der Gefäßarchitektur vorhanden ist, kann durch die sogenannte Residuenfunktion $\mathbf{R}(t)$ beschrieben werden:

$$\mathbf{R}(t) = \left[1 - \int_0^t \mathbf{h}(\tau)\,d\tau\right]. \tag{5.2}$$

Die Residuenfunktion ist in der Theorie eine monoton fallende, positive Funktion mit einem Startwert $\mathbf{R}(0) = 1$. Damit lässt sich nun implizit der pulmonale Blutfluss PBF in Abhängigkeit der Konzentration eines Tracers $\mathbf{C_{VOI}}$ innerhalb eines Referenzvolumens (VOI, vom Englischen Volume of Interest) berechnen mit

$$\mathbf{C_{VOI}}(t) = \mathrm{PBF} \int_0^t \mathbf{C_a}(\tau)\mathbf{R}(t-\tau)\,d\tau\,. \tag{5.3}$$

Hieraus ergibt sich, dass zur Berechnung von PBF eine Entfaltung der Konzentration innerhalb des VOI mit der AIF nötig ist, um die Residuenfunktion und PBF zu isolieren. Der pulmonale Blutfluss ergibt sich theoretisch aus dem Wert dieses Terms zum Zeitpunkt $t = 0$, jedoch sollte aufgrund einer stattfindenden Dispersion

und Verzögerung der gemessenen AIF, stattdessen das Maximum dieses Terms als PBF verwendet werden [132, 133]. Das pulmonale Blutvolumen PBV lässt sich berechnen mit

$$\mathrm{PBV} = \frac{\int_0^\infty \mathbf{C}_{\mathbf{VOI}}(t)\,dt}{\int_0^\infty \mathbf{C}_{\mathbf{a}}(t)\,dt} \tag{5.4}$$

und die mittlere Transitzeit MTT ergibt sich aus dem Quotienten aus PBV und PBF

$$\mathrm{MTT} = \frac{\mathrm{PBV}}{\mathrm{PBF}}\,. \tag{5.5}$$

Zur Bestimmung von PBF und MTT ist daher immer auch die Bestimmung der Residuenfunktion nötig. Zur Bestimmung der Residuenfunktion wurden von Østergaard et al. zwei wesentliche Ansätze vorgeschlagen. Der erste Ansatz ist die modellabhängige analytische Entfaltung, bei der die Residuenfunktion zusammen mit MTT und PBF mithilfe von $\mathbf{C}_\mathbf{a}$ und $\mathbf{C}_\mathbf{VOI}$ mit einem Ansatz der kleinsten Quadrate angepasst wird. Der zweite Ansatz ist die modellunabhängige, nichtparametrische Entfaltung, bei der nur die Residuenfunktion und PBF gleichzeitig bestimmt werden. Zur Entfaltung wird die Fouriertransformation verwendet und es ergibt sich nachfolgende Formel zur Berechnung der Residuenfunktion

$$\mathbf{R}(t) = \frac{1}{\mathrm{PBF}}\mathcal{F}^{-1}\left\{\frac{\mathcal{F}\{\mathbf{C}_{\mathbf{VOI}}(t)\}}{\mathcal{F}\{\mathbf{C}_\mathbf{a}(t)\}}\right\}\,. \tag{5.6}$$

Hierbei bezeichnet $\mathcal{F}$ die Fouriertransformation und $\mathcal{F}^{-1}$ die inverse Fouriertransformation. Die Berechnung in Gleichung (5.6) ist jedoch sehr anfällig gegenüber Rauschen und es handelt sich um ein inkorrekt gestelltes Problem. Daher ist für die Lösung des Problems eine Regularisierung notwendig. Mehrere Autoren schlagen hierzu die Verwendung der Singulärwertzerlegung (SVD, vom englischen Singular Value Decomposition) mit einer zusätzlichen Regularisierung vor [52, 55, 131]. Hierzu wird Gleichung (5.3) zunächst diskretisiert, weil die Signale

immer nur zu diskreten Zeitpunkten gemessen werden und angenommen werden kann, dass sie für diese Zeitpunkte konstant sind.

$$\Delta t \begin{bmatrix} C_a(t_1) & 0 & \cdots & 0 \\ C_a(t_2) & C_a(t_1) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ C_a(t_{\max}) & C_a(t_{\max}-1) & \cdots & C_a(t_1) \end{bmatrix} \cdot \begin{pmatrix} R(t_1) \\ R(t_2) \\ \vdots \\ R(t_{\max}) \end{pmatrix} = \begin{pmatrix} C(t_1) \\ C(t_2) \\ \vdots \\ C(t_{\max}) \end{pmatrix} \tag{5.7}$$

Definiert man nun die Matrix $\mathbf{A} = \Delta t \mathbf{C_a}(t)$, den Vektor $\mathbf{x} = \mathbf{R}(t)$ und den Vektor $\mathbf{b} = \mathbf{C}(t)$, so lässt sich mit der Methode der kleinsten Quadrate ein Minimierungsproblem definieren [52]

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 \ . \tag{5.8}$$

Mithilfe der SVD kann man die Matrix $\mathbf{A}$ darstellen mit

$$\mathbf{A} = \mathbf{V}\mathbf{W}\mathbf{U}^\top \ , \tag{5.9}$$

wobei $\mathbf{V}$ und $\mathbf{U}$ jeweils orthogonale Matrizen sind und $\mathbf{W}$ eine Diagonalmatrix ist, deren Werte außerhalb der Diagonalen null sowie entlang der Diagonalen absteigend sortiert sind. [131, 134] Entspricht der Rang von Matrix $\mathbf{A}$ dem Wert k, ergibt sich die einheitliche Lösung der Gleichung mit

$$\mathbf{x} = \mathbf{V}\, diag\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \cdots, \frac{1}{\sigma_k}, \cdots, 0, \cdots, 0\right)\mathbf{U}^\top \mathbf{b} \ . \tag{5.10}$$

Aufgrund des inkorrekt gestellten Problems kommt es jedoch dazu, dass diese Minimierung insbesondere beim Auftreten von Rauschen zu falschen Ergebnissen führt, weshalb eine Regularisierung notwendig ist. Zwei weitverbreitete Verfahren sind die abgeschnittene SVD (tSVD, vom Englischen Truncated Singular Value Decomposition), bei der nur Singulärwerte bis zu einem Schwellwert in Abhängigkeit von $\sigma_{\max}$ verwendet werden [131] sowie die Tikhonov-Regularisierung in Standardform (SFTR) [52, 135, 136]. Hierzu wird der Optimierungsterm aus

Gleichung (5.8) mit einem zusätzlichen Term zur Stabilisierung/Regularisierung ergänzt [52]:

$$\min_{\mathbf{x}} \left\{ \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 + \mu^2 \|\mathbf{x}\|_2^2 \right\} , \tag{5.11}$$

wobei die Wahl des Regularisierungsparameter μ die Stärke der Regularisierung bestimmt. Die Lösung des Minimierungsproblems für ein gewähltes μ berechnet sich dann mit

$$\mathbf{x}_\mu = \mathbf{V}\, diag\left(\frac{\sigma_1}{\mu^2 + \sigma_1^2}, \cdots, \frac{\sigma_n}{\mu^2 + \sigma_n^2} \right) \mathbf{U}^\top \mathbf{b} . \tag{5.12}$$

Für die Festlegung des Regularisierungsparameters μ können verschiedene Techniken eingesetzt werden. Die am häufigsten verwendeten Techniken sind das L-Kurven Kriterium (LCC, vom Englischen L-curve criterium) [136] und die generalisierte Kreuzvalidierung (GCV, vom Englischen general cross validation) [52]. Salehi Ravesh et al. [137] haben sowohl tSVD als auch SFTR-LCC, sowie zwei angepasste Varianten von SFTR-LCC getestet und festgestellt, dass sowohl tSVD und die beiden angepassten SFTR-LCC Varianten gut für die Berechnung von $\mathbf{R}(t)$ in der Lunge funktionieren, während die normale SFTR-LCC Variante nicht geeignet war. In der Arbeit von Konietzke [129] wurde festgestellt, dass SFTR-LCC zu größeren Streuungen der PBF-Werte im Vergleich zu tSVD führt, weshalb tSVD für die Erzeugung von vergleichbaren Ergebnissen präferiert werden sollte.

Nach der erfolgreichen Berechnung der Residuenfunktion mithilfe einer der zuvor definierten Verfahren, lässt sich final der pulmonale Blutfluss PBF als das Maximum der Residuenfunktion mit der Einheit [ml/100ml/min] definieren und das Blutvolumen PBV als die Fläche unter der Residuenfunktion mit der Einheit [ml/100ml]. Die mittlere Transitzeit berechnet sich wie in Gleichung (5.5) als Quotient von PBF und PBV und besitzt standardmäßig die Einheit [s]. Eine Visualisierung der Zusammenhänge zwischen der Residuenfunktion $\mathbf{R}(t)$, PBF, PBV und MTT ist in Abbildung 5.2 dargestellt.

Eine Perfusionsquantifizierung ohne die Messung der AIF ist nicht möglich, da diese wichtige Informationen liefert, wie das Kontrastmittel im zu untersuchenden

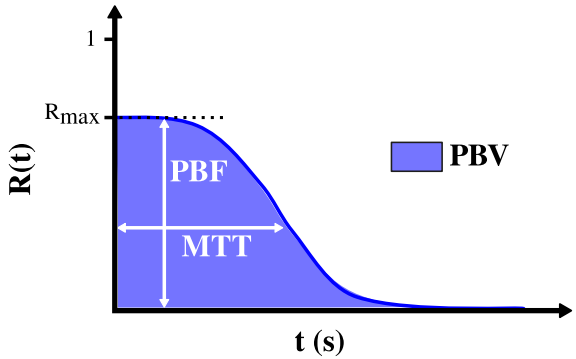


Abbildung 5.2: Darstellung der Herleitung der quantitativen Perfusionsparameter PBF, PBV und MTT anhand der bestimmten Residuenfunktion $R(t)$.

Bereich ankommt. Daher ist die richtige Messung der AIF für eine gute Perfusionsquantifizierung essenziell. Einen guten Überblick über die Einflüsse der AIF liefert die Arbeit von Calamante [138]. Das Wichtigste ist die AIF so nah wie möglich am Bereich, in dem die Perfusion quantifiziert werden soll, zu messen. Im Optimalfall sollte die Messung außerdem innerhalb einer Arterie geschehen, jedoch wurde gezeigt, dass auch eine (ungewollte) Messung in der Umgebung der Arterie zu guten quantitativen Ergebnissen führen kann. [138]

Zur Berechnung der Kontrastmittelkonzentration im zu untersuchenden Gewebe $\mathbf{C}(t)$ aus dem MRT Signal $\mathbf{S}(t)$ gibt es verschiedene Ansätze. Grundsätzlich wird meist ein linearer Zusammenhang zwischen dem MRT-Signal und der Kontrastmittelkonzentration angenommen [40]. Der einfachste Ansatz zur Berechnung von $\mathbf{C}(t)$ ist die Berechnung mittels eines Subtraktionsbildes, indem von jedem Zeitschritt jeweils das Ausgangssignal vor der Kontrastmittelankunft subtrahiert wird [53, 139]:

$$\mathbf{C}(t) \approx \mathbf{S}(t) - \mathbf{S_0}(t) \,, \tag{5.13}$$

wobei $\mathbf{S_0}(t)$ dem Signal vor der Kontrastmittelankunft entspricht. Eine alternative Möglichkeit ist die zusätzliche Normierung dieser Differenz mit dem Ausgangssignal [137]:

$$\mathbf{C}(t) \approx \frac{\mathbf{S}(t) - \mathbf{S_0}(t)}{\mathbf{S_0}(t)} \ . \tag{5.14}$$

Das schwache Signal des Lungenparenchyms [36], das im Grundlagenkapitel bereits erwähnt wurde, erzeugt das Problem, dass das Signal vor der Kontrastmittelankunft auch einen großen Anteil an Rauschen enthält, wodurch eine Normierung mit dem Ausgangssignal zu Artefakten führen kann. Daher wurde in der Arbeit von Konietzke von der Verwendung von Gleichung (5.14) abgeraten und eine Verwendung von Gleichung (5.13) empfohlen. [129]

Die hier beschriebene Perfusionsquantifizierungs-Technik wurde zunächst nur für die Quantifizierung der Perfusion in Gehirn-MRT Messungen verwendet [132, 140]. Einige Jahre später wurde sie auch erstmals für die Quantifizierung der Lungenperfusion eingesetzt [55, 141]. Aufbauend darauf gab es in den folgenden Jahren immer mehr Publikationen, die sich mit der Perfusionsquantifizierung in DCE-MRT Bildern der Lunge beschäftigt haben [44, 48, 53, 56, 57, 142–149].

Eine wesentliche Weiterentwicklung im Bereich der Perfusionsquantifizierung war die Einführung einer weiteren quantitativen Perfusionsmetrik, den Perfusionsdefekten in Prozent (QDP), die ursprünglich für die prozentuale Angabe von Ventilationsdefekten entwickelt [48, 150, 151] und für die Quantifizierung von Perfusionsdefekten im DCE-MRT weiterenwickelt wurde [56, 57]. Zur Berechnung von QDP können verschiedene Clustering- und Schwellwert-Verfahren eingesetzt werden, die die Lunge in einzelne Bereiche einteilen. In der Arbeit von Konietzke et al. [56] wurde hierfür die Residuen-Karte zum Zeitpunkt der größten Kontrastverstärkung $\mathbf{R}_{\text{max}}$ in einzelne Bereiche eingeteilt. In einer Arbeit von Schiwek et al. wurden vier verschiedene Möglichkeiten (Otsu Clustering, K-Means Clustering, Texturanalyse und 80. Perzentilansatz) miteinander verglichen. Hierbei zeigte die Variante mit Otsu Clustering die höchste Korrelation mit dem MRT-Perfusionsscore von Eichinger et al. [45], starke Korrelationen mit quantitativen CT Werten und eine stärkere Korrelation als PBF. [57]

Die Fleischner Society hat in einem Positionspapier die Relevanz von multizentrischen Vergleichsstudien von MRT-Techniken mit klinischen Alternativen für verschiedene Lungenerkrankungen herausgestellt [152]. Bisherige DCE-MRT Studien haben sich jedoch bisher meist nur mit kleinen [55, 56, 141–143, 146, 147] oder großen [48, 57, 144] monozentrischen Studien beschäftigt. Die wenigen multizentrischen Studien beinhalteten entweder nur zwei Zentren mit jeweils sehr kleinen Subkohorten [145, 148] oder viele verschiedene Zentren, die aber ebenso nur aus sehr kleinen Subkohorten bestanden [149]. Daher sind solche Studien nicht geeignet, um statistisch relevante Aussagen über die Robustheit des DCE-MRT im multizentrischen Einsatz zu liefern, wie beispielsweise die Untersuchung des Einflusses verschiedener Scanner. Ein weiteres Problem ist, dass keine der bisherigen Methoden eine vollautomatische Auswertung ohne menschlichen Einfluss ermöglicht, da meist entweder die Lunge manuell segmentiert oder die AIF manuell bestimmt werden musste. Die Verfügbarkeit von vollautomatischen quantitativen Methoden ist jedoch ein wichtiger Baustein für die Überwachung des Therapieerfolgs bei Lungenerkrankungen [18].

Um diese zwei Forschungslücken auszufüllen, soll in dieser Arbeit eine vollautomatische und robuste Perfusionsquantifizierungs-Pipeline gebaut und mithilfe einer großen multizentrischen Studie getestet werden.

5.2 Halbautomatische Perfusionsquantifizierung

Die Arbeit baut auf einer Vorarbeit von Konietzke [129] auf, in der eine halbautomatische Pipeline zur Perfusionsquantifizierung entwickelt wurde, die jedoch bisher nur in monozentrischen Studien eingesetzt wurde [56, 57]. Einen Überblick über die Ausgangslage dieser Arbeit gibt Abbildung 5.3, die die einzelnen Schritte der in Matlab implementierten Perfusionsquantifizierungs-Pipeline von M. Konietzke zeigt. Im ersten Schritt werden die Bilder, die im medizinischen

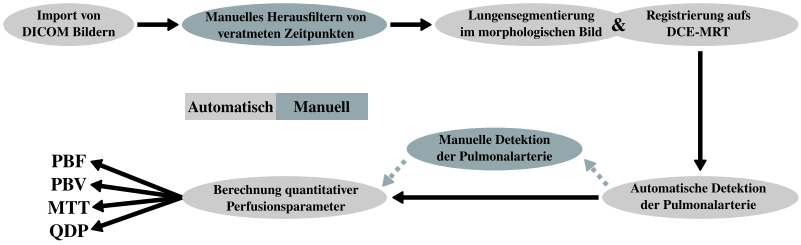


Abbildung 5.3: Überblick über die Perfusionsquantifizierungs-Pipeline von M. Konietzke, die als Ausgangslage für diese Arbeit dient.

DICOM-Format vorliegen, in Matlab importiert. Anschließend findet das manuelle Herausfiltern von veratmeten Zeitpunkten durch einen Anwendenden statt. Der dritte Schritt ist die Segmentierung der Lunge in einem morphologischen MRT-Bild. Um anschließend die Lungenmaske für das DCE-MRT zu erhalten, findet zusätzlich ein Registrierungsschritt statt, indem das morphologische MRT-Bild auf das DCE-MRT Bild registriert wird und damit dann die erhaltene Lungenmaske transferiert werden kann. Der vierte Schritt ist die automatische Detektion der Pulmonalarterie, die nötig ist, um darin die AIF zu bestimmen. Da dies nicht immer robust funktioniert, wurde ein optionaler zusätzlicher Schritt definiert, der nur dann ausgeführt wird, wenn die Pulmonalarterie nicht automatisch detektiert werden konnte. In diesem Fall wird die Pulmonalarterie von einem Anwendenden manuell definiert. In einem finalen Schritt findet dann die Berechnung der quantitativen Pefusionswerte PBF, PBV, MTT und QDP statt. Die einzelnen Schritte dieser Pipeline werden in den nachfolgenden Unterkapiteln 5.2.1 - 5.2.6 nochmals näher beschrieben, da einige davon auch in dieser Arbeit verwendet wurden. Die Beschreibungen der halbautomatischen Pipeline basieren auf den Arbeiten von Schiwek/Konietzke et al. [56, 57, 129]. Adaptionen und Ergänzungen an dieser Pipeline werden dann in den Unterkapiteln 5.3 - 5.8 behandelt.

5.2.1 Einlesen von DICOMs

MRT-Bilder werden meist im DICOM-Format abgespeichert. Dabei wird jede Schicht eines MRT-Bildes als einzelne DICOM-Datei mit zugehörigen Metadaten gespeichert. Dies führt zu extrem vielen Einzeldateien pro MRT-Bild und zusätzlich zu einem hohen Speicherplatzbedarf. Daher wurden die Bilder vor der Auswertung eingelesen und aus den einzelnen Schichten ein drei- bzw. vierdimensionales Bild generiert, das als Matrix abgespeichert wird. Zusätzlich wurden die benötigten Bildmetadaten extrahiert und ebenso abgespeichert. Für die Perfusionsquantifizierung wurden drei verschiedene MRT-Bilder verwendet. Die zwei verwendeten morphologischen MRT-Bilder sind 3D T1-gewichtete Gradientenechos (3D T1W GRE) und wurden für die Lungensegmentierung benötigt. Beide Bilder wurden mit der VIBE-Sequenz aufgenommen, das eine in koronaler Ausrichtung und das andere in transversaler Ausrichtung. Die dritte verwendete MRT-Messung ist die DCE-MRT Messung, die mit der TWIST-Sequenz gemessen wurde.

5.2.2 Manuelle Atmungsentfernung

Dieser Schritt wird verwendet, um Zeitpunkte mit Atembewegung vor oder nach dem Kontrastmitteldurchlauf zu entfernen, da diese die Quantifizierung negativ beeinflussen können. Der Grund hierfür ist, dass die Entfaltung nicht zwischen Signaländerungen durch Kontrastmittel oder Bewegung unterscheiden kann. Hierzu wurde ein graphisches User-Interface entwickelt, mit dem AnwenderInnen durch die einzelnen Zeitschritte scrollen und anschließend festlegen können, wann der Start- und Endzeitpunkt der Messung sein sollte. Zeitpunkte werden dabei entfernt, wenn diese Atmung enthalten und die Zeitpunkte vor der Kontrastmittelankunft oder aber nach dem vollständigen Kontrastmitteldurchlauf liegen.

5.2.3 Lungensegmentierung

Die Segmentierung der Lunge wurde von der Arbeit von Kohlmann et al. [111] motiviert. Hierzu wurde die Lunge zunächst im morphologischen VIBE-Bild segmentiert, da darin aufgrund der höheren Auflösung und dem höheren CNR eine Unterscheidung zwischen Lunge und umliegendem Gewebe leichter ist. Im Anschluss musste die generierte Lungenmaske vom morphologischen VIBE-Bild auf das DCE-MRT Bild transformiert werden, wofür ein Bildregistrierungsalgorithmus verwendet wurde.

Segmentierung im morphologischen Bild

Die Segmentierung der Lunge wurde schwerpunktmäßig im koronalen VIBE-Bild durchgeführt, wobei zusätzlich das transversale VIBE-Bild für die Segmentierung der Trachea verwendet wurde. Als Segmentierungsmethode diente die in Abschnitt 4.5 beschriebene algorithmische Lungensegmentierung.

Registrierung des morphologischen MRT auf das DCE-MRT

Zur Registrierung wurde in der ursprünglichen Pipeline die deformierbare Demon-Registrierung von Kroon [153] verwendet, die in MATLAB implementiert ist. Eine deformierbare Registrierung ist notwendig, weil sich der Atemzustand zwischen dem morphologischen und dem DCE-MRT Bild unterscheiden kann und die Korrektur der Atemzustände nicht mit einer klassischen affinen Registrierung abbildbar ist. Zunächst wurde für das DCE-MRT Bild die Zeitachse entfernt, indem der Mittelwert entlang der Zeitachse für jeden Voxel berechnet wurde. Das DCE-MRT Bild diente anschließend als statisches Bild, während das morphologische Bild als bewegliches Bild definiert wurde. Bei der Demon-Registrierung werden jeweils Punkte entlang der Kanten der zwei zu registrierenden Objekte definiert, die als Membran dienen. Durch diese Membran können nun Voxel des

beweglichen Objektes in das statische Objekt gelangen. Die Voxel des beweglichen Objektes werden mithilfe von Kräften so verschoben, dass die Entropie zwischen den beiden Objekten minimiert wird. Eine Minimierung der Entropie wird erreicht, wenn die Grauwerte zweier übereinanderliegender Voxel ähnlich sind. Das Verschiebungsfeld eines jeden Schrittes wird berechnet mit

$$\mathbf{u} = \frac{(S_m - S_s)\,\Delta S_s}{|\Delta S_s|^2 + (S_m - S_s)^2}\ , \tag{5.15}$$

wobei $\mathbf{u}$ das Verschiebungsfeld ist, S_m und S_s jeweils der Intensität eines Voxels im beweglichen bzw. statischen Objekt entsprechen und ΔS_s den räumlichen Voxelabständen im statischen Bild entspricht. [154] Eine Erweiterung dieses Konzeptes, bei der zusätzlich auch die Voxel des ursprünglich statischen Objektes verschiebbar sind, führt zu deutlich kürzeren Registrierungszeiten [155]. Da sich der Bildkontrast zwischen dem morphologischen MRT und dem DCE-MRT aufgrund unterschiedlicher T1-Gewichtungen unterscheidet, wurde ein multimodaler Registrierungsansatz verwendet. Dabei wird für beide Bilder ein gemeinsames Histogramm generiert und jedem Voxel im beweglichen Bild der passende Grauwert aus dem statischen Bild zugewiesen. [156]

5.2.4 Detektion der Pulmonalarterie

Die für die Perfusionsquantifizierung benötigte AIF sollte möglichst nah am zu untersuchenden Gewebe gemessen werden. In der Lunge bietet sich hierfür daher die Pulmonalarterie an. In der verwendeten Perfusionsquantifizierungs-Pipepline wurde die AIF innerhalb der Pulmonalarterie am Punkt der Aufspaltung in die zwei Lungenarterien gemessen. Die Segmentierung der Pulmonalarterie geschah innerhalb des DCE-MRT Bildes. Zur Identifizierung dieses Punktes wurden zunächst alle Gefäße und das Herz detektiert. Dies geschah, indem eine Projektion der maximalen Werte (MIP, vom Englischen Maximum Intensity Projection) entlang der Zeitachse für einen Quader rund um die Herzregion berechnet wurde. Die Pulmonalgefäße und das Herz wurden dann als die Voxel definiert, die oberhalb des 95. Perzentils liegen. Eine Gruppierung von verbundenen Voxeln lieferte dann

zwei Objekte: Die linke Herzkammer und die rechte Herzkammer zusammen mit der Pulmonalarterie. Die linke Herzkammer wurde von der rechten Herzkammer und der Pulmonalarterie anschließend durch ein Clustering per Otsu Methode unterschieden. Das Signal innerhalb der linken Herzkammer unterscheidet sich von dem Signal der rechten Herzkammer bzw. der Pulmonalarterie deutlich, da die linke Herzkammer vom Kontrastmittel erst nach dem Passieren der Lunge durchflutet wird. Mithilfe von anatomischem Wissen über die Form der Pulmonalarterie lässt sich nun die Pulmonalarterie detektieren, indem nach einem runden Objekt mit einer Zirkularität von mindestens 60 % gesucht wird, das nicht mit dem Rand verbunden ist. Zusätzlich wurde das Wissen genutzt, dass sich das Objekt an der Position der Aufspaltung von einem runden Objekt in ein längliches, horizontales Objekt befindet.

5.2.5 Atmungskorrektur

Die halbautomatische Perfusionsquantifizierungs-Pipeline beinhaltet auch einen Algorithmus, der Bereiche aus der Lungenmaske entfernt, in denen eine flache Atmung, also eine kleine Atembewegung, stattfindet. Denn wird während der dynamischen DCE-MRT Messung geatmet, so führt dies zu künstlich, hohen Signalen im Subtraktionsbild $\mathbf{C}(t)$. Damit die Quantifizierung dadurch nicht verfälscht wird, wurden diese mithilfe eines Korrekturschrittes von der Auswertung ausgeschlossen. Für das automatische Auffinden der Voxel mit flacher Atmung wurde ausgenutzt, dass sich das Signal dieser Voxel deutlich vom Signal der AIF-Funktion unterscheidet. Mithilfe der normalisierten Kreuzkorrelation (NCC, vom Englischen Normalized Cross Correlation) kann die Ähnlichkeit zweier Kurven berechnet werden und damit auch die Ähnlichkeit des Signals eines Voxels mit der AIF. Die Formel der normalisierten Kreuzkorrelation zwischen einem Signal x und einem Signal y lautet:

$$NCC = \frac{1}{t_{\max}} \frac{\sum_{t=0}^{t_{\max}} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\sum_{i=0}^{t_{\max}} (x_i - \bar{x})^2\right)\left(\sum_{i=0}^{t_{\max}} (y_i - \bar{y})^2\right)}}\,, \tag{5.16}$$

wobei $\bar{x}$ und $\bar{y}$ jeweils den Mittelwerten der beiden Signale entsprechen. Diese normalisierte Kreuzkorrelation wurde sowohl für das Original-Signal, als auch für ein zeitlich verschobenes Signal in jedem Voxel berechnet. Für die Korrelation mit dem verschobenen Signal wurde zusätzlich auch der Zeitpunkt der maximalen Korrelation berechnet. Alle Voxel mit einer geringen Kreuzkorrelation oder einem verspäteten Korrelationsmaximum wurden als Voxel mit flacher Atmung identifiziert und von der Analyse ausgeschlossen. Zusätzlich wurden ebenso alle Voxel ausgeschlossen, deren PBV-Werte einen Wert von 200 ml/100 ml überschritten, da solche physikalisch unlogischen Werte ebenso von Atemartefakten verursacht werden.

5.2.6 Perfusionsquantifizierung

Die Perfusionsquantifizierung basierte auf den in Abschnitt 5.1 vorgestellten Formeln. Zunächst wurde aus dem DCE-MRT Bild ein Subtraktionsbild mithilfe von Gleichung (5.13) berechnet. Die ebenso benötigte AIF wurde berechnet, indem das mittlere Signal in der detektierten Pulmonalarterie aus Abschnitt 5.2.4 für jeden Zeitpunkt bestimmt wurde. Die gesuchte Residuenfunktion wurde basierend auf Gleichung (5.7) mithilfe von tSVD wie in Gleichung (5.12) gelöst. Für die Regularisierung wurden alle Werte der Diagonalmatrix der SVD, die kleiner als ein fester Schwellwert waren, gleich null gesetzt. Der Schwellwert wurde in Übereinstimmung mit den Veröffentlichungen von Østergaard et al. [132] und Sourbron et al. [157] auf 20 % des maximalen Singulärwertes $\sigma_{\max}$ festgelegt. Nach der erfolgreichen Berechnung der Residuenfunktion ergab sich der pulmonale Blutfluss PBF als das Maximum der Residuenfunktion und das Blutvolumen PBV als die Fläche unterhalb der Residuenfunktion. Die mittlere Transitzeit berechnete sich dann als Quotient von PBF und PBV. Zusätzlich zu den drei klassischen Perfusionsmetriken wurden eine Perfusionsdefektkarte und die Perfusionsdefekte in Prozent (QDP) berechnet. Basierend auf den Vergleichsergebnissen in [56], wurde die Perfusionsdefektkarte mithilfe der Otsu Methode berechnet, indem drei Cluster gesucht werden, die die Residuenkarte zum Zeitpunkt der größten Kontrastverstärkung in die Klassen Gefäße, normale Perfusion und Perfusionsdefekte

unterteilen. QDP berechnete sich als der Quotient der Anzahl an Voxeln, die als Perfusionsdefekte gekennzeichnet wurden sowie der Gesamtanzahl an Voxeln in der Lungenmaske.

5.3 Alternative Registrierungsmethoden

Die verwendete deformierbare Demon-Registrierung ist relativ langsam und macht einen Großteil der Zeitkosten der Gesamtauswertung aus. Außerdem hat sie insbesondere bei größeren Deformationen, wie z. B. bei großen Atemunterschieden Probleme, diese vollständig zu korrigieren. Daher wurden alternativ die zwei weitverbreiteten Registrierungsmethoden *elastix* [158, 159] und *ANTs* [160] getestet.

Bei der *elastix*-Registrierung bestand der erste Registrierungsschritt aus einer starren Registrierung, bei der das koronale VIBE-Bild als bewegliches Bild und das entlang der Zeitachse gemittelte DCE-MRT Bild als statisches Bild gewählt wurde. Zusätzlich wurde die Lungenmaske als bewegliche Maske für den Registrierungsschritt mitgegeben. Nach der starren Registrierung schloss sich ein affiner Registrierungsschritt an, bei dem die Lungenmaske jedoch nicht mehr verwendet wurde. Da die Atembewegung nicht mit einer reinen affinen Registrierung korrigiert werden kann, folgte ein deformierbarer BSpline-Registrierungsschritt. Für alle drei Registrierungsschritte wurde eine Registrierung in mehreren Auflösungen durchgeführt, bei der zunächst niedrig aufgelöste Varianten der Bilder registriert werden und die Auflösung nach und nach erhöht wird. Für die starre und affine Registrierung gab es vier Stufen mit je einem Achtel, einem Viertel, der halben und der kompletten Auflösung. Bei der BSpline-Registrierung wurde die Stufe mit einem Achtel der Auflösung nicht ausgeführt. Als Optimierungsmetrik wurde *AdvancedMattesMutualInformation* gewählt. Abschließend wurden alle drei bestimmten Transformationen kombiniert und zu einem Verschiebungsfeld zusammengefügt. Dieses Verschiebungsfeld wurde dann auf die Lungenmaske angewendet, wobei hier die Interpolationsmethode der nächsten Nachbarn gewählt wurde. Die verwendeten *elastix*-Registrierungsparameter sind im Detail in Anhang A.1 aufgelistet.

Für die *ANTs*-Registrierung wurden nur zwei Registrierungsschritte, eine affine Registrierung und eine Kombination aus einer affinen und deformierbaren Registrierung („SyN“) durchgeführt. Auch hier wurden das koronale VIBE-Bild als bewegliches Bild und das entlang der Zeitachse gemittelte DCE-MRT Bild als statisches Bild gewählt. Die gewählten Parameter für die *ANTs*-Registrierung sind in Anhang A.2 aufgelistet. Zusätzlich wurde für beide Registrierungsmethoden jeweils ein zusätzlicher Nachbearbeitungsschritt durch die Verwendung eines Chan-Vese aktiven Konturenmodells [161] getestet. Hierzu wurde das aktive Konturenmodell jeweils einzeln auf jede Schicht des zeitlich gemittelten DCE-MRT Bildes mit einem Glättungsfaktor von 4,0 angewendet.

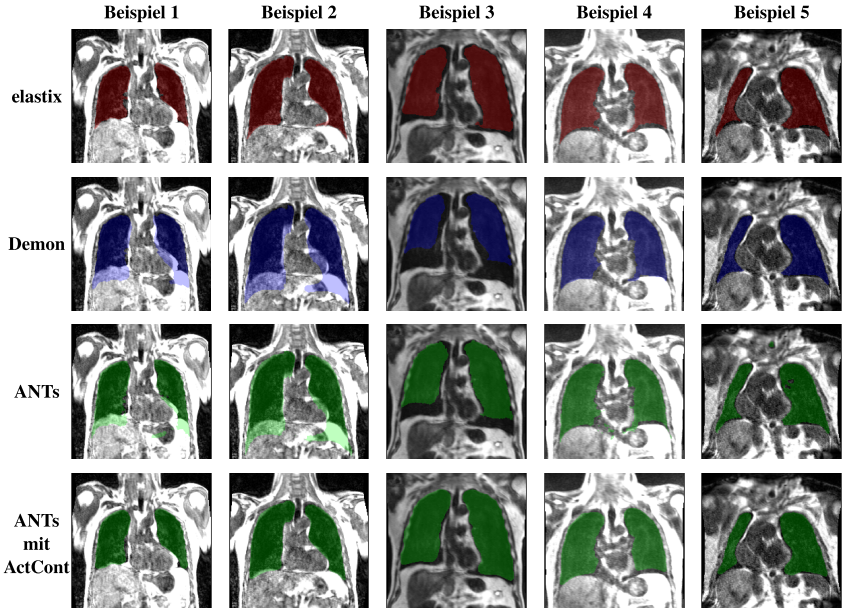


Abbildung 5.4: Vergleich zwischen *elastix*-, Demon- und *ANTs*-Registrierung für fünf verschiedene Beispielbilder. Erste Reihe: *elastix*-Registrierung, zweite Reihe: Demon-Registrierung, dritte Reihe: *ANTs*-Registrierung und vierte Reihe: *ANTs*-Registrierung mit dem aktiven Konturenmodell. Die Farben dienen nur der besseren Unterscheidbarkeit der einzelnen Methoden.

In Abbildung 5.4 sind die registrierten Lungenmasken für die *elastix*-, Demon- und *ANTs*-Registrierung für fünf verschiedene Beispielbilder dargestellt. Für die zwei Bilder in der ersten Spalte führte die *elastix*-Registrierung zu einem nahezu perfekten Registrierungsergebnis, während die Demon- und *ANTs*-Registrierung den Atemzustand nur unzureichend registrierten und die Lungenmaske nach unten deutlich zu groß war. Beim Bild in der mittleren Spalte führte keine der Registrierungsvarianten zu optimalen Ergebnissen, da die Lungenmaske unten nicht passte, wobei auch hier die Demon- und *ANTs*-Registrierungen die schlechtesten Ergebnisse erzeugten. Für die Bilder in der vierten und fünften Spalte führte die Demon-Registrierung zu den besten Ergebnissen. In der vierten Reihe sieht man die *ANTs*-Registrierung mit Nachbearbeitung mittels aktivem Konturenmodell. Man erkennt, dass die Nachbearbeitung die Lungenmaske der *ANTs*-Registrierung für alle Bilder deutlich verbesserte. Auch allgemein betrachtet erzeugte diese Registrierungsvariante im Mittel die konstantesten und besten Ergebnisse.

5.4 Automatische Atmungserkennung

Das DCE-MRT sollte im Optimalfall innerhalb eines inspiratorischen Atemanhalts gemessen werden, um den Kontrastmitteldurchlauf korrekt abzubilden. Manche der PatientInnen können ihren Atem jedoch nicht für die komplette Dauer der Messung (ca. 35 s) anhalten. In Abbildung 5.5 ist beispielhaft die Atmung zwischen zwei benachbarten Zeitpunkten einer DCE-MRT Messung dargestellt. Da die Quantifizierung auf Voxellevel geschieht, können schon kleine Verschiebungen des Lungenparenchyms zu fehlerhaften quantitativen Perfusionswerten führen. Außerdem können durch Bewegungen während der DCE-MRT Messung mittels der TWIST-Sequenz Artefakte im gemessenen Bild entstehen. Aus diesem Grund ist es essenziell, eine mögliche Atembewegung zwischen zwei Zeitpunkten bzw. generell innerhalb einer DCE-MRT Messung zu detektieren. In der ursprünglichen Perfusionsquantifizierungs-Pipeline wurden solche Atemzeitpunkte manuell bestimmt (siehe Abschnitt 5.2.2) und dann von der Quantifizierung ausgeschlossen, sofern diese vor oder nach dem Kontrastmitteldurchlauf lagen. Dieser Schritt wird

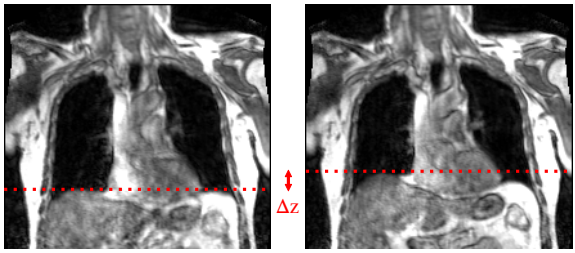


Abbildung 5.5: Atmung zwischen zwei Zeitpunkten einer DCE-MRT Messung.

dabei einerseits durch die Subjektivität der Anwendenden beeinflusst, andererseits nimmt die manuelle Bearbeitung erneut Zeit in Anspruch. Um den Prozess der Atmungserkennung zu automatisieren, wurde eine CNN zur Atmungserkennung entwickelt und trainiert.

5.4.1 Daten-Vorverarbeitung

Für das Trainieren, Validieren und Testen wurden DCE-MRT Messungen von 75 zufälligen PatientInnen gewählt, die jeweils zwischen 0–1 normiert wurden. Von jedem Zeitpunkt einer DCE-MRT Messung wurde jeweils die mittlere Schicht extrahiert. Für jedes Paar von zwei angrenzenden Zeitpunkten wurde ein manuelles Label definiert, ob Atmung zwischen den Zeitpunkten stattgefunden hat oder nicht. Insgesamt konnten so 1615 Paare von je zwei aneinandergrenzenden Zeitpunkten generiert werden. Diese wurden dann mithilfe einer patientenweisen Teilung in einen Trainings-, einen Validierungs- und einen Testdatensatz mit einer Aufteilung von $(65\%/15\%/20\%)$ aufgeteilt. Die patientenweise Aufteilung garantiert hier, dass Bilder der gleichen PatientInnen nicht in verschiedenen Datensätzen vorhanden sind.

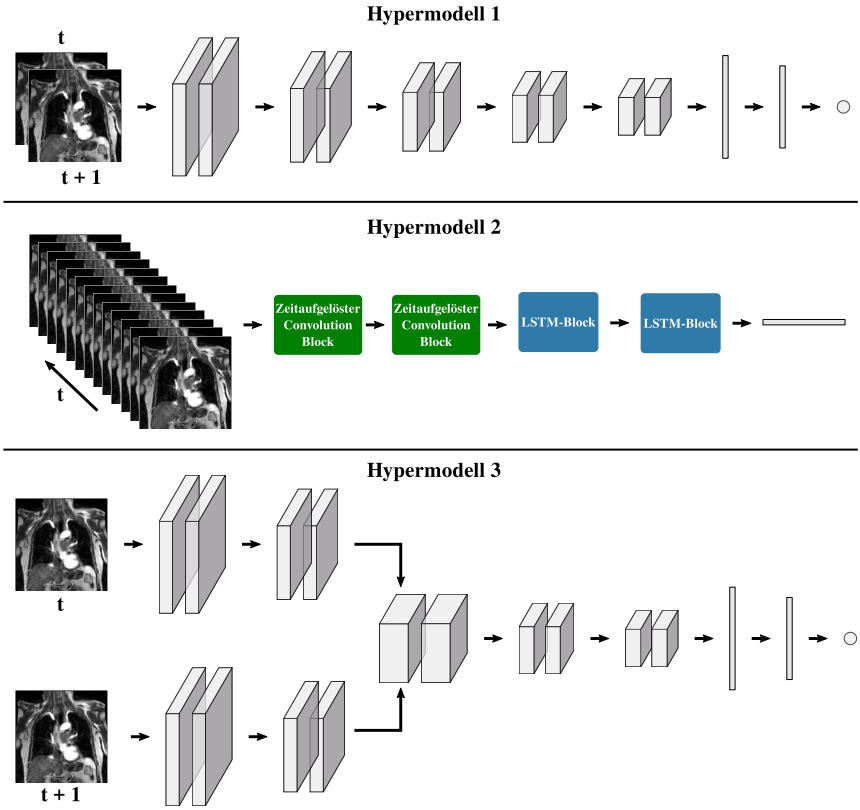


Abbildung 5.6: Überblick über die drei Hypermodelle, die für die automatische Atemdetektion getestet wurden.

5.4.2 Auswahl einer passenden Modellarchitektur

In einem ersten Schritt wurden drei verschiedene Modellarchitekturen (Hypermodelle) untersucht, die schematisch in Abbildung 5.6 dargestellt sind. Die erste Architektur ist ein klassisches CNN, das ein Bildpaar zweier Zeitpunkte, die entlang der Channel-Achse gestapelt werden, als Input erhält. Die Inputdaten werden mit mehreren Convolution-Blöcken verarbeitet und abschließend wird eine

Entscheidung zurückgegeben, ob Atmung stattgefunden hat oder nicht. Die zweite Architektur erhält die komplette Zeitreihe der mittleren Schicht des DCE-MRT als Input. Diese Zeitreihe wird zunächst mit zeitaufgelösten Convolution-Blöcken und anschließend mit sogenannten LSTM-Blöcken (Long Short-Term Memory) verarbeitet. Ein LSTM ist ein rekurrentes neuronales Netz, das speziell für die Analyse von Sequenzdaten ausgelegt ist, da es auch die Signale von vorherigen Zeitpunkten in die Verarbeitung des aktuellen Zeitpunktes mit einbezieht. Abschließend werden für die komplette Zeitreihe die jeweiligen Atemzeitpunkte vorhersagt. Die dritte Architektur erhält die zwei benachbarten Zeitpunkte separat. Diese werden dann zunächst jeweils in ihrem eigenen Zweig mit Convolution-Blöcken verarbeitet. Anschließend werden die Zweige zusammengeführt und der gemeinsame Zweig dann mit weiteren Convolution-Blöcken verarbeitet, bis wieder eine Vorhersage über das Stattfinden von Atmung zurückgegeben wird.

Alle drei Hypermodelle wurden in Python mit TensorFlow 2.11 implementiert und es wurde jeweils eine Hyperparametersuche mit dem Hyperband-Algorithmus von *keras-tuner* [162] für alle drei Modelle durchgeführt. Die untersuchten Hyperparameter und deren Bereiche sind in Anhang A.3 aufgelistet. Die jeweils besten Hyperparameterkombinationen aus der Hyperparametersuche wurden anschließend ausgewählt und diese Modelle für eine größere Anzahl an Epochen trainiert. Alle Modelle wurden mithilfe von Online Datenaugmentierung trainiert, um die Varianz und Größe des Datensatzes weiter zu erhöhen. Für die Augmentierung wurden zufällige Kombinationen von Bildskalierung mit Skalierungsfaktoren im Bereich [0.8, 1.0], Rotationen im Bereich [0°, 8°] und Kontrastveränderungen im Bereich [0.6, 0.9] genutzt. Das Training wurde auf einer Nvidia A100 mit 80 GB GPU-RAM und 128 AMD EPYC 7513 32-Core Prozessoren mit einem Gesamt RAM von 2 100 GB durchgeführt.

Nach dem Training wurden die Modelle auf den ungesehenen Testdatensatz angewendet und hierfür jeweils die Genauigkeit, also die prozentuale Anzahl der richtig erkannten Atemvorgänge, berechnet. Die Ergebnisse sind in Tabelle 5.1 dargestellt. Man erkennt, dass die beiden regulären CNN Modelle (Hypermodell 1 und Hypermodell 3) sehr ähnlich abgeschnitten haben, wobei das Modell mit nur

Tabelle 5.1: Vergleich der Performance der einzelnen Modellarchitekturen zur automatischen Atmungserkennung auf dem ungesehenen Testdatensatz.

Modell	Genauigkeit [%]
Hypermodell 1	93,49
Hypermodell 2	70,36
Hypermodell 3	93,16

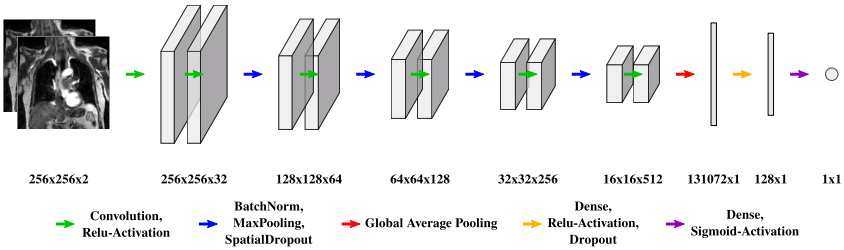


Abbildung 5.7: Aufbau des CNN zur automatischen Atemdetektion

einem Zweig (Hypermodell 1) minimal besser war. Da dieses Modell jedoch deutlich weniger Parameter als Hypermodell 3 besitzt und daher schneller zu trainieren ist, wurde im Nachfolgenden nur noch diese Modellarchitektur verwendet.

5.4.3 Ausgewählte Modellarchitektur

Die exakte Modellarchitektur des besten Modells ist in Abbildung 5.7 dargestellt. Das Modell erhielt als Input jeweils die mittlere Schicht von zwei aufeinanderfolgenden Zeitpunkten des DCE-MRT, die entlang der Channel-Richtung gestapelt sind. Das Eingangsbild wurde anschließend mit fünf Convolution-Blöcken verarbeitet, wobei jeder Convolution-Block aus je zwei 2D-Convolution Schichten und einer ReLU-Aktivierungsschicht besteht. Zwischen den Convolution-Blöcken wurde je eine Batch-Normalisierungsschicht, eine Maximum-Pooling Schicht und eine räumliche Dropout-Schicht mit einer Dropout-Rate von 0,2 platziert.

Nach dem letzten Convolution-Block wurde der 2D-Ausgang mithilfe einer globalen Average-Pooling Schicht geebnet. Der geebnete Output wurde dann mit einer dichten Schicht (engl. Dense-Layer) mit einer ReLU-Aktivierung und einer Dropout-Schicht mit einer Dropout-Rate von 0,2 weiterverarbeitet. Die finale Schicht des Modells war eine dichte Schicht mit einem einzelnen Neuron und Sigmoid-Aktivierung.

5.4.4 Validierung des Modells

Zur Validierung des trainierten Modells wurde ein künstlicher Testdatensatz erzeugt. Hierbei sollte sowohl untersucht werden, wie das Modell auf unterschiedlich starke Verschiebungen reagiert, als auch welchen Einfluss ein Kontrastunterschied im Bild hat. Hierzu wurde von insgesamt 527 PatientInnen jeweils ein zentrales Bild zu Beginn der Messung extrahiert, in dem noch kein Kontrastmittel anwesend war. Anschließend wurden jeweils Abwandlungen dieses Bildes generiert, um Atembewegung zu simulieren. Hierzu wurde ein Verschiebungsfeld auf das Bild angewendet, bei dem die obere Hälfte gleich null ist und die untere Hälfte jeweils einem Wert Δz entspricht. Insgesamt wurden 13 Varianten mit Verschiebungen im Bereich $[-3.0, 3.0]$ in Schritten von 0,5 Voxeln erzeugt. Das Gleiche wurde für ein zentrales Bild zu einem späteren Messzeitpunkt wiederholt, an dem Kontrastmittel im Lungenparenchym sichtbar ist. Dieser Testdatensatz simuliert sowohl die Ein- als auch Ausatmung für Zeitpunkte mit und ohne Kontrastmittel. Er erlaubt damit eine Überprüfung der Sensitivität des trainierten Atemdetektions-Modells und gibt Einblicke, wie robust das Modell auf mögliche Kontrastunterschiede im Bild reagiert.

Die Ergebnisse sind in Tabelle 5.2 dargestellt. Man erkennt, dass das Modell bereits bei kleinsten Verschiebungen von einem Voxel eine Atmung mit einer Sicherheit von fast 96 % für eine künstliche Ausatembewegung und mit ca. 92 % für eine künstliche Einatembewegung detektiert. Dies schien unabhängig von der Kontrastmittelanwesenheit zu sein. Der Grund für die bessere Vorhersage

Tabelle 5.2: Validierung des trainierten CNN zur Atmungserkennung auf einem künstlich erzeugten Testdatensatz. Mittelwerte und Standardabweichungen der vorhergesagten Atemwahrscheinlichkeit für den kompletten künstlichen Testdatensatz für verschiedene Verschiebungen. Nativ: Kein Kontrastmittel im Bild, Mit KM: Mit Kontrastmittel im Bild.

Verschiebung	**Mean**		**Std**	
	Nativ	**Mit KM**	**Nativ**	**Mit KM**
-3.0	0,9955	0,9938	0,0573	0,0624
-2.5	0,9952	0,9942	0,0590	0,0623
-2.0	0,9949	0,9942	0,0608	0,0620
-1.5	0,9886	0,9916	0,0863	0,0696
-1.0	0,9591	0,9451	0,1627	0,1820
-0.5	0,0535	0,0190	0,1579	0,0875
0.0	0,0062	0,0031	0,0293	0,0375
0.5	0,0505	0,0192	0,1579	0,0858
1.0	0,9230	0,8909	0,2266	0,2583
1.5	0,9856	0,9894	0,0939	0,0809
2.0	0,9952	0,9930	0,0576	0,0697
2.5	0,9951	0,9926	0,0584	0,0708
3.0	0,9954	0,9920	0,0557	0,0705

einer Ausatembewegung lag vermutlich an einer deutlich größeren Menge an Ausatembewegung im Trainingsdatensatz, da die Messungen ja im inspiratorischen Atemanhalt stattfanden und daher eine Ausatembewegung wahrscheinlicher ist. Ebenso lässt sich anhand der mittleren Standardabweichungen erkennen, dass eine Verschiebung von einem Voxel vermutlich der Übergangsbereich im Modell zwischen Atmung und keiner Atmung ist, da hier die Standardabweichungen deutlich größer ausfielen. Die Ergebnisse lassen darauf schließen, dass das Modell robust gegenüber Kontrastunterschieden und sehr sensitiv war und daher gut geeignet ist, um Atmung in DCE-MRT Messungen automatisch zu detektieren.

5.5 Charakterisierung des Kontrastmitteldurchlaufs

Ein weiterer nützlicher Aspekt bei der Auswertung von DCE-MRT Messungen ist die Charakterisierung des Kontrastmittelbolus. Hierzu zählt sowohl die automatische Bestimmung der Kontrastmittelankunft t_{BA} in der Lunge, als auch des Zeitpunkts des vollständigen Kontrastmitteldurchlaufs t_{BC}. Die Kontrastmittelankunft wurde bestimmt, indem die Position des Maximums der diskreten Ableitung des AIF-Signals bestimmt wurde, also der Zeitpunkt der maximalen positiven Steigung der AIF. Für die Berechnung des Zeitpunkts der kompletten Boluspassage wurden dagegen alle zeitlichen Signalverläufe in der Lunge betrachtet. Für jeden Voxel wurde hierbei der Zeitpunkt des zweiten Maximums des Signals bestimmt und dann in einem Array gespeichert. Nachdem dieser Array mit den bestimmten Zeitpunkten aller Lungenvoxel gefüllt wurde, wurde das 90. Perzentil dieses Arrays bestimmt und als Zeitpunkt der kompletten Boluspassage festgelegt.

Eine genaue Betrachtung der gemessenen AIF liefert Informationen, wie das Kontrastmittel injiziert worden ist. Eine fehlerhafte Kontrastmittelinjektion kann zu einer schlechten DCE-MRT Messung führen. Wurde das Kontrastmittel zu früh injiziert, so kommt dieses möglicherweise bereits im ersten gemessenen Zeitpunkt in der Lunge an, wodurch ein kontrastmittelfreier Referenzzeitpunkt fehlt, der für die Berechnung des Subtraktionsbildes benötigt wird. Auch das Gegenteil, eine zu späte Injektion des Kontrastmittels, kann für die Perfusionsquantifizierung problematisch sein, da dies dazu führen kann, dass das Kontrastmittel die Lunge zum Ende der Messung noch nicht komplett durchlaufen hat und somit Teile der Perfusion erst nach dem Abschluss der Messung stattfinden. Neben dem Kontrastmittel selbst wird noch ein sogenannter Chaser injiziert, der das Kontrastmittel vor sich herschiebt. Dies führt zu einem wohl definierten Kontrastmittelbolus. Wurde die Chaser-Injektion jedoch vergessen oder wurde zu wenig Chaser injiziert, so kommt es zu keiner wohldefinierten Bolusform.

In Abbildung 5.8 sind vier verschiedene AIF Signalverläufe dargestellt. Abbildung 5.8 (a) zeigt dabei eine optimale Bolusinjektion. Die AIF besitzt einen

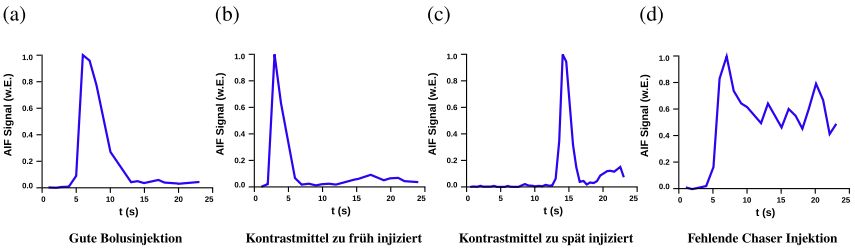


Abbildung 5.8: Darstellung der AIF-Kurve für verschiedene Fehler bei der Kontrastmittelinjektion: (a) Gute Bolusinjektion (b) Zu frühe Injektion des Kontrastmittels (c) Zu späte Injektion des Kontrastmittels (d) Falsche Injektion eines Chasers.

steilen Anstieg, einen schmalen wohldefinierten Höhepunkt und anschließend einen schnellen Abfall. In Abbildung 5.8 (b) ist zwar erneut ein wohldefinierter Bolus zu erkennen, jedoch ist die Position des Bolus viel zu früh und es kann kein gutes Referenzbild ohne Kontrastmittel bestimmt werden. Den Fall einer zu späten Kontrastmittelinjektion zeigt Abbildung 5.8 (c). Eine solche Injektion kann dazu führen, dass möglicherweise nicht die komplette Boluspassage gemessen werden kann und es zu Berechnung von fehlerhaften quantitativen Perfusionswerten kommt. Eine Kontrastmittelgabe mit falscher Chaser-Injektion ist in Abbildung 5.8 (d) dargestellt. Man erkennt keine wohldefinierte Form des Bolus in der AIF-Kurve, was zu Problemen bei der Entfaltung führen kann.

5.6 Automatisches Herausfiltern veratmeter Zeitpunkte

Das Modell zur automatischen Identifizierung von Atembewegung und die zuvor beschriebene Bestimmung der Bolusankunft und des vollständigen Bolusdurchlaufs bieten nun die Möglichkeit, den in Abschnitt 5.2.2 beschriebenen manuellen Schritt der Atmungsentfernung durch eine automatisierte Variante zu ersetzen. Das hierzu entwickelte Vorgehen ist in Abbildung 5.9 visualisiert.

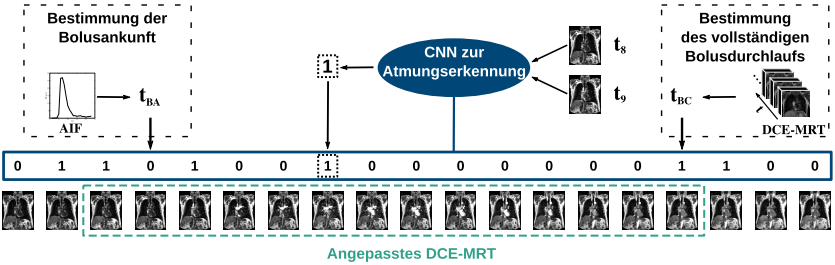


Abbildung 5.9: Darstellung des automatischen Vorgehens zur Entfernung von veratmeten Zeitpunkten aus der DCE-MRT Messung.

Zunächst wurde der Zeitpunkt der Bolusankunft und der Zeitpunkt des vollständigen Bolusdurchlaufs bestimmt. Anschließend wurde für jedes Paar von aufeinanderfolgenden Zeitpunkten bestimmt, ob zwischen den beiden Zeitpunkten Atmung stattgefunden hat oder nicht. Hierzu wurde das entwickelte CNN zur Atmungserkennung genutzt. Mithilfe der nun gewonnenen Informationen ließ sich die DCE-MRT Messung automatisch anpassen, falls Atmung während der Messung stattgefunden hat. Hierfür wurden alle Zeitpunkte nach dem ersten Atemzeitpunkt nach dem vollständigen Kontrastmitteldurchlauf t_{BC} entfernt. Da für die Berechnung des Subtraktionsbildes mindestens ein Zeitpunkt vor der Kontrastmittelankunft vorhanden sein muss, wurde die Messung am letzten Atmungszeitpunkt vor der Kontrastmittelankunft abgeschnitten.

In Abbildung 5.10 ist für einen beispielhaften Patienten dargestellt, wie sich die quantitativen Perfusionskarten nach der erfolgten Anpassung verändert haben. Ohne die Entfernung von Atmung waren große Atemartefakte in Zwerchfellnähe in der PBF-Karte in Abbildung 5.10 (a) erkennbar. Diese Atemartefakte verschwanden für die Quantifizierung der angepassten DCE-MRT Messung nahezu komplett (weiße Pfeile). Die Perfusionsdefektkarten (QDP-Karte) in Abbildung 5.10 (b) zeigten, dass das Herausschneiden der veratmeten Zeitpunkte auch hier besonders wichtig war. In der unteren linken Lunge wurden in der Quantifizierung der Originalmessung links sehr viele Bereiche als Perfusionsdefekte markiert (weiße Pfeilspitzen). Betrachtet man jedoch die korrigierte PBF-Karte, so erkennt man, dass diese Bereiche eigentlich gut durchblutet waren. Diese Bereiche wurden in

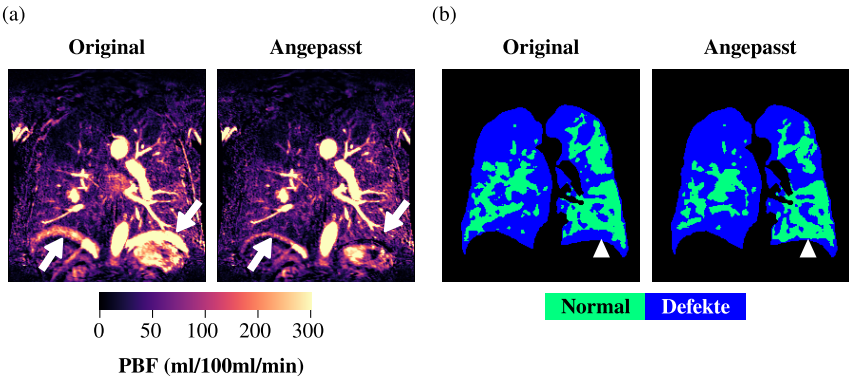


Abbildung 5.10: Vergleich der quantitativen Perfusionswerte für die Quantifizierung mit und ohne Entfernung von veratmeten Zeitpunkten. (a) Vergleich der PBF-Karte (b) Vergleich der berechneten Perfusionsdefekte.

der korrigierten QDP-Karte auch richtigerweise als normale Bereiche markiert. Im Gegensatz dazu waren im unteren rechten Lungenlappen zwar auch Atemartefakte sichtbar, aber nach der Korrektur zeigt sich, dass diese Bereiche sehr schlecht durchblutet waren und daher ein Perfusionsdefekt vorlag. Auch dieser wurde in der Quantifizierung der angepassten Messung richtig als defekter Bereich erkannt.

5.7 Quantifizierung der Perfusionsverzögerung

Die Bestimmung der Stärke der Perfusion oder die Detektion von Perfusionsdefekten sind zwei wichtige Aspekte der Perfusionsquantifizierung. Manchmal kann es jedoch auch interessant sein, Bereiche in der Lunge zu erkennen, die eine verzögerte Perfusion besitzen. Eine verzögerte Perfusion kann nämlich beispielsweise auf eine Obstruktion der Atemwege hinweisen. Aus diesem Grund wurde eine Methode zur Quantifizierung der Perfusionsverzögerung entwickelt. Das Subtraktionsbild $\mathbf{C}(t)$ wurde zunächst mit einem Tiefpassfilter gefiltert, wodurch das tiefpassgefilterte Subtraktionsbild $\mathbf{C_f}(t)$ entsteht. Anschließend wurde sowohl $\mathbf{C_f}(t)$ als auch $\mathbf{C_a}(t)$ mithilfe von kubischer Interpolation zeitlich auf eine

Auflösung von 100 Zeitschritten hochskaliert. Die daraus entstandenen hochaufgelösten Matrizen werden als $\mathbf{C_{f,i}}(t)$ und $\mathbf{C_{a,i}}(t)$ bezeichnet. Die Kreuzkorrelation k zwischen jedem Voxel von $\mathbf{C_{f,i}}(t)$ mit einem beweglichen Fenster $\omega_{\mathbf{C_{a,i}}}$ der mit am Rand mit Nullen aufgefüllten und interpolierten arteriellen Inputfunktion $\mathbf{C_{a,i}}(t)$ lässt sich berechnen zu:

$$k(t,x,y,z) = \frac{1}{n}\sum_{x,y,z}\frac{1}{\sigma_{C_{f,i}}\sigma_{\omega_{C_{a,i}}}}\left(\mathbf{C_{f,i}}(x,y,z)-\mu_{C_{f,i}}\right)\left(\omega_{\mathbf{C_{a,i}}}-\mu_{\omega_{C_{a,i}}}\right)\ , \qquad (5.17)$$

wobei $\mu_{C_{f,i}}$ und $\mu_{\omega_{C_{a,i}}}$ jeweils die Mittelwerte sowie $\sigma_{C_{f,i}}$ und $\sigma_{\omega_{C_{a,i}}}$ jeweils die Standardabweichungen von $\mathbf{C_{f,i}}$ und $\omega_{\mathbf{C_{a,i}}}$ sind. Die Perfusionsverzögerung t_{delay} ergibt sich dann als die Position der maximalen Kreuzkorrelation:

$$t_{delay}(x,y,z) = \arg\max_{t}(\mathbf{k}(x,y,z,t))\ , \qquad (5.18)$$

wobei $\mathbf{k}(t)$ dem Vektor aller Kreuzkorrelationen der verschiedenen Verschiebungen entspricht. Das Vorgehen zur Berechnung der Perfusionsverzögerung ist auch in Abbildung 5.11 visualisiert.

5.8 DCE-MRT Qualitätsscore

Die entwickelte Pipeline zur Perfusionsquantifizierung wird fast immer quantitative Perfusionswerte berechnen und ausgeben. Dabei spielt es keine Rolle, ob die zugrundeliegende DCE-MRT Messung eine gute oder schlechte Qualität besitzt. Die Qualität der DCE-MRT Messung hat jedoch einen Einfluss auf die quantitativen Werte. Ein gutes Beispiel hierfür sind die in Abschnitt 5.5 beschriebenen fehlerhaften Kontrastmittelinjektionen, die zu fehlerhaften quantitativen Perfusionswerten führen können. Ein weiteres Beispiel ist Atmung während der DCE-MRT Messung. Da es sich bei der Perfusionsquantifizierung um eine voxelbasierte Methode handelt, führen schon leichte Verschiebungen der Voxel zwischen den einzelnen Zeitpunkten der Messung zu fehlerhaften quantitativen Perfusionswerten. Außerdem kann die Atembewegung auch zu Artefakten in der

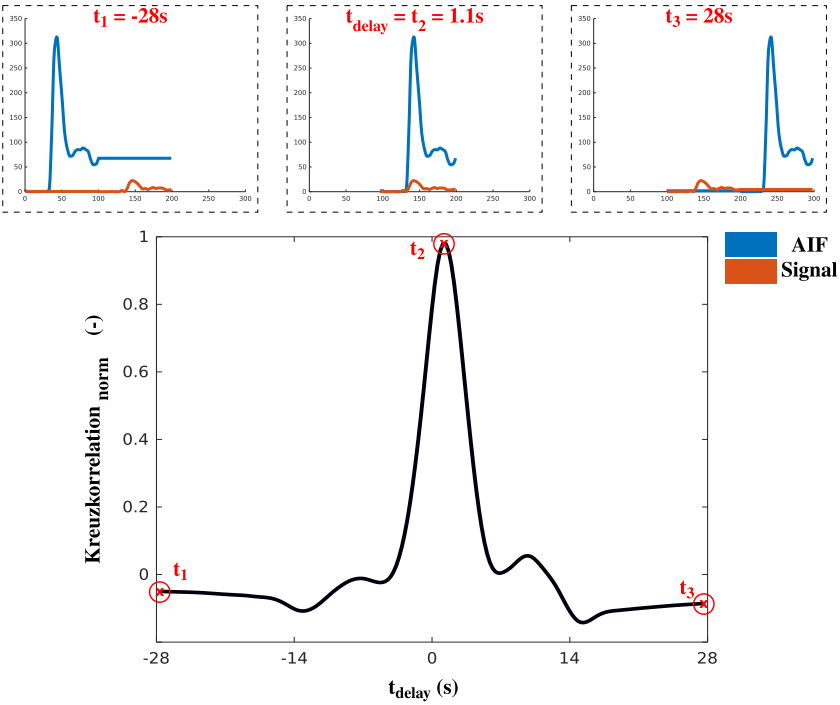


Abbildung 5.11: Darstellung des Vorgehens zur Berechnung der Perfusionsverzögerung.

TWIST-Messung führen. Um auf diese möglichen Probleme hinzuweisen, kam die Idee, einen DCE-MRT Qualitätsscore zu entwickeln, der zusätzlich zu den quantitativen Perfusionswerten ausgegeben wird. Der Qualitätsscore besteht aus zwei Bestandteilen, die die DCE-MRT Qualität in Bezug auf die Atmung und auf die Kontrastmittelinjektion bewerten. Einen Überblick über den DCE-MRT Qualitätsscore gibt Abbildung 5.12.

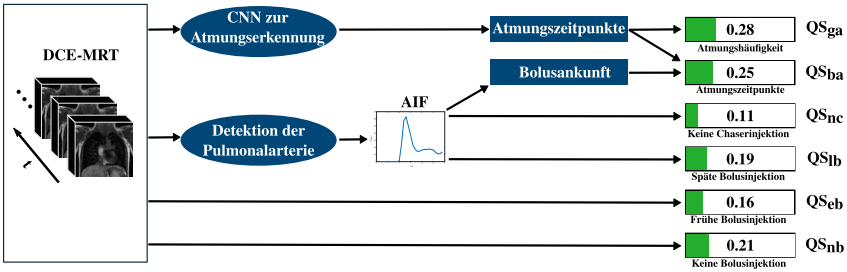


Abbildung 5.12: Überblick über den DCE-MRT Qualitätsscore.

5.8.1 Messung der Qualität der Bolusinjektion

Der Qualitätsscore für die Bolusinjektion besteht aus vier Teilscores, wobei für jeden Teilscore zunächst eine Wahrscheinlichkeit p_* für das jeweilige Event berechnet wurde und sich der Teilscore dann mit $QS_* = 1 - p_*$ berechnete. Der erste Teilscore basierte auf der Wahrscheinlichkeit einer unzureichenden oder fehlenden Chaser-Injektion p_{nc}. Die Wahrscheinlichkeit wurde basierend auf dem Maximum der AIF $a_{\max}$ sowie der Standardabweichung des AIF-Signals zwischen Maximum und Ende der AIF $\sigma_{a,me}$ berechnet:

$$p_{nc} = \frac{0.25\, a_{\max} - \sigma_{a,me}}{0.15\, a_{\max}}\, \mathrm{e}^{-0.001\left(\sigma_{a,me} - 0.1\, a_{\max}\right)} \tag{5.19}$$

Der zweite Teilscore basierte auf der Wahrscheinlichkeit p_{nb}, die misst, ob der Kontrastmittelbolus nicht injiziert wurde. Der Teilscore wurde mithilfe der Standardabweichung $\sigma_{S_{h,\max}}$ des maximalen Signals für jeden Zeitpunkt innerhalb einer Region, die das Herz umfasst, berechnet:

$$p_{nb} = \begin{cases} \frac{30 - \sigma_{S_{h,\max}}}{28}\, \mathrm{e}^{-0.001\left(\sigma_{S_{h,\max}} - 2\right)}, & 2 < \sigma_{S_{h,\max}} < 30 \\ 0, & \sigma_{S_{h,\max}} \geq 30 \\ 1, & \sigma_{S_{h,\max}} \leq 2 \end{cases} \tag{5.20}$$

Der dritte Teilscore wurde basierend auf der Wahrscheinlichkeit p_{eb}, die die Wahrscheinlichkeit einer fehlenden Bolusinjektion misst, bestimmt. Zur Berechnung

wurde das maximale Signal für jeden Zeitpunkt innerhalb einer Region, die das Herz umfasst $S_{h,\max}$ verwendet:

$$p_{eb} = \mathrm{e}^{-0.2\left(\arg\max\ \{S_{h,\max}\}-1\right)} \tag{5.21}$$

Der letzte Teilscore misst die Wahrscheinlichkeit einer zu späten Bolusinjektion p_{lb} und basierte auf dem AIF Signalverlauf:

$$p_{lb} = \begin{cases} \left(\frac{10-\left(l_{C_a}-\arg\max\{C_a\}\right)}{5}\right)^2, & 5 < (l_{C_a} - \arg\max\{C_a\}) < 10 \\ 0\,, & (l_{C_a} - \arg\max\{C_a\}) \geq 10 \\ 1\,, & (l_{C_a} - \arg\max\{C_a\}) \leq 5 \end{cases} \tag{5.22}$$

Hierbei entspricht l_{C_a} der Anzahl der Zeitpunkte der AIF und $\arg\max\{C_a\}$ ist der Zeitpunkt des Maximums der AIF.

5.8.2 Bestimmung der Atmungshäufigkeit

In Bezug auf die Atmung besitzt der DCE-MRT Qualitätsscore zwei verschiedene Komponenten. Für den ersten Teilscore wurde die relative Atemhäufigkeit während der kompletten Messung h_{ga} bestimmt. Hierzu wurde die Anzahl der Zeitpunkte mit Atmung durch die Anzahl der Gesamtzeitpunkte dividiert. Der zweite Teilscore h_{ba} misst dagegen, wie viel Atmung während der Boluspassage vorhanden ist. Dies ist besonders wichtig, da aktuell keine Möglichkeit besteht, diese Atmung zu korrigieren und Atmung während der Boluspassage daher zwangsläufig zu verfälschten Ergebnissen führt. Hierzu wurde die Anzahl der veratmeten Zeitpunkte während der Boluspassage mit der Länge der Boluspassage dividiert. Beide berechneten Werte stellen daher eine relative Häufigkeit dar. Um daraus einen Qualitätsscore zu generieren, wurde die relative Häufigkeit vom Wert eins abgezogen. Dadurch ergab sich für eine DCE-MRT Messung ohne Atmung der Teilscore der Gesamtatmung $QS_{ga} = 1 - h_{ga} = 1$ und für Atmung während der

gesamten Messung $QS_{ga} = 0$. In gleicher Weise berechnete sich der Teilscore für die Atmung während der Boluspassage mit $QS_{ba} = 1 - h_{ba}$.

5.9 EigenLung-Distanz Metrik

Die in Abschnitt 3.2.1 vorgestellten Segmentierungsmetriken lassen sich alle den beiden Kategorien der volumenbasierten und der randbasierten Metriken zuordnen. Neben dem Volumen und dem Rand ist jedoch auch oft die Form einer Segmentierung charakteristisch für ein Objekt. Bei der Segmentierung der Lunge wird erwartet, dass die Segmentierung einer Lunge gleicht, also aus zwei Lungenhälften besteht, deren oberer Rand eine konvexe Form und deren unterer Rand eine konkave Form besitzen. Um dies zu nutzen, wurde im Zuge dieser Arbeit eine neue formbasierte Metrik entwickelt, die von den *EigenFaces*-Papern [163, 164] motiviert wurde. Die wesentliche Idee dieser Veröffentlichungen war, eine Hauptkomponentenanalyse (PCA) auf Bilder von menschlichen Gesichtern anzuwenden. Das Resultat der PCA sind Eigenwerte und Eigenvektoren, die genutzt werden können, um jedes Gesicht aus dem Datensatz zu rekonstruieren und sogar neue, unbekannte Gesichter damit zu generieren. [163, 164]

Angelehnt an diese beiden Arbeiten wurde in dieser Arbeit das EigenLung-Modell entwickelt. Die einzelnen Schritte zur Berechnung des EigenLung-Modells sind in Abbildung 5.13 (a) visualisiert. Zunächst wurden die Lungenmasken von 527 PatientInnen mittels einer affinen Registrierung mithilfe von *ANTs* [160] auf eine Referenzlunge registriert. Durch Mittelung aller registrierten Lungenmasken ergab sich dann eine durchschnittliche Lunge (MeanLung). Die registrierten Masken wurden anschließend mithilfe einer SVD in Eigenvektoren und Eigenwerte zerlegt. Das EigenLung-Modell ergibt sich dann aus der Kombination aus MeanLung und den Eigenvektoren. Mit dem damit erstellten EigenLung-Modell kann nun jede neue Lungenmaske mithilfe ihrer Eigenwerte in einer niedrigeren Dimension dargestellt werden.

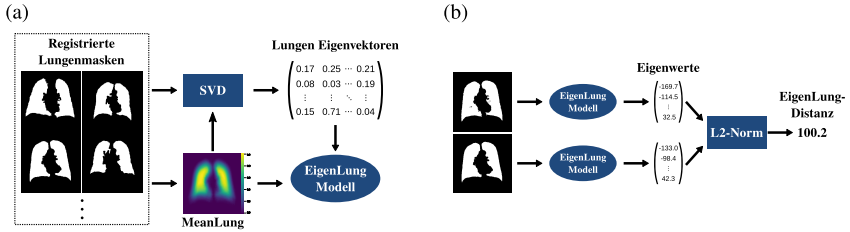


Abbildung 5.13: Darstellung der einzelnen Schritte für die Berechnung der EigenLung-Distanz Metrik. (a) Berechnung des EigenLung-Modells aus Lungenmasken (b) Berechnung der EigenLung-Distanz zwischen zwei Lungenmasken.

Die neue Metrik EigenLung-Distanz (ELD) wurde wie in Abbildung 5.13 (b) beschrieben berechnet. Für zwei Lungenmasken, die miteinander verglichen werden sollen, wurden jeweils die Eigenwerte basierend auf dem EigenLung-Modell berechnet. Die EigenLung-Distanz ergab sich dann aus der L2-Norm der beiden Eigenwert-Vektoren. Bei MRT-Bildern kommt es aufgrund von partiellen Volumeneffekten oft zu einer leichten Verschiebung der Kante von Objekten. Aus diesem Grund wurde ebenso eine unscharfe (fuzzy) Version dieser Metrik entwickelt. Hierzu wurden von der zu untersuchenden Maske zwei alternative Varianten generiert, die den Rand der Lunge jeweils um einen Voxel nach innen bzw. einen Voxel nach außen verschieben. Die Fuzzy EigenLung-Distanz (FELD) berechnete sich dann als die minimale EigenLung-Distanz von den drei generierten Maskenvarianten mit der Referenzmaske. Bei beiden Metriken entspricht ein höherer Wert einer schlechteren Segmentierung.

In Abbildung 5.14 sind für verschiedene künstlich erzeugte Beispielfälle die berechneten ELD- und FELD-Werte dargestellt. Man erkennt, dass die ELD gut zwischen verschieden ausgeprägten Fehlern skalierte. Der ELD-Wert war beispielsweise für ein kleines Loch am Rand kleiner, als für ein großes Loch am Rand. Ebenso waren die ELD-Werte für das Fehlen von nur einem unteren Lungenteil im Vergleich ca. halb so groß, als wenn beide unteren Lungenteile fehlten. Es zeigt sich jedoch auch, dass die normale Metrik (ELD) zu hohen Werten führte, wenn die Lungenmaske um nur je einen Voxel am Rand geschrumpft war. Diese Werte

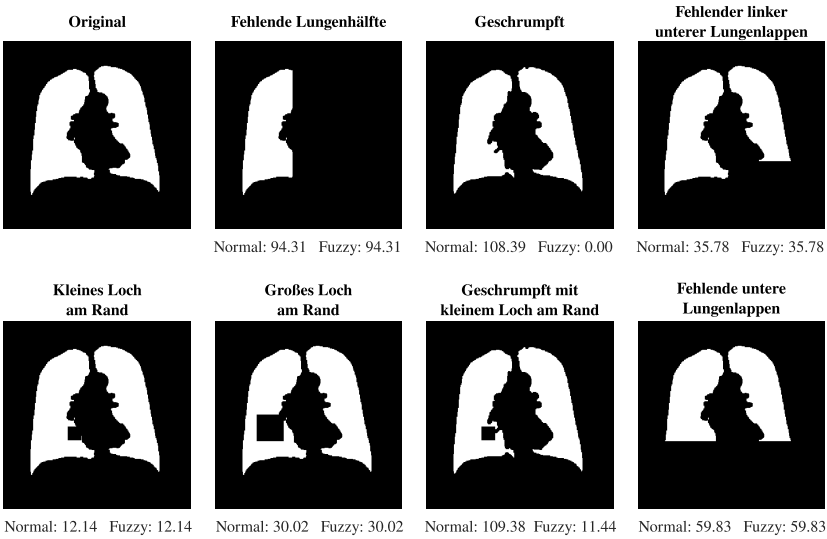


Abbildung 5.14: Vergleich der EigenLung-Distanz Metrik für verschiedene Segmentierungsfehler.

waren ungefähr in der gleichen Größenordnung wie für den Fall, in dem eine komplette Lungenhälfte fehlte. Dieses Problem wurde jedoch mit der unscharfen Variante (FELD) behoben. Hier erreichte die geschrumpfte Lunge einen perfekten Score, während die Maske mit fehlender Lungenhälfte weiterhin einen schlechten Score erhielt. Basierend auf diesen Erkenntnissen lässt sich feststellen, dass die EigenLung-Distanz, insbesondere in der unscharfen Version, eine gute Ergänzung zu den herkömmlichen Metriken ist. Ein Nachteil dieser Methode ist jedoch, dass sie im Vergleich zu vielen der anderen Segmentierungsmetriken nicht normiert ist und die Werte daher schwerer zu interpretieren sind.

5.10 Validierung der algorithmischen Lungensegmentierung

Die Wahl der passenden Segmentierungsmetriken ist für die Evaluation der Qualität eines neuen Segmentierungsalgorithmus essenziell. In der Literatur wird die Kombination verschiedener Arten von Metriken empfohlen, die jeweils unterschiedliche Eigenschaften untersuchen. Eine gute Kombination ist beispielsweise die Verwendung einer volumenbasierten und einer randbasierten Metrik. [165] Die Berechnung der meisten Segmentierungsmetriken benötigt eine Art von Groundtruth, also einer richtigen Referenzsegmentierung. Im vorliegenden Datensatz lagen jedoch keine solchen Groundtruth Segmentierungen für den Großteil der Lungen-MRT Bilder vor, wodurch eine alternative Referenz verwendet wurde. Hierbei wurde der Studienaufbau der multizentrischen COSYCONET-Studie [166] ausgenutzt, die detailliert in Abschnitt 5.12.1 vorgestellt wird. In der COSYCONET Unterstudie „MR-COPD“ (NCT02629432) [167] wurden CT- und MRT-Messungen am selben Tag mit einheitlichen Protokollen durchgeführt. Der quantitative CT Analyzer YACTA (Yet Another CT Analyzer) [168] besitzt eine gründlich validierte Methode zur Lungensegmentierung in CT Bildern [29]. Daher wurden die YACTA Segmentierungen der Lunge aus den CT-Bildern gewählt und auf das MRT-Bild transferiert. Hierfür wurde das inspiratorische CT-Bild auf das morphologische VIBE MRT-Bild mithilfe von *ANTs* [160] registriert. Das CT-Bild wurde hierbei als bewegliches Bild und das MRT-Bild als statisches Bild definiert. Zunächst wurde in einem mehrstufigen Prozess eine affine Registrierung durchgeführt und anschließend eine mehrstufige Kombination einer affinen und deformierbaren Registrierung („SyN“) verwendet. Die daraus resultierende Transformation wurde sodann verwendet, um die CT-Lungenmaske auf das VIBE-Bild zu registrieren. Bei der Transformation der Maske wurde eine k-nächste Nachbar-Interpolation verwendet.

Die registrierte CT-Lungenmaske konnte anschließend für die Berechnung von verschiedenen Segmentierungsmetriken verwendet werden. Als Metriken wurden volumenbasierte Metriken wie der DICE-Score (DSC), randbasierte Metriken wie

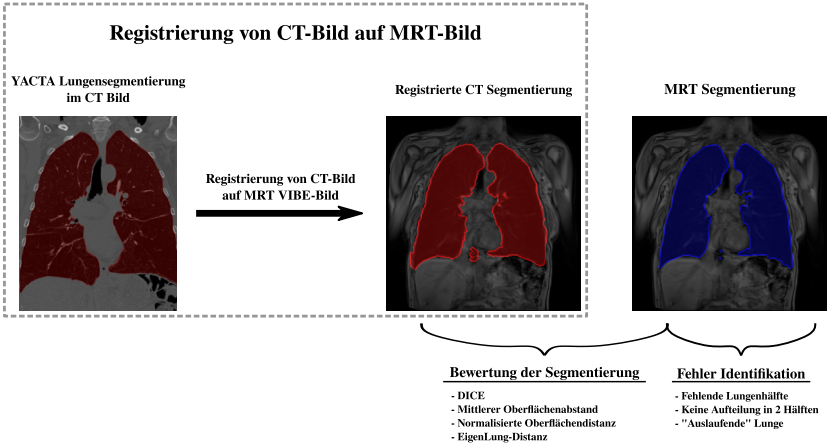


Abbildung 5.15: Vorgehen zur Erzeugung von Referenzsegmentierungen für die Validierung der algorithmischen Lungensegmentierungen.

der mittlere Oberflächenabstand (MASD) und die normalisierte Oberflächendistanz (NSD) sowie die in dieser Arbeit neu entwickelte EigenLung-Distanz (ELD) verwendet. Zusätzlich wurden jeweils die unscharfen Varianten des DICE-Score (FDSC), des mittleren Oberflächenabstands (FMASD) und der EigenLung-Distanz (FELD) berechnet. Diese unscharfen Versionen wurden genauso berechnet, wie in Abschnitt 5.9 für die FELD-Metrik beschrieben. Unabhängig von den Segmentierungsmetriken wurden auch Algorithmen zur Detektion typischer Segmentierungsfehler implementiert. Hierbei wurden Fälle wie das Fehlen einer kompletten Lungenhälfte, das Verschmelzen der beiden Lungenhälften und das „Auslaufen“ der Lunge detektiert. Beim Auslaufen der Lunge handelt es sich um einen typischen Fehler der verwendeten Floodfill-Operation, bei dem die Segmentierung der Lungen in den Hintergrund läuft, da dunkle Verbindungen zwischen diesen beiden Bereichen im Bild existieren. Das komplette Vorgehen zur Validierung der Lungensegmentierung ist in Abbildung 5.15 dargestellt. Die registrierten CT-Lungenmasken wurden als Referenz verwendet, um verschiedene Segmentierungsmetriken zu berechnen. Zusätzlich wurden direkt auf der

MRT-Segmentierung Tests durchgeführt, die typische Segmentierungsfehler identifizieren.

5.11 Monozentrische Perfusionsquantifizierung in CF

Dieses Unterkapitel basiert auf der eigenen Publikation mit dem Titel „Quantitative MRI Detects Delayed Perfusion and Impact of Bronchial Artery Dilatation on Pulmonary Circulation in Patients with Cystic Fibrosis“ [169]. Die untersuchte Studienkohorte wurde in einem monozentrischen Rahmen mit einem einheitlichen MRT-Protokoll aufgenommen. Der besondere Fokus dieser Studie lag in der Entwicklung und Validierung der quantitativen Bestimmung der Perfusionsverzögerung sowie in der allgemeinen Bestimmung von Korrelationen der quantitativ berechneten Perfusionswerte mit klinischen Parametern.

5.11.1 Klinische Einordnung

Die Bronchialarterien sind kleine Gefäße (Durchmesser < 2 mm), die der Aorta entspringen und das Lungengewebe mit Blut und Nährstoffen versorgen. Schwere Lungenblutungen, die meist lebensbedrohlich sind, werden in fast 90 % der Fälle von dilatierten Bronchialarterien verursacht. Daher ist die frühe Identifikation dilatierter Bronchialarterien (BAD) essenziell, um diese frühzeitig behandeln zu können und damit einer Lungenblutung vorzubeugen. Im Zuge ihrer Erkrankung entwickeln PatientInnen mit CF überdurchschnittlich oft solche BADs. [170] In Abbildung 5.16 ist eine dilatierte Bronchialarterie visualisiert. Die BAD lässt sich dabei allein aufgrund der Vergrößerung des Gefäßes in der MR Angiographie darstellen. Bisher war nur wenig über den quantitativen Beitrag der Bronchialarterien zur gesamten Perfusion bekannt. Die Hypothese dieser Studie war, dass die quantitative Auswertung des DCE-MRT die Perfusionsverzögerung berechnen

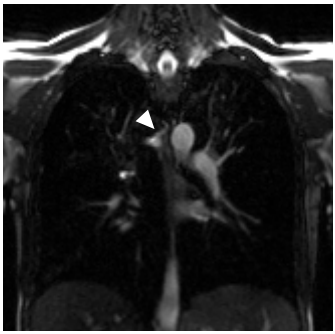


Abbildung 5.16: Visualisierung einer dilatierten Bronchialarterie (weiße Pfeilspitze) in der MR Angiographie. Physiologische Bronchialarterien sind im MRT nicht sichtbar, erst durch die Dilatation können BADs in der MR Angiographie dargestellt werden.

und damit die Perfusionsanteile von pulmonaler arterieller Perfusion und die Perfusion durch die Bronchialarterien eventuell unterschieden werden kann.

5.11.2 Studienkohorte

Die Studienkohorte bestand aus insgesamt 75 PatientInnen mit zystischer Fibrose. Hiervon hatten 48 PatientInnen mindestens eine diagnostizierte dilatierte Bronchialarterie, während die restlichen 27 PatientInnen keine dilatierten Bronchialarterien besaßen. Einen Überblick über die Kohorte gibt Tabelle 5.3. Alle PatientInnen wurden mit einem einheitlichen MRT-Protokoll [45] gemessen, das verschiedene T1- und T2-gewichtete MRT-Sequenzen sowie ein DCE-MRT enthielt. Zusätzlich wurden Spirometrie und Gasauswaschverfahren zur klinischen Diagnostik eingesetzt. Hierbei wurden die Lungenbelüftung (LCI) und das forcierte exspiratorische Volumen in einer Sekunde in Prozent vom Soll (ppFEV1) bestimmt. Die MRT-Messungen wurden mit klinischen 1,5 T MRT-Scannern (Siemens Avanto und Siemens Aera) durchgeführt.

Tabelle 5.3: Überblick über die CF-BAD Studienkohorte. Daten sind dargestellt als Prozentsatz, Mittelwert ± Standardabweichung oder Median ± Interquartilsabstand für die MRT-Scores. LCI2% : Anzahl Lungenumsätze bis 98% des Markergases ausgespült wurden, ppFEV1: Forciertes exspiratorisches Volumen in einer Sekunde in Prozent vom Soll, MBW: Gasauswaschverfahren. * $p < 0{,}05$, ** $p < 0.01$ und *** $p < 0.001$ jeweils im Vergleich zu ohne BAD.

	Gesamt	**Ohne BAD**	**BAD**
Anzahl	75	27	48
Alter (Jahre) (Bereich)	$18{,}7 \pm 7{,}6$ (6–39)	$12{,}7 \pm 3{,}7$ (6–18)	$22{,}1 \pm 7{,}2$ (7–39)***
Männlich/Weiblich	45/30	15/12	30/18
Größe (cm)	$161{,}1 \pm 16{,}1$	$153{,}3 \pm 19{,}9$	$165{,}4 \pm 11{,}4$*
Größe, SDS	$0{,}0 \pm 1{,}2$	$0{,}0 \pm 1{,}1$	$-0{,}1 \pm 1{,}3$
Gewicht (kg)	$50{,}9 \pm 15{,}4$	$44{,}0 \pm 17{,}1$	$54{,}7 \pm 12{,}9$*
Gewicht, SDS	$-0{,}6 \pm 0{,}9$	$-0{,}5 \pm 0{,}8$	$-0{,}8 \pm 1{,}0$
Spirometrie, *n*	64	27	37
ppFEV1	$73{,}0 \pm 18{,}3$	$84{,}8 \pm 16{,}0$	$64{,}4 \pm 14{,}8$***
MBW, *n*	24	21	3
LCI2%	$8{,}5 \pm 2{,}4$	$8{,}2 \pm 2{,}2$	$10{,}7 \pm 2{,}8$
Morphologie MRT-Score	$14{,}0 \pm 12{,}7$	$8{,}0 \pm 3{,}8$	$17{,}5 \pm 9{,}5$***
Perfusions MRT-Score	$6{,}0 \pm 3{,}0$	$5{,}0 \pm 3{,}8$	$7{,}5 \pm 3{,}0$***
Globaler MRT-Score	$22{,}0 \pm 16{,}0$	$13{,}0 \pm 5{,}8$	$25{,}0 \pm 12{,}0$***

5.11.3 Auswertung der MRT Messungen

Die gemessenen MRT-Bilder wurden mithilfe des Eichinger Scores [45] von einem Radiologen mit mehr als 15 Jahren Erfahrung bewertet. Quantitative MRT-Werte wurden mit der zuvor beschriebenen Perfusionsquantifizierungs-Pipeline berechnet. Die Lungen wurden zunächst in der morphologischen VIBE-Messung segmentiert und anschließend mithilfe der *elastix*-Registrierung auf das DCE-MRT registriert. Die Pulmonalarterie wurde automatisch detektiert und die Residuenfunktion basierend auf der Indikator-Verdünnungs-Theorie mittels tSVD bestimmt. Daraus wurden anschließend die quantitativen MRT Werte PBF, MTT und QDP bestimmt. Zusätzlich wurde die Perfusionsverzögerung für jede Messung anhand

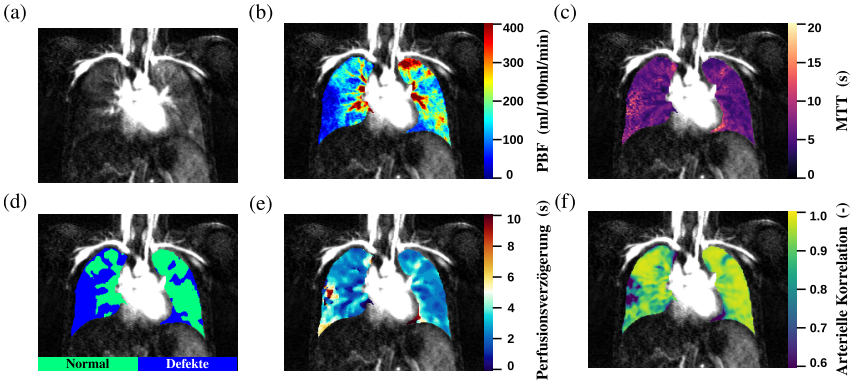


Abbildung 5.17: Quantitative Perfusionskarten eines 13-jährigen männlichen CF-Patienten ohne dilatierte Bronchialarterien: (a) Subtraktionsbild (b) PBF (c) MTT (d) Perfusionsdefekte (e) Perfusionsverzögerung (f) Arterielle Korrelation.

des in Abschnitt 5.7 beschriebenen Vorgehens mit Gleichung (5.18) berechnet. Darüber hinaus wurde die arterielle Korrelation r_{art} als weiterer quantitativer Wert definiert, die sich als die berechnete maximale Korrelation eines Voxels mit der AIF ergibt:

$$r_{art}(x,y,z) = \max(\mathbf{r}(x,y,z,t)) \ . \tag{5.23}$$

Die arterielle Korrelation ist hierbei eine Messung der Ähnlichkeit des Signals eines Voxels mit der AIF und stellt eine Möglichkeit dar, zwischen dem arteriellen Anteil und dem Bronchialarterien-Anteil der Perfusion zu unterscheiden.

In Abbildung 5.17 sind die quantitativen Perfusionskarten für einen CF-Patienten ohne dilatierte Bronchialarterien dargestellt. Man erkennt, dass die Bereiche mit Perfusionsdefekten (blaue Bereiche in Abbildung 5.17 (d)) kleinere PBF-Werte, höhere MTT-Werte, längere Perfusionsverzögerungen und kleinere arterielle Korrelationen besaßen.

Für alle PatientInnen wurde jeweils der Median innerhalb der Lunge für die quantitativen Metriken PBF, MTT, Perfusionsverzögerung und arterielle Korrelation berechnet. Anschließend wurde für verschiedene Gruppierungsarten jeweils der Mittelwert der Mediane berechnet. Zusätzlich wurden für jede untersuchte

Person Histogramme der quantitativen Parameter berechnet. Für verschiedene Gruppierungsarten wurden diese Histogramme dann jeweils zu einem gemittelten Histogramm zusammengefügt, indem jeweils die Mittelwerte der Bins berechnet wurden. Die Signifikanz zwischen zwei Gruppen wurde je nach Anwendungsfall für die Mediane der Perfusionswerte mittels eines Wilcoxon-Vorzeichen-Rang-Tests bzw. eines Wilcoxon-Rangsummentests berechnet. Bei den Histogrammen wurde zur Berechnung der Signifikanz das Bootstrapping-Verfahren [171] verwendet.

5.11.4 Korrelation mit klinischen Parametern

Zunächst wurden jeweils die Spearman Korrelationen zwischen QDP, den Medianwerten von PBF, MTT und der Perfusionsverzögerung mit den Bestandteilen des MRT-Scores sowie mit ppFEV1 und LCI berechnet. Die Korrelationen sind in Tabelle 5.4 dargestellt. Sowohl QDP und PBF korrelierten moderat und signifikant mit allen drei MRT-Scores und mit ppFEV1 ($p < 0.001$). Die Perfusionsverzögerung korrelierte schwach und signifikant mit den MRT-Scores ($p < 0.001$) und mit ppFEV1 ($p < 0.05$). Diese Resultate decken sich sehr gut mit den Ergebnissen von Schiwek et al. [57]. Während die Korrelationen zwischen PBF und dem Perfusions MRT-Score in der aktuellen Studie leicht stärker waren, fielen die Korrelation mit QDP etwas schwächer aus als in der Arbeit von Schiwek et al. Jedoch fielen die in Tabelle 5.4 berechneten Korrelationen von QDP und dem Perfusions MRT-Score stärker als in der Arbeit von Konietzke et al. [56] aus, in der ebenso CF-PatientInnen untersucht wurden. Einen Erklärungsansatz hierfür könnte die Erkrankung selbst sein, die möglicherweise zu unterschiedlich starken Korrelationen in COPD- und CF-PatientInnen führt. Ein wesentlicher Unterschied zwischen COPD- und CF-PatientInnen ist außerdem meist, dass insbesondere die jungen CF-PatientInnen häufig während der Messung atmen. Ein weiterer Grund für Diskrepanzen bei den Korrelationen könnte die grobe Granularität des MRT-Scores sein. Diese macht sich insbesondere in Randfällen bemerkbar, in denen ein QDP-Wert von 49 % einen Score von 1 erhalten würde, während ein QDP-Wert von 51 % einen Score von 2 erhält. In der aktuellen Studie besaßen

Tabelle 5.4: Spearman Korrelationen zwischen den quantitativen MRT-Werten und den MRT-Scores, ppFEV1 und LCI für die CF-PatientInnen in dieser Studie.
* $p < 0,05$ und *** $p < 0.001$

	QDP (-)	**PBF** (ml/100ml/min)	**MTT** (s)	**Perfusions-verzögerung** (s)
Morphologie MRT-Score, *r*	0.66***	-0.56***	-0.11	0.46***
| 95% CI |	0.53, 0.75	-0.68, -0.41	-0.29, 0.09	0.29, 0.60
Perfusions MRT-Score, *r*	0.54***	-0.56***	-0.09	0.50***
| 95% CI |	0.38, 0.66	-0.68, -0.41	-0.28, 0.10	0.34, 0.63
Globaler MRT-Score, *r*	0.66***	-0.59***	-0.12	0.49***
| 95% CI |	0.53, 0.75	-0.70, -0.45	-0.31, 0.07	0.33, 0.62
ppFEV1	-0.48***	0.45***	0.07	-0.30*
| 95% CI |	-0.62, -0.30	0.27, 0.60	-0.14, 0.27	-0.47, -0.09
LCI	0.20	-0.33	0.20	0.36
| 95% CI |	-0.15, 0.51	-0.61, 0.01	-0.15, 0.51	0.02, 0.63

17 PatientInnen einen QDP-Wert zwischen 40–50 %, wobei fünf dieser PatientInnen sogar einen QDP-Wert zwischen 49–51 % besaßen. Wurden nur ein paar dieser PatientInnen beim MRT-Score in die jeweils andere Kategorie gepackt, so könnte dies die beobachteten Abweichungen in der Korrelation erklären. Die fehlenden Korrelationen mit LCI könnten entweder an der geringen Anzahl an LCI-Messungen liegen, da diese für weniger als ein Drittel der untersuchten Personen vorlagen. Außerdem könnte es sein, dass kleine funktionale Veränderungen in der Lunge von LCI nicht identifiziert werden können, während die quantitativen Perfusionsmetriken sensitiv auf diese kleinen Änderungen reagieren.

Grundsätzlich zeigen die Korrelationen, dass PBF, aber insbesondere auch QDP robuste Metriken für die Quantifizierung der Perfusion sind und eine sinnvolle komplementäre Metrik für die klinische Diagnostik darstellen. Ein wesentlicher

Tabelle 5.5: Quantitative Perfusionsparameter aufgeteilt nach normalen und defekten Perfusionsbereichen für die Gesamtkohorte sowie die zwei Gruppen mit und ohne BAD. Dargestellt als Mittelwert ± Standardabweichung. * $p < 0,05$, ** $p < 0,01$ und *** $p < 0.001$ im Vergleich zu ohne BAD, ### $p < 0,001$ im Vergleich zu alle normal perfundiert.

		PBF (ml/100ml/min)	MTT (s)	Perfusions-verzögerung (s)	Arterielle Korrelation (-)
Gesamt	Alle	123,0 ± 75,9	5,7 ± 1,7	6,8 ± 7,2	0,87 ± 0,08
	Ohne BAD	178,3 ± 77,4	6,2 ± 2,2	5,2 ± 5,7	0,91 ± 0,06
	BAD	91,9 ± 54,8***	5,5 ± 1,2	7,8 ± 7,8**	0,85 ± 0,08***
Normal	Alle	161,2 ± 77,9	5,4 ± 1,7	4,6 ± 5,3	0,91 ± 0,07
	Ohne BAD	198,5 ± 80,7	6,0 ± 2,1	4,2 ± 2,5	0,92 ± 0,06
	BAD	140,2 ± 68,6***	5,1 ± 1,2*	4,9 ± 6,4	0,90 ± 0,07
Defekte	Alle	57,5 ± 26,4###	6,9 ± 2,3###	13,4 ± 16,2###	0,78 ± 0,07###
	Ohne BAD	75,1 ± 27,0	8,0 ± 2,6	12,8 ± 23,7	0,80 ± 0,08
	BAD	47,7 ± 20,5***	6,3 ± 1,9***	13,8 ± 10,1*	0,77 ± 0,06

Vorteil dieser Methode ist, dass QDP feingranularer als der Perfusions MRT-Score ist und dass die Ergebnisse keiner Interobserver-Variabilität unterliegen.

5.11.5 Zusammenhang von Perfusionsdefekten und quantitativen Pefusionswerten

In einem nächsten Schritt sollte untersucht werden, wie sich die quantitativen Perfusionswerte zwischen Bereichen mit und ohne Perfusionsdefekten unterscheiden. In Tabelle 5.5 sind jeweils die Mittelwerte und Standardabweichungen für die einzelnen quantitativen Perfusionswerte dargestellt und nach Bereichen mit und ohne Perfusionsdefekten sowie nach den Gruppen der PatientInnen mit und ohne BAD unterteilt worden.

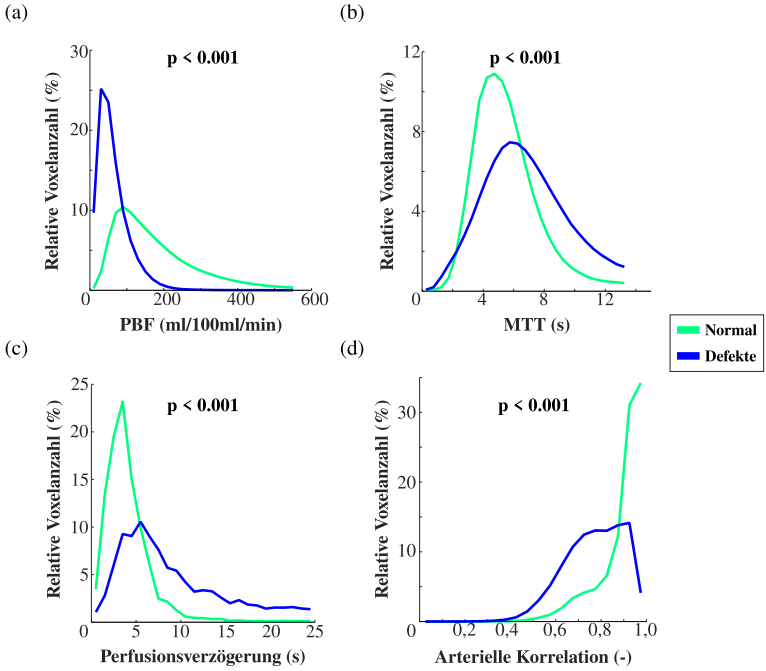


Abbildung 5.18: Mittlere Histogramme der quantitativen Perfusionswerte für normale und defekte Perfusionsbereiche. (a) PBF (b) MTT (c) Perfusionsverzögerung (d) Arterielle Korrelation.

Für alle vier Perfusionswerte ergaben sich signifikante Unterschiede zwischen normalen und defekten Perfusionsbereichen. Innerhalb der Perfusionsdefekte war PBF geringer, die mittlere Transitzeit MTT war länger und es gab längere Perfusionsverzögerungen sowie eine geringere arterielle Korrelation. Diese Beobachtung wird auch durch Betrachtung von Abbildung 5.18 bestätigt, in der die mittleren Histogramme für die normalen und defekten Bereiche dargestellt sind. Für PBF war das grüne Histogramm der normalen Bereiche breiter und in Richtung höherer PBF-Werte verschoben. Bei MTT und der Perfusionsverzögerung war dagegen jeweils das blaue Histogramm der Perfusionsdefekte breiter und in Richtung höherer Zeiten verschoben. Insbesondere bei der arteriellen Korrelationen waren deutliche Unterschiede in der Form der Histogrammkurven zu erkennen. Während

für die normalen Bereiche fast alle Voxel den Bins oberhalb von 0,8 zugewiesen wurden, waren die Voxel für die defekten Bereiche deutlich stärker gestreut und verteilten sich auf die Bins zwischen 0,5–0,9.

Eine Ursache für die verspätete Perfusion in den Perfusionsdefekten könnte sein, dass die hier gemessene Perfusion die schwerpunktmäßige oder reine Perfusion durch die Bronchialarterien darstellt, weil diese Bereiche aufgrund von Lungenschädigung nicht mehr richtig durchblutet werden. Da die bronchialarterielle Perfusion standardmäßig später stattfindet als die arterielle Perfusion, wird daher eine Perfusionsverzögerung beobachtet. Für die sehr hohen Perfusionsverzögerungen > 20 s könnte die gemessene Perfusion auch von einer Rezirkulation im pulmonalen Gefäßbett ausgehend von den Pulmonalarterien stammen.

5.11.6 Vergleich von PatientInnen mit und ohne BAD

Nachdem zuvor nachgewiesen werden konnte, dass Perfusionsdefekte andere quantitative Perfusionswerte besitzen als normal perfundierte Bereiche der Lunge, sollte nun untersucht werden, welche Rolle dilatierte Bronchialarterien für die Perfusion spielen. In Abbildung 5.19 sind die Karten der quantitativen Perfusionsparameter für einen CF-Patienten mit BAD dargestellt. Man erkennt hier die verzögerte Perfusion in den Bereichen der Perfusionsdefekte deutlich. Außerdem besaßen die Perfusionsdefekte auch eine geringere arterielle Korrelation im Vergleich zu normalen Bereichen.

Vergleicht man die PatientInnen mit und ohne BAD in Tabelle 5.5, so hatten die PatientInnen mit dilatierten Bronchialarterien signifikant höhere QDP-Werte ($44{,}6 \pm 20{,}8$ vs. $17{,}3 \pm 11{,}0$), signifikant kleinere PBF-Werte ($91{,}9 \pm 54{,}8$ vs. $178{,}3 \pm 77{,}4$ ml/100ml/min) und kürzere MTT-Werte ($5{,}5 \pm 1{,}2$ vs. $6{,}2 \pm 2{,}2$ s). Dagegen war die Perfusionsverzögerung bei PatientInnen mit BAD im Mittel signifikant höher ($7{,}8 \pm 7{,}8$ vs. $5{,}2 \pm 5{,}7$ s) und die arterielle Korrelation signifikant kleiner ($0{,}85 \pm 0{,}08$ vs. $0{,}91 \pm 0{,}06$). Schaut man sich nur die Bereiche der Perfusionsdefekte an, so hatten PatientInnen mit BAD auch hier signifikant niedrigere PBF-Werte ($47{,}7 \pm 20{,}5$ vs. $75{,}1 \pm 27{,}0$ ml/100ml/min) und hier sogar

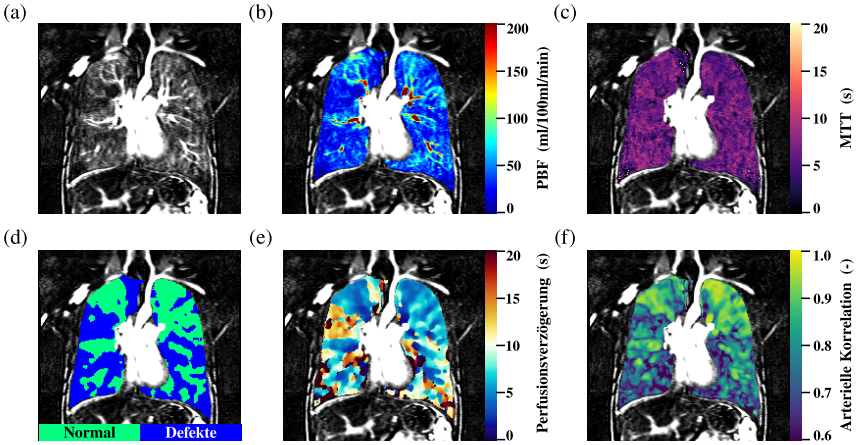


Abbildung 5.19: Quantitative Perfusionskarten eines 23-jährigen männlichen CF-Patienten mit dilatierten Bronchialarterien: (a) Subtraktionsbild (b) PBF (c) MTT (d) Perfusionsdefekte (e) Perfusionsverzögerung (f) Arterielle Korrelation.

auch signifikant niedrigere MTT-Werte ($6{,}3 \pm 1{,}9$ vs. $8{,}0 \pm 2{,}6$ s). Die Perfusionsverzögerungen sind in den Defektbereichen für die PatientInnen mit BAD ebenso größer, jedoch nur für das kleine Signifikanzniveau ($13{,}8 \pm 10{,}1$ vs. $12{,}8 \pm 23{,}7$ s), während es bei den arteriellen Korrelationen keinen signifikanten Unterschied gab.

Betrachtet man die gemittelten Histogramme für die vier Gruppen in Abbildung 5.20 so erkennt man für PBF deutlich unterschiedliche Verläufe. Während die beiden Histogramme für die Defektbereiche sehr schmal waren und ihr Maximum bei sehr niedrigen PBF-Werten hatten, waren die Histogramme für die normalen Bereiche deutlich breiter, wobei das Histogramm der normalen Bereiche für die PatientInnen ohne BAD am breitesten war. Auch für MTT unterschieden sich die Histogrammverläufe deutlich. Eine interessante Beobachtung konnte für die Perfusionsverzögerung gemacht werden. Während die beiden Histogramme der normalen Bereiche nahezu identisch waren, unterschieden sich die Histogramme für die defekten Bereiche jedoch signifikant. Das Histogramm der PatientInnen mit BAD zeigte eine Verschiebung hin zu höheren Perfusionsverzögerungen. Auch

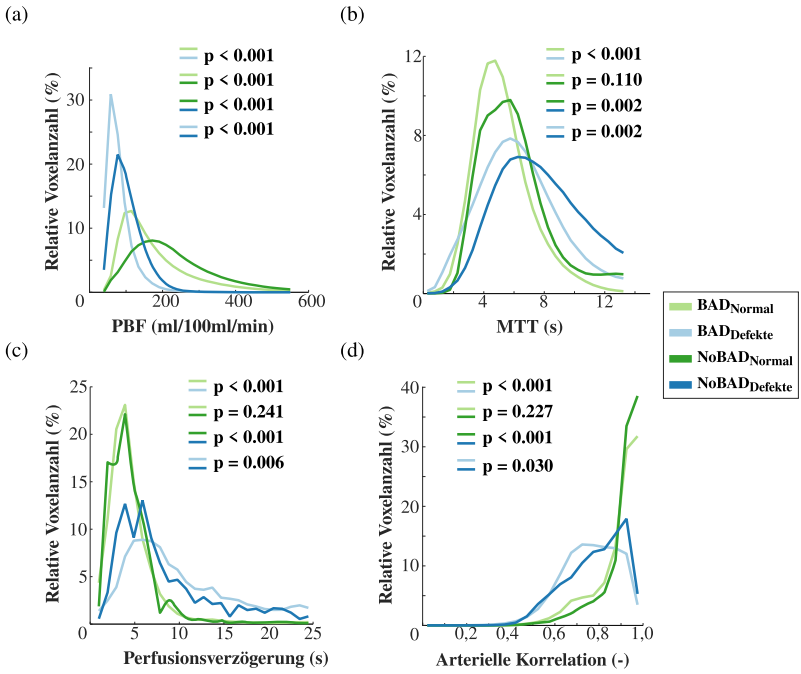


Abbildung 5.20: Mittlere Histogramme der quantitativen Perfusionswerte für PatientInnen mit und ohne BAD jeweils für normale und defekte Perfusionsbereiche: (a) PBF (b) MTT (c) Perfusionsverzögerung (d) Arterielle Korrelation.

für die arterielle Korrelation waren die Histogramme der normalen Bereiche fast identisch, während die Histogramme der Defektbereiche Unterschiede aufwiesen. Das Histogramm der BAD PatientInnen wies hier eine Verschiebung hin zu kleineren Korrelationen auf und hatte einen breiteren maximalen Bereich.

Die beobachteten längeren Perfusionsverzögerungen für PatientInnen mit dilatierten Bronchialarterien könnten zwei mögliche Ursachen haben. Basierend auf der Kontinuitätsgleichung

$$\dot{V} = vA \tag{5.24}$$

führt bei einem konstanten Volumenstrom $\dot{V}$ eine Vergrößerung des Querschnitts A zu einer niedrigeren Fließgeschwindigkeit v, was wiederum zu einer verzögerten

Ankunft des Kontrastmittels führen kann. Ebenso könnte der sehr gewundene Verlauf von dilatierten Bronchialarterien [172] zu Veränderungen des laminaren Flusses und dadurch zu einer verzögerten Perfusion durch die Bronchialarterie führen. Die kürzeren mittleren Transitzeiten in Defektbereichen für die BAD-PatientInnen bedeuten einen schnelleren Durchlauf des Kontrastmittels durch die Perfusionsdefekte. Basierend auf diesen Beobachtungen besteht die Vermutung, dass die dilatierten Bronchialarterien nur einen Teil des verringerten Blutflusses kompensieren können. Der gemessene pulmonale Blutfluss ist nicht geeignet, die Anteile der arteriellen und bronchoarteriellen Perfusion voneinander zu unterscheiden, da die Signale sehr ähnlich sind und kurz nacheinander stattfinden. Das gemessene Signal ist immer ein Mischsignal aus beiden Anteilen, die je nach Anteil des bronchoarteriellen Blutflusses unterschiedlich stark ausfallen können.

Eine Möglichkeit das Mischsignal zumindest grob in seine Bestandteile aufzuteilen, könnte die neu vorgeschlagene Metrik der arteriellen Korrelation sein. Der Blutfluss durch die Bronchialarterien generiert ein Shuntvolumen vom Körperkreislauf in den Lungenkreislauf [54]. Shunts zwischen den Bronchialarterien und Pulmonalarterien ermöglichen einen Blutfluss von den Bronchialarterien in das pulmonale Gefäßbett und sind der Grund, warum das Signal der Bronchialarterien überhaupt gemessen werden kann. Normalerweise sind diese Shunts geschlossen, können jedoch bei chronischen Erkrankungen zur Korrektur eines beeinträchtigten Blutflusses geöffnet werden. Das gemessene DCE-MRT Signal ist daher immer eine Überlagerung der Signale des Körper- und Lungenkreislaufs. Im Datensatz konnte beobachtet werden, dass die arterielle Korrelation für PatientInnen mit dilatierten Bronchialarterien geringer ausfällt, was auf einen größeren Beitrag des bronchoarteriellen Perfusionsanteils bzw. des Anteils aus dem Körperkreislauf schließen lässt. Eine Erweiterung der Bronchialarterien führt nämlich zu einem erhöhten Blutfluss und Shuntvolumen und damit zu einem vergrößerten Perfusionsanteil des Körperkreislaufes. Dies bedeutet zusätzlich, dass die arterielle Korrelation möglicherweise eine Messgröße zur frühen Detektion dilatierter Bronchialarterien sein könnte, bevor diese im regulären MRT nachgewiesen werden können.

5.11.7 Einordnung und Zusammenfassung der Ergebnisse

Die hier durchgeführte Studie besitzt Limitationen. Aufgrund des teilweise jungen Alters der PatientInnen und der Schwere der Erkrankung konnten nicht alle während der Messung ihren Atem anhalten. Diese Atmung kann, wie zuvor schon erwähnt, die voxelweise Perfusionsquantifizierung negativ beeinflussen. Daher wäre die Entwicklung einer Methode zur Korrektur der Atmung während der Boluspassage sehr nützlich. Außerdem wurde die Lage der Bronchialarterien in dieser Studie nicht berücksichtigt. Damit hätten manche bronchoarteriellen Perfusionsbeiträge eventuell aufgrund der Anatomie der Bronchialarterie ausgeschlossen werden können. Die Studie konnte jedoch zeigen, dass die entwickelte Perfusionsquantifizierungs-Pipeline vollautomatisch und ohne äußere Einflüsse in der Lage ist, quantitative Perfusionsmetriken zu berechnen. Insbesondere wurde erneut das Potenzial von QDP als quantitativem, feingranularem Messwert für Perfusionsdefekte aufgezeigt. Bisher wurde die Perfusionsquantifizierungsmethode nur an einer kleinen Studie mit 15 CF-PatientInnen getestet [56]. Diese Studie hat nun gezeigt, dass auch die Verwendung in einer deutlich größeren monozentrischen CF-Kohorte möglich ist. Eine weitere wichtige Erkenntnis war der erfolgreiche Einsatz der beiden neuen quantitativen Perfusionsmetriken der Perfusionsverzögerung und der arteriellen Korrelation, die sich als nützliche, komplementäre Metriken herausgestellt haben.

5.12 Multizentrische Perfusionsquantifizierung in COPD

Dieses Unterkapitel dient der Validierung der entwickelten Pipeline zur Perfusionsquantifizierung in einer multizentrischen Kohorte. Insbesondere ist das Ziel dieses Kapitels die Robustheit der Methode zu testen und den Einfluss verschiedener MRT-Scanner auf die Auswertung zu untersuchen.

5.12.1 Studienkohorte

Die verwendete Studienkohorte entstammt einer Unterstudie der deutschlandweiten multizentrischen COPD Studie COSYCONET (NCT01245933). Die COSYCONET-Studie ist eine große multizentrische Studie mit COPD-PatientInnen mit verschiedenen GOLD-Klassen. Die Diagnose für COPD basierte auf dem standardmäßig verwendeten Schwellwert des Quotienten von FEV1 und FVC mit $\mathrm{FEV1}/\mathrm{FVC} < 0,7$. Die GOLD-Klassen beinhalteten sowohl die regulären GOLD-Klassen 1-4, als auch die alte GOLD-Klasse 0, die alle PatientInnen erhielten, die oberhalb des Schwellwerts von 0,7 lagen, aber dennoch COPD-spezifische Symptome zeigten. Zusätzlich wurden auch PatientInnen mit einem Risiko für COPD inkludiert, sofern diese keine der fünf GOLD-Klassen zuzuordnen waren. [166] Die detaillierten Einschlusskriterien für die COSYCONET-Studie sind im Artikel von Jörres et al. [173] beschrieben. Die COSYCONET-Studie enthält außerdem Blutproben, Lungenfunktionstests, Gesundheitsfragebögen und weitere Informationen von mehr als 2 000 COPD-PatientInnen.

Die Teilstudie „MR-COPD“ (NCT02629432) enthält CT- und MRT-Messungen des gleichen Tages von einer Teilkohorte mit 602 PatientInnen, die an 15 verschiedenen Studienstandorten mit einem einheitlichen Protokoll aufgenommen wurden. Das Hauptziel dieser Teilstudie war zu untersuchen, ob für die strukturelle und funktionale Phänotypisierung der COPD in Zukunft das CT durch MRT ergänzt bzw. ersetzt werden kann. An jedem Zentrum wurde ein institutionelles Ethikvotum eingeholt und alle PatientInnen haben ihre schriftliche Einwilligung zur Datenverarbeitung erteilt. [167] Von der ursprünglichen Teilkohorte aus 602 PatientInnen wurden insgesamt 86 PatientInnen von der nachfolgenden Auswertung ausgeschlossen, da diese eine unzureichende Bildqualität besaßen oder Daten zur Auswertung fehlten. Damit ergab sich eine Kohortengröße von 516 PatientInnen. Einen Überblick über die Studienkohorte liefert Tabelle 5.6.

Die CT-Bilder der Studie wurden mit der quantitativen CT-Software YACTA [28, 29, 168, 174, 175] analysiert. Die berechneten QCT-Werte umfassten unter anderem $\mathrm{PRM}_{\mathrm{normal}}$, $\mathrm{PRM}_{\mathrm{fSAD}}$, $\mathrm{PRM}_{\mathrm{emph}}$, den Emphysemindex, das Lungenvolumen

Tabelle 5.6: Überblick über die COSYCONET-Studienkohorte. Daten sind dargestellt als Prozentsatz, Mittelwert ± Standardabweichung oder Median ± Interquartilsabstand für die MRT-Scores. ppFEV1: Forcierte exspiratorische Volumen in einer Sekunde in Prozent vom Soll, ppFVC: Forcierte Vitalkapazität in Prozent vom Soll, DLCO: Kohlenmonoxid-Diffusionskapazität, EI: Emphysemindex.

	COSYCONET
Anzahl	516
Alter (Jahre) (Bereich)	$65{,}5 \pm 8{,}6$ (42–84)
Männlich/Weiblich	320 / 196
GOLD (Risiko/0/1/2/3/4)	27 / 48 / 52 / 207 / 146 / 34
DLCO-Prozentsatz	$59{,}0 \pm 20{,}5$
Spirometrie, n	514
ppFEV1	$59{,}6 \pm 20{,}7$
ppFVC	$83{,}0 \pm 19{,}1$
Atemwegs-Score	$5{,}2 \pm 3{,}3$
Funktioneller Krankheits-Score	$8{,}6 \pm 3{,}0$
Globaler MRT-Score	$13{,}8 \pm 4{,}7$
$\mathbf{PRM_{normal}}$	$57{,}4 \pm 22{,}2$
$\mathbf{PRM_{fSAD}}$	$28{,}7 \pm 13{,}8$
$\mathbf{PRM_{emph}}$	$12{,}7 \pm 12{,}4$
EI	$15{,}1 \pm 13{,}4$

und das Atemwegsvolumen. Eine genaue Beschreibung und quantitative Auswertung der gemessenen Werte kann dem Paper von Konietzke et al. [30] entnommen werden.

In der MRT-Studie wurden insgesamt zehn verschiedene Varianten von MRT-Scannern des gleichen Herstellers (Siemens Healthineers, Erlangen, Deutschland) verwendet. Dabei kamen vier verschiedene Modelle (Aera, Avanto, Espree und Trio) zum Einsatz. Diese vier Modelle unterscheiden sich im Wesentlichen in der Geometrie der Bohrung und der B_0 Feldstärke. Während der Aera und der Espree jeweils eine Bohrung von 70 cm Durchmesser besitzen, fällt die Bohrung beim Avanto und Trio mit 60 cm kleiner aus. Ebenso unterscheidet sich die Länge der

Bohrung zwischen den vier Modellen. Der Trio besitzt ein Hauptmagnetfeld mit einer Feldstärke von 3 T, wohingegen die anderen drei Scanner eine Feldstärke von 1,5 T besitzen. Weitere Unterschiede sind die Stärke des Gradientensystems (33 mT/T vs. 45 mT/T), die Anzahl der Empfangskanäle und die Softwareversion. Die jeweiligen Eigenschaften der zehn Scanner-Varianten sind detailliert in Tabelle A.1 aufgelistet.

Für die MRT-Studie wurden alle PatientInnen mit dem gleichen MRT-Protokoll aufgenommen, das zuvor im Zuge einer Studie mit einem MRT-Phantom validiert wurde [176]. Die nachfolgend genutzten MRT-Sequenzen sind die koronalen und transversalen T1-gewichteten Gradientenecho Bilder (VIBE), die in einem Atemanhalt gemessen wurden sowie eine zeitaufgelöste T1-gewichtete 3D Keyhole Pulssequenz (TWIST). Die verwendete TWIST-Sequenz beinhaltete die intravenöse Gabe einer festen Dosis von 2 ml eines Gadolinium-basierten Kontrastmittels (Gadobutrol, Bayer Vital GmbH) mit einer Rate von 4 ml/s gefolgt von einem Chaser mit Kochsalzlösung [177]. In Tabelle A.2 sind die genauen Sequenzparameter dargestellt. Diese waren zwischen den einzelnen MRT-Scannern leicht unterschiedlich, um jeweils die bestmögliche Bildqualität aus der Messung zu erhalten. [176]

5.12.2 Validierung der algorithmischen Lungensegmentierung

In einem ersten Schritt wurde die COSYCONET-Kohorte verwendet, um die algorithmische Lungensegmentierung zu validieren. Hierzu wurde das in Abschnitt 5.10 beschriebene Vorgehen verwendet. In Tabelle 5.7 sind die über die gesamte Kohorte gemittelten Segmentierungsmetriken dargestellt. Als direkter Vergleich wurde auch das entwickelte HMRF-UNet mit normaler Potts-Nachbarschaft mit $\lambda_n = 0,5$ und $\sigma_{\text{thresh}} = 0,1$ auf alle Bilder der Studie angewendet. Als externes Referenzmodell diente ein überwachtes CNN-Modell von Ringwald et al. Auch dieses von Ringwald et al. trainierte nn-U-Net [106] wurde auf alle Bilder der

Tabelle 5.7: Vergleiche zwischen algorithmischer, HMRF-UNet und AI-Segmentierung aus [106]. Alle Werte sind dargestellt als Mittelwert ± Standardabweichung. ### $p < 0.001$ Algorithm. Seg vs. HMRF-UNet-Seg *** $p < 0.001$ Algorithm. Seg vs. AI-Seg

	Algorithm. Seg	**HMRF-UNet-Seg**	**AI-Seg [106]**
DSC	0,932 ± 0,027	0,883 ± 0,061$^{###}$	0,900 ± 0,016***
FDSC	0,960 ± 0,020	0,948 ± 0,051$^{###}$	0,950 ± 0,010***
MASD	2,035 ± 1,511	3,320 ± 3,127$^{###}$	2,624 ± 0,426***
FMASD	1,539 ± 1,324	2,235 ± 2,940$^{###}$	1,772 ± 0,368***
NSD	0,951 ± 0,051	0,833 ± 0,138$^{###}$	0,913 ± 0,057***
EL	89,503 ± 34,335	176,112 ± 72,776$^{###}$	147,516 ± 33,220***
FEL	46,911 ± 27,874	108,383 ± 69,416$^{###}$	93,071 ± 28,755***

Studie angewendet, um auch hiermit die Lungen zu segmentieren. Für beide Vergleichsmodelle wurden jeweils die mittleren Metriken berechnet und sind ebenso in Tabelle 5.7 dargestellt.

Bei allen berechneten Metriken schnitt die in dieser Arbeit verwendete algorithmische Segmentierung signifikant besser ab, als beide anderen Modelle. Das in dieser Arbeit entwickelte HMRF-UNet schnitt dagegen schlechter ab, als das überwachte Referenzmodell. Die Unterschiede zwischen algorithmischer und überwachter nn-U-Net-Segmentierung waren ungefähr ähnlich groß wie zwischen HMRF-UNet und dem überwachten nn-U-Net. Vergleicht man die berechneten DICE-Scores mit anderen Segmentierungsmodellen für Lungen-MRT Bilder, so schnitt die algorithmische Segmentierung teilweise besser [105, 107] und teilweise ähnlich gut bzw. leicht schlechter ab [108, 178]. Jedoch gilt es zu beachten, dass sich die Größe der Datensätze hier teilweise enorm unterscheidet und die COSYCONET-Kohorte die mit Abstand größte Testkohorte darstellte. Außerdem sind diese Ergebnisse nicht direkt vergleichbar, da jeweils unterschiedliche MRT-Sequenzen für die Lungensegmentierung verwendet wurden. Für einen direkten Vergleich der Methoden müssten alle Segmentierungsmethoden auf den gleichen Datensatz angewendet werden.

Ein wesentlicher Vorteil der hier genutzten algorithmischen Segmentierung ist, dass hierfür keine manuell annotierten Trainingsdaten nötig sind, um das Modell zu trainieren. Alle oben genannten Modelle sind überwachte ML-Modelle, die jeweils manuell segmentierte Lungenmasken zum Training benötigen. Diese sind jedoch oft nur unter erheblichem zeitlichen Aufwand zu generieren. Eine Limitation der hier verwendeten Validierungsmethode ist zudem, dass als Referenz keine manuell segmentierten Lungenmasken verwendet wurden, sondern CT-Segmentierungen auf das MRT-Bild registriert wurden. Hierdurch könnten sich Fehler aus der Registrierung direkt auf die Berechnung der Metriken auswirken.

Die schlechteren Ergebnisse des HMRF-UNet könnten unterschiedliche Ursachen haben. Zum einen könnte das Hinzufügen der Trachea zur Lungenmaske eine mögliche Ursache sein. Um dies zu vermeiden, besitzt die hier vorgestellte algorithmische Lungensegmentierung einen extra Schritt, der die Trachea segmentiert und anschließend von der Segmentierung ausschließt. Ein solches Vorgehen könnte ebenso für das HMRF-UNet implementiert werden, indem die Trachea in der transversalen VIBE-Messung segmentiert und anschließend ebenso aus der koronalen Segmentierung der Lunge entfernt wird. Eine weitere Auffälligkeit waren die deutlich höheren Standardabweichungen für das HMRF-UNet Modell. Dies deutet darauf hin, dass die Segmentierungsgüte sehr unterschiedlich ist und es daher sowohl sehr gute als auch eher schlechtere Segmentierungen gab.

In einem nächsten Schritt sollte untersucht werden, ob es Unterschiede in der Segmentierungsgüte zwischen den verschiedenen MRT-Scannern gibt. Hierzu wurden die Segmentierungsmetriken für die algorithmische Lungensegmentierung jeweils nach Scanner gruppiert. In Abbildung 5.21 sind Boxplots für drei verschiedene Segmentierungsmetriken dargestellt. Die Boxen sind jeweils nach MRT-Scannern gruppiert. Es wurde jeweils nach Scanner-Typ, Gradientensystem und Softwareversion unterschieden. Man erkennt, dass die Segmentierung für alle Scanner-Modelle ähnlich gut war. Für die Metriken ohne Unschärfe (fuzzy), fielen die Unterschiede zwischen den Scanner-Modellen etwas deutlicher aus, diese wurden jedoch bei der Verwendung der fuzzy Metriken (siehe Abbildung A.1)

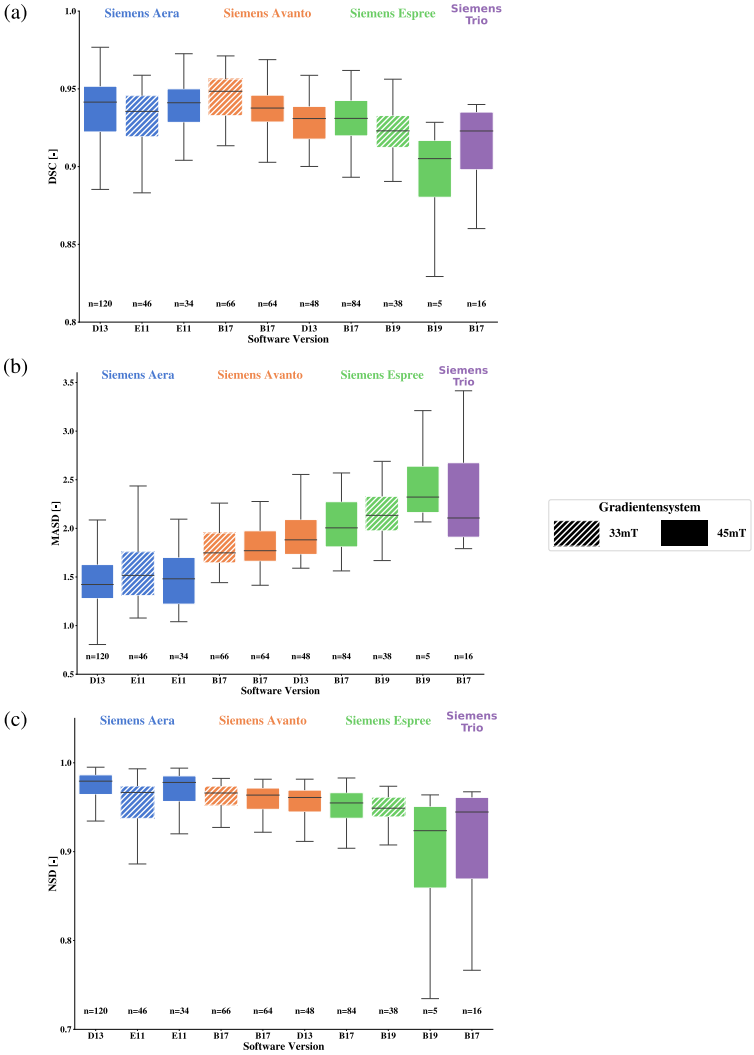


Abbildung 5.21: Boxplots verschiedener Segmentierungsmetriken aufgeteilt nach MRT-Scanner. (a) DSC (b) MASD und (c) NSD. Die schraffierten Boxen entsprechen den MRT-Scannern mit 33 mT Gradientensystem.

deutlich reduziert. Dieser Effekt ist insbesondere für die randbasierte MASD-Metrik sichtbar. Die ähnlicheren Ergebnisse bei den fuzzy Metriken könnten darauf hindeuten, dass die Segmentierungsfehler für die Scanner-Modelle Avanto, Espree und Trio möglicherweise auf eine minimale Über- oder Untersegmentierung zurückzuführen sind. Hinsichtlich der Softwareversion oder des verwendeten Gradientensystems konnten keine Einflüsse beobachtet werden. Die Boxplots verdeutlichen jedoch, dass die Verwendung von verschiedenen Metrikarten angebracht ist, da beispielsweise für die volumenbasierte DSC-Metrik die Unterschiede zwischen den Scanner-Modellen wesentlich kleiner ausfielen, als für die randbasierten Metriken MASD und NSD. Die Boxplots der in dieser Arbeit entwickelte EigenLung-Distanz Metrik sind in Abbildung 5.22 dargestellt. Man erkennt auch hier Unterschiede zwischen den einzelnen MRT-Scannern und dass die fuzzy Version der Metrik (FELD) zu deutlich besseren EigenLung-Distanzen führte. Im Vergleich zu den randbasierten und volumenbasierten Metriken zeigte die EigenLung-Distanz Metrik leicht unterschiedliche Ergebnisse, was die zusätzliche Aussagekraft der Metrik unterstreicht.

5.12.3 Vergleich verschiedener QDP-Varianten

Es gibt verschiedene Möglichkeiten QDP zu berechnen. Da die verschiedenen Clustering-Varianten bereits in Konietzke et al. [56] detailliert untersucht wurden und die Berechnung mit der Otsu Methode dort am besten abgeschnitten hatte, wurden nachfolgend nur Vergleiche verschiedener Varianten mit der Otsu Variante berücksichtigt. Neben der Cluster-Methode spielt die Wahl der zu erwartenden Klassen eine wesentliche Rolle. Hierbei wurden drei Varianten untersucht, die jeweils von zwei Klassen (normal perfundiert und defekt), drei Klassen (Gefäße, normal perfundiert und defekt) und vier Klassen (Gefäße, normal perfundiert, defekt und Hintergrund) ausgingen. Des Weiteren können zusätzliche Schritte zur Berechnung hinzugefügt werden. Eine Möglichkeit ist die Einführung eines Korrekturschrittes, der zur Berücksichtigung des Gravitationsgradienten im MRT den Schwellwert mit einem Gradienten entlang der dorsalen Richtung versieht. Zudem kann die dadurch generierte Defektkarte anschließend noch mit

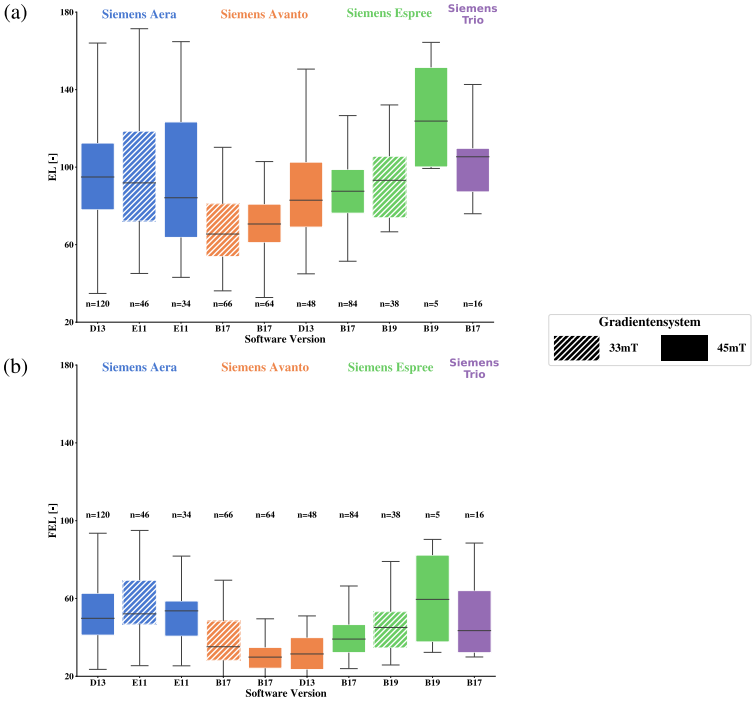


Abbildung 5.22: Boxplots der EigenLung-Distanz Metrik aufgeteilt nach MRT-Scanner. (a) Normale EigenLung-Distanz EL und (b) Fuzzy EigenLung-Distanz FEL. Die schraffierten Boxen entsprechen den MRT-Scannern mit 33 mT Gradientensystem.

einem (7×7)-Medianfilter gefiltert werden, um kleine rauschbasierte Artefakte zu korrigieren.

Für die Perfusionsquantifizierung gibt es keine direkte Referenz, die die wahren Perfusionswerte liefern könnte. Die beste Möglichkeit zur Bewertung einer Perfusionsquantifizierungsmethode ist die Berechnung der Korrelation der quantitativen Perfusionswerte (z. B. QDP) mit klinischen Parametern oder quantitativen

Tabelle 5.8: Spearman-Korrelationen zwischen verschiedenen QDP-Varianten und klinischen Parametern. Die Zahl hinter Otsu steht für die Anzahl an angenommenen Klassen. DLCO: Kohlenmonoxid-Diffusionskapazität, FEV1: Forciertes exspiratorisches Atemvolumen in 1 s, FVC: Forcierte Vitalkapazität. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

		Atemwegs Score	Funktioneller Score	Globaler Score	FEV1/FVC	DLCO
Otsu2	**Grav.**	$0{,}100^{*}$	$0{,}653^{***}$	$0{,}489^{***}$	$-0{,}419^{***}$	$-0{,}444^{***}$
	Grav.-7x7	$0{,}100^{*}$	$0{,}659^{***}$	$0{,}492^{***}$	$-0{,}421^{***}$	$-0{,}448^{***}$
	Normal	$0{,}107^{*}$	$0{,}665^{***}$	$0{,}501^{***}$	$-0{,}433^{***}$	$-0{,}462^{***}$
Otsu3	**Grav.**	$0{,}127^{**}$	$0{,}664^{***}$	$0{,}521^{***}$	$-0{,}462^{***}$	$-0{,}404^{***}$
	Grav.-7x7	$0{,}126^{**}$	$0{,}667^{***}$	$0{,}522^{***}$	$-0{,}465^{***}$	$-0{,}408^{***}$
	Normal	$0{,}130^{**}$	$0{,}671^{***}$	$0{,}527^{***}$	$-0{,}471^{***}$	$-0{,}417^{***}$
Otsu4	**Grav.**	0,088	$0{,}563^{***}$	$0{,}425^{***}$	$-0{,}412^{***}$	$-0{,}312^{***}$
	Grav.-7x7	0,088	$0{,}575^{***}$	$0{,}434^{***}$	$-0{,}419^{***}$	$-0{,}321^{***}$
	Normal	$0{,}093^{*}$	$0{,}586^{***}$	$0{,}445^{***}$	$-0{,}429^{***}$	$-0{,}334^{***}$

CT-Werten (QCT-Werten), die ebenso für die PatientInnen der COSYCONET-Kohorte bestimmt wurden. Daher wurden alle nachfolgenden Validierungen basierend auf den Korrelationen mit klinischen Parametern und QCT-Werten durchgeführt. In Tabelle 5.8 sind jeweils für neun verschiedene Berechnungsarten für QDP die Spearman-Korrelationen mit ausgewählten klinischen Parametern dargestellt. Alle neun Varianten basierten dabei auf einer Registrierung mit der Demon-Registrierung. Man erkennt, dass bei der Verwendung von vier Klassen zum Clustern (Otsu4) die schlechtesten Korrelationen gemessen wurden. Bei den Varianten mit zwei (Otsu2) bzw. drei (Otsu3) Klassen war kein durchgängiger Trend für höhere oder geringere Korrelationen erkennbar, wobei Otsu3 für alle klinischen Parameter außer für die Kohlenmonoxid-Diffusionskapazität (DLCO) etwas höhere Korrelationen besaß. Ebenso ist ersichtlich, dass die Verwendung der beiden Varianten zur Gravitationsgradienten-Korrektur immer zu leicht geringeren Korrelationen im Vergleich zu der Variante ohne Korrektur führten. Ebenso zeigte sich, dass sich für den funktionellen Score die stärksten Korrelationen ergaben.

Tabelle 5.9: Spearman-Korrelationen zwischen verschiedenen QDP-Varianten und QCT-Werten. Die Zahl hinter Otsu steht für die Anzahl an angenommenen Klassen. EI: Emphysem-Index. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

		EI	$\text{PRM}_{\text{normal}}$	PRM_{fSAD}	PRM_{emph}
Otsu2	**Grav.**	$0{,}606^{***}$	$-0{,}564^{***}$	$0{,}389^{***}$	$0{,}613^{***}$
	Grav.-7x7	$0{,}608^{***}$	$-0{,}570^{***}$	$0{,}400^{***}$	$0{,}616^{***}$
	Normal	$0{,}624^{***}$	$-0{,}591^{***}$	$0{,}420^{***}$	$0{,}632^{***}$
Otsu3	**Grav.**	$0{,}634^{***}$	$-0{,}584^{***}$	$0{,}388^{***}$	$0{,}642^{***}$
	Grav.-7x7	$0{,}639^{***}$	$-0{,}592^{***}$	$0{,}398^{***}$	$0{,}647^{***}$
	Normal	$0{,}651^{***}$	$-0{,}606^{***}$	$0{,}411^{***}$	$0{,}658^{***}$
Otsu4	**Grav.**	$0{,}562^{***}$	$-0{,}517^{***}$	$0{,}333^{***}$	$0{,}574^{***}$
	Grav.-7x7	$0{,}571^{***}$	$-0{,}529^{***}$	$0{,}345^{***}$	$0{,}583^{***}$
	Normal	$0{,}588^{***}$	$-0{,}547^{***}$	$0{,}359^{***}$	$0{,}600^{***}$

Dies war allerdings zu erwarten, da der funktionelle Score auf der Bewertung des DCE-MRT beruht und ein Maß für die Perfusionsdefekte darstellt.

Die Spearman-Korrelationen mit den quantitativen CT-Werten sind in Tabelle 5.9 abgebildet. Auch hier erkennt man, dass Otsu4 die geringsten Korrelationen mit den QCT-Parametern besaß. Otsu3 führte zu leicht höheren Korrelationen für alle Parameter außer PRM_{fSAD}. Die Korrekturen für den Gravitationseffekt führten genauso wie bei den klinischen Parametern zuvor zu niedrigeren Korrelationen.

5.12.4 Vergleich der Registrierungsmethoden

In einem nächsten Schritt sollte untersucht werden, welche der Registrierungsmethoden zu den besten Ergebnissen führt. Hierzu wurden basierend auf den Ergebnissen aus Abschnitt 5.12.3 jeweils nur noch die beiden Varianten Otsu2 und Otsu3 ohne Korrektur für den Gravitationseffekt betrachtet. In Tabelle 5.10 sind für fünf Registrierungsvarianten jeweils die Spearman-Korrelationen mit ausgewählten klinischen Parametern dargestellt. Man erkennt, dass alle fünf Regis-

Tabelle 5.10: Vergleich verschiedener Registrierungsvarianten durch Betrachtung der Spearman-Korrelationen mit klinischen Parametern. Die Zahl hinter Otsu steht für die Anzahl an angenommenen Klassen. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

		Atemwegs Score	Funktioneller Score	Globaler Score	FEV1/FVC	DLCO
Matlab	**Otsu2**	$0{,}107^{*}$	$0{,}665^{***}$	$0{,}501^{***}$	$-0{,}433^{***}$	$-0{,}462^{***}$
	Otsu3	$0{,}130^{**}$	$0{,}671^{***}$	$0{,}527^{***}$	$-0{,}471^{***}$	$-0{,}417^{***}$
Elastix	**Otsu2**	$0{,}103^{*}$	$0{,}712^{***}$	$0{,}528^{***}$	$-0{,}471^{***}$	$-0{,}432^{***}$
	Otsu3	$0{,}121^{**}$	$0{,}671^{***}$	$0{,}516^{***}$	$-0{,}470^{***}$	$-0{,}380^{***}$
ANTs	**Otsu2**	$0{,}122^{**}$	$0{,}696^{***}$	$0{,}530^{***}$	$-0{,}446^{***}$	$-0{,}429^{***}$
	Otsu3	$0{,}130^{**}$	$0{,}663^{***}$	$0{,}516^{***}$	$-0{,}458^{***}$	$-0{,}383^{***}$
ElastixActCont	**Otsu2**	$0{,}109^{*}$	$0{,}716^{***}$	$0{,}534^{***}$	$-0{,}457^{***}$	$-0{,}412^{***}$
	Otsu3	$0{,}122^{**}$	$0{,}682^{***}$	$0{,}524^{***}$	$-0{,}456^{***}$	$-0{,}369^{***}$
ANTsActCont	**Otsu2**	$0{,}106^{*}$	$0{,}709^{***}$	$0{,}526^{***}$	$-0{,}454^{***}$	$-0{,}408^{***}$
	Otsu3	$0{,}118^{*}$	$0{,}670^{***}$	$0{,}511^{***}$	$-0{,}452^{***}$	$-0{,}363^{***}$

trierungsvarianten zu sehr ähnlichen Korrelationen mit den klinischen Parametern führten. Für alle Registrierungsvarianten außer Matlab, wurden die höheren Korrelationen für die QDP-Berechnung mit Otsu2, also dem Clustering in zwei Klassen, erreicht. Beim Vergleich der Korrelationen zwischen den einzelnen klinischen Parametern zeigte sich erneut, dass der funktionelle Score am stärksten mit QDP korrelierte. Bezüglich der Verwendung der Nachbearbeitung mit dem aktiven Konturenmodell zeigte sich kein einheitlicher Trend, da die Korrelationen mit dem Nachbearbeitungsschritt für manche klinischen Parameter anstiegen, während die Korrelationen für andere klinische Parameter sanken. Insgesamt ließ sich anhand der Korrelation mit den klinischen Parametern keine Präferenz für eine Registrierungsvariante ableiten.

In Tabelle 5.11 sind jeweils die Spearman-Korrelationen mit ausgewählten QCT-Werten für die untersuchten Registrierungsvarianten dargestellt. Betrachtet man

Tabelle 5.11: Vergleich verschiedener Registrierungsvarianten durch Betrachtung der Spearman-Korrelationen mit QCT-Werten. Die Zahl hinter Otsu steht für die Anzahl an angenommenen Klassen. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

		EI	$\text{PRM}_{\text{normal}}$	PRM_{fSAD}	PRM_{emph}
Matlab	**Otsu2**	$0{,}624^{***}$	$-0{,}591^{***}$	$0{,}420^{***}$	$0{,}632^{***}$
	Otsu3	$0{,}651^{***}$	$-0{,}606^{***}$	$0{,}411^{***}$	$0{,}658^{***}$
Elastix	**Otsu2**	$0{,}619^{***}$	$-0{,}584^{***}$	$0{,}396^{***}$	$0{,}629^{***}$
	Otsu3	$0{,}606^{***}$	$-0{,}572^{***}$	$0{,}384^{***}$	$0{,}615^{***}$
ANTs	**Otsu2**	$0{,}606^{***}$	$-0{,}568^{***}$	$0{,}389^{***}$	$0{,}612^{***}$
	Otsu3	$0{,}607^{***}$	$-0{,}559^{***}$	$0{,}371^{***}$	$0{,}614^{***}$
ElastixActCont	**Otsu2**	$0{,}601^{***}$	$-0{,}572^{***}$	$0{,}393^{***}$	$0{,}611^{***}$
	Otsu3	$0{,}593^{***}$	$-0{,}561^{***}$	$0{,}377^{***}$	$0{,}601^{***}$
ANTsActCont	**Otsu2**	$0{,}598^{***}$	$-0{,}569^{***}$	$0{,}390^{***}$	$0{,}608^{***}$
	Otsu3	$0{,}584^{***}$	$-0{,}553^{***}$	$0{,}372^{***}$	$0{,}594^{***}$

die Korrelationen mit den QCT-Werten, so schien die Variante mit Matlab-Registrierung und Otsu3 zu den stärksten Korrelationen über alle QCT-Werte hinweg zu führen. Außerdem zeigte sich, dass die Nachbearbeitung mit dem aktiven Konturenmodell bei den QCT-Werten überall zu niedrigeren Korrelationen führte. Die beobachteten höchsten Korrelationen für die Matlab-Registrierung wirken in Anbetracht des visuellen Vergleichs der registrierten Lungenmasken in Abbildung 5.4 erstaunlich, da die Matlab-Registrierung hier teilweise zu schlechteren Lungenmasken geführt hat. Sowohl die *elastix*-Registrierung, als auch die *ANTs*-Registrierung mit der aktiven Konturenkorrektur erzeugten hier bessere Lungenmasken. Jedoch zeigte die Abbildung nur ein paar zufällig ausgewählte PatientInnen. Die hier berechneten Korrelationen beziehen sich allerdings auf die komplette Kohorte und liefern daher eine Bewertung der Registrierungsvarianten für alle untersuchten Personen. Daher wurde diesen Ergebnissen eine höhere Bedeutung beigemessen, als den zufällig ausgewählten Registrierungen in Abbildung 5.4, die möglicherweise ein verzerrtes Bild darstellen. Unter Anbetracht der Ergebnisse des Registrierungsvergleiches beziehen sich alle nachfolgenden

Ergebnisse auf das Modell mit Matlab-Registrierung und der QDP-Berechnung mit Otsu3.

5.12.5 Korrelation mit klinischen Parametern & QCT

Die zuvor betrachteten klinischen und QCT Parameter sind nur ein kleiner Teil der verfügbaren Werte. Nachdem basierend auf den vorherigen Untersuchungen ein Vorgehen zur Perfusionsquantifizierung ausgewählt wurde, sollten in einem nächsten Schritt die Korrelationen mit allen verfügbaren klinischen Parametern sowie den relevantesten QCT-Werten berechnet werden. In Tabelle 5.12 sind jeweils die Spearman-Korrelationen für PBF und QDP dargestellt. Für PBF konnten fast ausschließlich keine oder geringe Korrelationen mit den klinischen Parametern beobachtet werden. Nur mit DLCO und dem funktionellen Krankheits-Score konnte eine moderate Korrelation nachgewiesen werden. Die Korrelationen von QDP mit den klinischen Parametern fielen dagegen grundsätzlich stärker aus als für PBF. Hier konnten starke Korrelationen für den funktionellen Score und den globalen Score beobachtet werden. Zusätzlich ergaben sich für ppFEV1, DLCO, GOLD und ppFEV1/ppFVC jeweils moderate Korrelationen.

Für die QCT-Werte konnten sowohl für PBF als auch für QDP im Mittel höhere Korrelationen als mit den klinischen Parametern beobachtet werden. Hier ergaben sich für PBF moderate Korrelationen mit EI, PRM_{normal}, PRM_{emph}, Perc15, MLD und Gefäßvolumen/Lungenvolumen, wobei keine der Korrelationen so stark wie die Korrelation mit dem funktionellen Score ausfiel. Im Gegensatz dazu lieferte QDP starke Korrelationen mit EI, PRM_{normal}, PRM_{emph}, Perc15 und MLD sowie moderate Korrelationen mit PRM_{fSAD}, EI-MLA und dem Gefäßvolumen/-Lungenvolumen. Die deutlich stärkeren Korrelationen von QDP gegenüber PBF für diverse klinische und QCT-Werte bestätigen erneut die Überlegenheit von QDP gegenüber PBF zur Bewertung des Gesundheitszustands einer untersuchten Person.

Tabelle 5.12: Spearman-Korrelationen von PBF und QDP mit allen klinischen Parametern sowie QCT-Werten. AWPi10: mittlere Wanddicke eines Atemweglumens von 10 mm, BE-Index: Index der Verstärkung der Bronchialwände, Perc15: 15. Perzentil des Lungen HU-Histogramms, MLD: mittlere Lungendichte, MLA: Mittlere Lungenabschwächung. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	PBF	**QDP**
6-Minuten Gehdistanz	0,029	−0,089
Timed Up and Go	0,011	0,019
ppFEV1	$0{,}181^{***}$	$-0{,}421^{***}$
ppFVC	0,049	$-0{,}118^{*}$
DLCO Prozentsatz	$0{,}301^{***}$	$-0{,}417^{***}$
GOLD	$-0{,}233^{***}$	$0{,}456^{***}$
Bode	$-0{,}175^{***}$	$0{,}349^{***}$
Atemwegs Score	0,008	$0{,}130^{**}$
Funktioneller Score	$-0{,}473^{***}$	$0{,}671^{***}$
Globaler Score	$-0{,}296^{***}$	$0{,}527^{***}$
ppFEV1/ppFVC	$0{,}212^{***}$	$-0{,}471^{***}$
Wand-Prozentzahl	$0{,}179^{***}$	$-0{,}257^{***}$
AWPi10	$0{,}131^{**}$	$-0{,}173^{***}$
BE-Index-Lunge	−0,074	$0{,}211^{***}$
EI	$-0{,}348^{***}$	$0{,}651^{***}$
PRM_{normal}	$0{,}372^{***}$	$-0{,}606^{***}$
PRM_{fSAD}	$-0{,}285^{***}$	$0{,}411^{***}$
PRM_{emph}	$-0{,}363^{***}$	$0{,}658^{***}$
Perc15	$0{,}353^{***}$	$-0{,}646^{***}$
MLD	$0{,}396^{***}$	$-0{,}564^{***}$
EI-MLA	$-0{,}248^{***}$	$0{,}426^{***}$
Gefäßvolumen/Lungenvolumen	$0{,}357^{***}$	$-0{,}368^{***}$

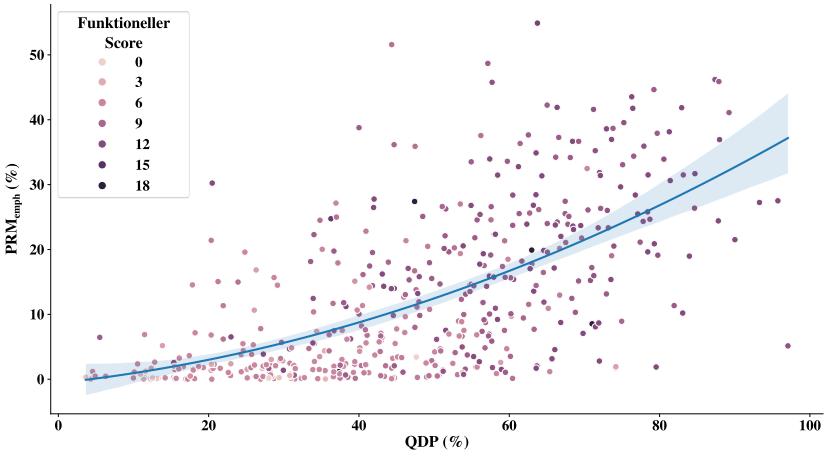


Abbildung 5.23: Korrelationsplot für die Korrelation zwischen QDP und PRM_{emph} aufgeteilt nach funktionellen Krankheitsscore.

Abschließend wurde noch untersucht, ob der Krankheitszustand einen Einfluss auf die Korrelation besitzt. Hierzu wurden die QDP-Werte gegenüber PRM_{emph} geplottet und die Punkte anhand des funktionellen Scores der PatientInnen eingefärbt. Die Ergebnisse sind in Abbildung 5.23 dargestellt. Man erkennt, dass es einen nicht-linearen Zusammenhang zwischen QDP und PRM_{emph} zu geben schien. Des Weiteren wird deutlich, dass die relativ gesunden PatientInnen mit niedrigem funktionellen Score nicht gut korrelierten. Eine mögliche Erklärung hierfür könnte sein, dass diese PatientInnen laut Definition keine Perfusionsdefekte aufwiesen. Bei der QDP-Berechnung wird jedoch zwangsläufig angenommen, dass es einen Cluster von normalen und defekten Perfusionsbereichen gibt, wodurch fehlerhafte QDP-Werte entstehen könnten. In diesem Fall wäre es sinnvoll, Referenzwerte zu verwenden, um zu bewerten, ob alle Signale in der Lunge normalen Werten entsprechen. Dies ist jedoch im MRT, im Gegensatz zum CT nicht möglich. Daher wäre es empfehlenswert, eine alternative Methode zu ermitteln, um im Falle einer komplett „physiologischen“ Perfusion keine künstlichen Perfusionsdefekte zu generieren.

Tabelle 5.13: Quantitative Perfusionswerte für die gesamte COSYCONET-Kohorte sowie aufgeteilt nach GOLD-Klassen.

	n	PBF (ml/100ml/min)	PBV (ml/100ml)	MTT (s)	QDP (%)
Gesamt	503	83,3 ± 65,7	6,6 ± 5,6	4,7 ± 1,0	47,8 ± 19,7
GOLD -1	27	105,7 ± 91,9	9,3 ± 9,1	5,1 ± 1,1	30,5 ± 15,4
GOLD 0	48	111,2 ± 74,2	9,1 ± 6,9	4,8 ± 1,0	31,5 ± 16,1
GOLD 1	49	76,6 ± 42,1	6,5 ± 4,1	4,9 ± 0,9	38,8 ± 18,0
GOLD 2	203	82,9 ± 58,0	6,4 ± 4,5	4,7 ± 1,0	47,9 ± 18,6
GOLD 3	144	78,6 ± 74,0	6,1 ± 6,2	4,6 ± 0,8	55,6 ± 16,2
GOLD 4	32	56,1 ± 35,1	4,5 ± 3,4	4,6 ± 0,9	65,1 ± 17,1

5.12.6 Quantitative Perfusionswerte

Die quantitativen Perfusionswerte wurden für alle PatientInnen der Kohorte berechnet. In Abbildung 5.24 sind für zwei Personen der Kohorte die quantitativen Perfusionskarten dargestellt. Die linke Spalte zeigt die quantitativen Perfusionskarten für eine 68-jährige weibliche COPD-Patientin, die der GOLD-Klasse GOLD-0 angehört. Die rechte Spalte stellt die quantitativen Perfusionskarten für einen 57-jährigen männlichen COPD-Patienten der GOLD-4 Klasse dar. Man erkennt sehr gut, wie sich die unterschiedlichen Krankheitsstadien in den quantitativen Perfusionskarten widerspiegelten. Während die linken Karten alle sehr homogen waren, besaß der rechte Patient deutliche Perfusionsunterschiede zwischen der oberen und unteren Lungenhälfte. Allgemein waren ca. 67 % der Lunge des rechten Patienten defekt, während die linke Patientin mit ca. 28 % deutlich weniger defekte Lungenbereiche besaß.

Die Einteilung von COPD-PatientInnen nach GOLD-Klassen ist eine üblich verwendete Gruppierung. Daher wurden die berechneten quantitativen Perfusionswerte in Tabelle 5.13 jeweils für die Gesamtkohorte und aufgeteilt nach GOLD-Klassen dargestellt. Entgegen der Erwartung fällt auf, dass die PBF-Werte

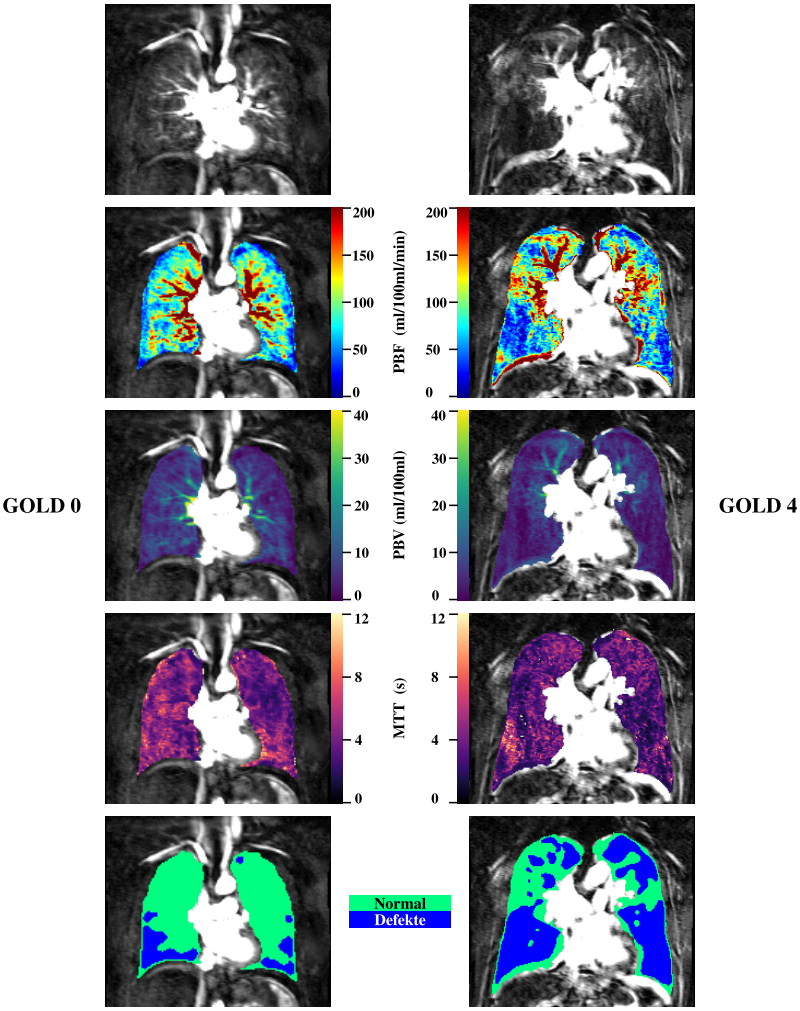


Abbildung 5.24: Quantitative Perfusionskarten für zwei verschiedene COPD-PatientInnen. Links 68-jährige weibliche COPD-Patientin mit GOLD-0 Klasse. Rechts 57-jähriger männlicher COPD-Patient mit GOLD-4 Klasse.

keinen klaren absteigenden Trend mit steigender GOLD-Klasse zeigten. Das pulmonale Blutvolumen PBV zeigte dagegen einen klar erkennbaren absteigenden Trend. Die mittlere Transitzeit schien sich zwischen den GOLD-Klassen 0–4 nicht nennenswert zu verändern. Die Tabelle zeigt erneut, dass QDP ein sehr vielversprechender Parameter ist, um unterschiedliche Erkrankungsstadien nachzuweisen. Die Werte zwischen den Goldstadien 0–4 unterschieden sich hier sehr deutlich und QDP nahm pro Goldstadium ca. 7–10 % zu. Einzig die Unterschiede zwischen den PatientInnen mit Risiko für COPD (GOLD -1) und Goldstadium 0 fielen sehr gering aus.

5.12.7 Einfluss des MRT-Scanners auf die Quantifizierung

Für die Bewertung der Robustheit der Quantifizierungsmethode ist die Betrachtung des Einflusses der verschiedenen MRT-Scanner von großer Bedeutung. In den Tabellen 5.14 und 5.15 sind die Korrelationen mit den klinischen Parametern bzw. den QCT-Werten jeweils aufgeteilt nach Scanner-Modell dargestellt. Die Korrelationen fielen je nach Scanner-Modell unterschiedlich stark aus. Für die klinischen Parameter waren die Korrelationen der 33 mT Scanner des Typs Avanto und Espree schwächer als die jeweiligen 45 mT Varianten der beiden Modelle. Beim Aera lieferte der Scanner mit dem 33 mT-Gradientensystem dagegen die stärksten Korrelationen. Auch für die QCT-Werte besaß der Aera mit 33 mT-Gradientensystem über alle Scanner hinweg die stärksten Korrelationen mit allen QCT-Parametern. Auch der Avanto mit 33 mT-Gradientensystem besaß hier stärkere Korrelationen als die gleichen Modelle mit stärkerem Gradientensystem.

Bei der Betrachtung der Boxplots der einzelnen Scanner für verschiedene Perfusionswerte in Abbildung 5.25 konnten sehr ähnliche Verteilungen beobachtet werden. Insbesondere bei QDP überlappten fast alle Boxen, ausgenommen von der Box des Espree-B19-45mT, die jedoch auch nur die Werte von sechs PatientInnen repräsentierte. Dies zeigt erneut die Robustheit der QDP-Metrik. Jedoch konnten auch für die anderen quantitativen Perfusionsparameter sehr ähnliche Boxplots beobachtet werden. Insbesondere die Aeras und die Avantos zeigten eine sehr

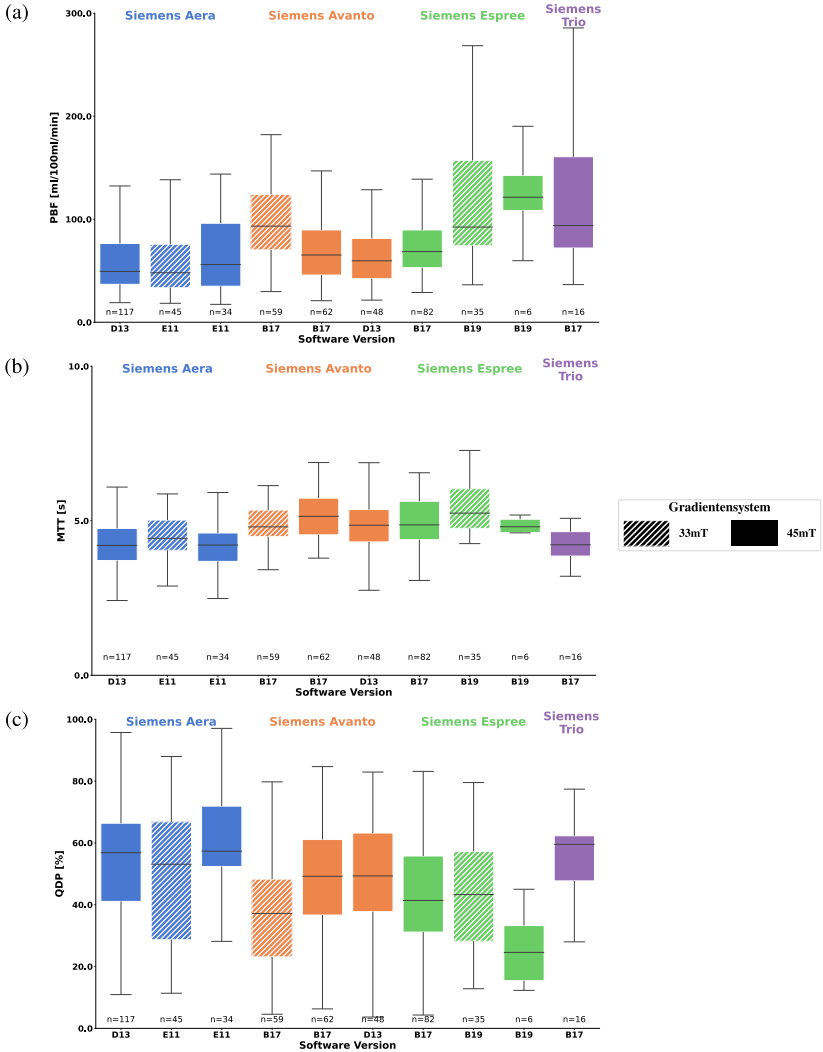


Abbildung 5.25: Boxplots verschiedener quantitativer Perfusionswerte aufgeteilt nach MRT-Scanner. (a) PBF (b) MTT und (c) QDP. Die schraffierten Boxen entsprechen den MRT-Scannern mit 33 mT Gradientensystem.

Tabelle 5.14: Spearman-Korrelationen zwischen QDP und klinischen Parametern für die Gesamtkohorte und aufgeteilt nach MRT-Scanner. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

		n	Atemwegs Score	Funktioneller Score	Globaler Score	FEV1/FVC	DLCO
Gesamt		455	$0{,}130^{**}$	$0{,}671^{***}$	$0{,}527^{***}$	$-0{,}471^{***}$	$-0{,}417^{***}$
Siemens Aera	**45mT D13**	112	0,183	$0{,}677^{***}$	$0{,}628^{***}$	$-0{,}494^{***}$	$-0{,}471^{***}$
	33mT E11	39	$0{,}368^{*}$	$0{,}788^{***}$	$0{,}718^{***}$	$-0{,}593^{***}$	$-0{,}563^{***}$
	45mT E11	32	0,078	$0{,}597^{***}$	$0{,}459^{**}$	$-0{,}367^{*}$	−0,070
Siemens Avanto	**33mT B17**	56	0,219	$0{,}549^{***}$	$0{,}521^{***}$	$-0{,}509^{***}$	$-0{,}305^{*}$
	45mT B17	57	$0{,}270^{*}$	$0{,}744^{***}$	$0{,}610^{***}$	$-0{,}541^{***}$	$-0{,}315^{*}$
	45mT D13	40	−0,152	$0{,}590^{***}$	0,261	$-0{,}533^{***}$	$-0{,}445^{**}$
Siemens Espree	**45mT B17**	70	0,134	$0{,}743^{***}$	$0{,}538^{***}$	$-0{,}613^{***}$	$-0{,}338^{**}$
	33mT B19	28	0,038	$0{,}411^{*}$	0,242	$-0{,}471^{*}$	$-0{,}488^{**}$
	45mT B19	5	−0,224	0,783	$0{,}975^{**}$	$-0{,}900^{*}$	$-0{,}900^{*}$
Siemens Trio	**45mT B17**	16	0,087	0,337	0,286	$-0{,}574^{*}$	−0,165

ähnliche statistische Verteilung. Ebenso war ersichtlich, dass die Avantos und Aeras sehr ähnliche Ergebnisse lieferten, es jedoch scheinbar einen Unterschied dieser beiden Scanner zu den Esprees und dem Trio gab.

Ein wesentlicher Schwachpunkt der reinen Betrachtung der Medianwerte einer untersuchten Person ist jedoch, dass dies eine starke Verallgemeinerung der gemessenen Werte erzeugt. Der Vorteil, dass für jeden Voxel quantitative Werte vorliegen, wird hiermit obsolet. Eine Betrachtung der Histogramme der quantitativen Werte bietet hierfür eine robustere Vergleichsvariante. In Abbildung 5.26 sind die mittleren Histogramme für jeden Scanner dargestellt, die jeweils alle Histogramme von PatientInnen, die mit diesem Scanner untersucht wurden, mitteln. Anhand der Histogramme ließen sich vier wesentliche Gruppen ausmachen. Die erste Gruppe bestand aus dem Espree-B19-45mT, der sich vermutlich rein aufgrund der geringen PatientInnenanzahl von den anderen Scannern abgrenzte. Die zweite Gruppe

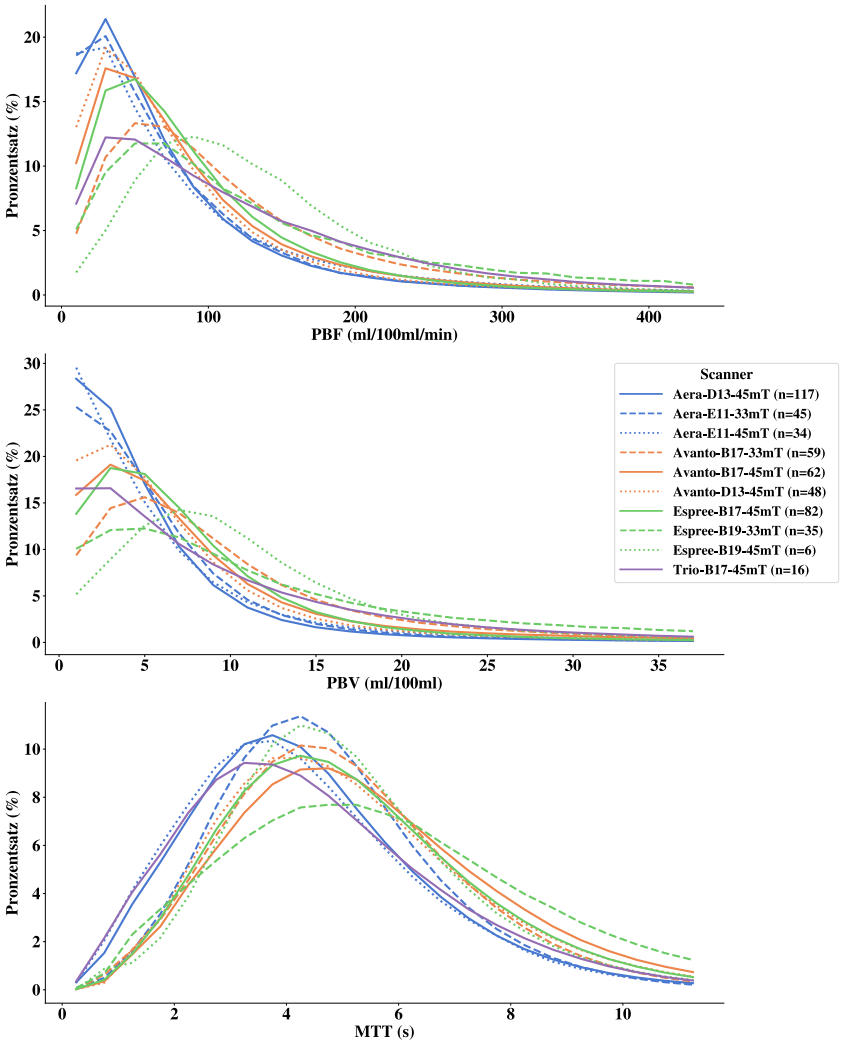


Abbildung 5.26: Mittlere Histogramme verschiedener quantitativer Perfusionswerte aufgeteilt nach MRT-Scanner.

Tabelle 5.15: Spearman-Korrelationen zwischen QDP und QCT-Werten für die Gesamtkohorte und aufgeteilt nach MRT-Scanner. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

		n	EI	PRM_{normal}	PRM_{fSAD}	PRM_{emph}
Gesamt		471	$0{,}651^{***}$	$-0{,}606^{***}$	$0{,}411^{***}$	$0{,}658^{***}$
Siemens Aera	**45mT D13**	110	$0{,}669^{***}$	$-0{,}599^{***}$	$0{,}327^{***}$	$0{,}661^{***}$
	33mT E11	44	$0{,}828^{***}$	$-0{,}748^{***}$	$0{,}522^{***}$	$0{,}820^{***}$
	45mT E11	33	$0{,}485^{**}$	$-0{,}486^{**}$	0,216	$0{,}535^{**}$
Siemens Avanto	**33mT B17**	57	$0{,}767^{***}$	$-0{,}687^{***}$	$0{,}486^{***}$	$0{,}752^{***}$
	45mT B17	54	$0{,}578^{***}$	$-0{,}477^{***}$	$0{,}294^{*}$	$0{,}589^{***}$
	45mT D13	39	$0{,}668^{***}$	$-0{,}613^{***}$	$0{,}477^{**}$	$0{,}679^{***}$
Siemens Espree	**45mT B17**	82	$0{,}715^{***}$	$-0{,}578^{***}$	$0{,}245^{*}$	$0{,}700^{***}$
	33mT B19	34	$0{,}495^{**}$	$-0{,}355^{*}$	0,170	$0{,}414^{*}$
	45mT B19	5	$1{,}000^{***}$	$-1{,}000^{***}$	$1{,}000^{***}$	$1{,}000^{***}$
Siemens Trio	**45mT B17**	13	$0{,}786^{**}$	$-0{,}599^{*}$	0,132	$0{,}791^{**}$

bildeten der Avanto und der Espree mit 33 mT-Gradientensystem. Der Avanto, Espree und Trio mit 45 mT-Gradientensystem und Softwareversion B17 bildeten die dritte Gruppe. Eine vierte Gruppe bildeten die drei Aeras und der Avanto mit Softwareversion D13. Diese Beobachtungen lassen darauf schließen, dass die Softwareversion im Wesentlichen für die unterschiedlichen Quantifizierungsergebnisse verantwortlich war, da der Avanto mit Softwareversion D13 zusammen mit den Aeras in eine Gruppe fiel. Zusätzlich schien auch das Gradientensystem einen Einfluss auf die Quantifizierung zu haben. Der Einfluss des Gradientensystems wurde bereits in der begleitenden Phantom-Studie gezeigt, in der die gemessenen Durchmesser eines Stabes für die Scanner mit 33 mT-Gradientensystem deutlich über den Werten der anderen Scanner lagen [176].

Tabelle 5.16: Vergleich der MRT-Qualitätsscores für Atmung zwischen den einzelnen Scannern. Alle Angaben in Mittelwert ± Standardabweichung. QS_{ga}: Qualitätsscore für Gesamtatmung, QS_{ba}: Qualitätsscore für Atmung während Boluspassage.

		n	QS_{ga}	QS_{ba}
Gesamt		504	0,68 ± 0,32	0,71 ± 0,32
Siemens Aera	**45mT D13**	117	0,77 ± 0,29	0,80 ± 0,28
	33mT E11	45	0,85 ± 0,25	0,88 ± 0,24
	45mT E11	34	0,61 ± 0,34	0,65 ± 0,37
Siemens Avanto	**33mT B17**	59	0,65 ± 0,32	0,68 ± 0,31
	45mT B17	62	0,71 ± 0,34	0,74 ± 0,34
	45mT D13	48	0,73 ± 0,27	0,75 ± 0,28
Siemens Espree	**45mT B17**	82	0,64 ± 0,29	0,67 ± 0,30
	33mT B19	35	0,48 ± 0,27	0,53 ± 0,31
	45mT B19	6	0,66 ± 0,27	0,69 ± 0,26
Siemens Trio	**45mT B17**	16	0,29 ± 0,31	0,32 ± 0,34

5.12.8 MRT-Qualitätsscore

Wie zuvor erwähnt liefert die Perfusionsquantifizierungs-Pipeline im Normalfall immer quantitative Ergebnisse. Diese können jedoch durch eine schlechte Qualität der DCE-MRT Messung verfälscht bzw. ungenau sein. Hierfür wurde, wie in Abschnitt 5.8 beschrieben, der MRT-Qualitätsscore entwickelt. In Tabelle 5.16 sind die mittleren Qualitätsscores der Atmung für die gesamte Kohorte und jeweils für jedes Scanner-Modell dargestellt. Zur Erinnerung, ein Qualitätsscore von $QS_* = 1$ ist optimal und die prozentuale Anzahl an veratmeten Zeitpunkten ergibt sich, wenn man den Qualitätsscore von eins abzieht. Für alle Scanner und die gesamte Kohorte waren die Qualitätsscores für die Atmung während der Boluspassage höher als für die gesamte Messung. Dies ist ein gutes Zeichen, da dies bedeutet, dass vor oder nach der Boluspassage mehr Atmung stattfand, welche durch das in Abschnitt 5.6 vorgestellte Zuschneiden der DCE-MRT Messung

korrigiert werden kann. Jedoch zeigte sich auch, dass für fast alle Scanner zwischen 20–30 % der Boluspassage veratmet waren. Diese Atmung während der Boluspassage führt zwangsläufig zu fehlerhaften Quantifizierungen. Dies zeigt erneut, dass es essenziell ist, eine Methode zu entwickeln, die die Atmung während der Messung korrigieren kann. Besonders viel Atmung wurde in den Messungen des Trio-Scanners festgestellt, aber auch der Espree mit 33 mT-Gradientensystem besaß im Mittel fast 50 % veratmete Zeitpunkte. Des Weiteren liefert die Tabelle eine mögliche Erklärung für die niedrigeren beobachteten Korrelationen des Aera 45mT-E11 im Vergleich zu den anderen beiden Aera-Scannern. Der mittlere Qualitätsscore für die Atmung während der Boluspassage war hier deutlich geringer (0,65 vs. 0,88 bzw. 0,80), was einer deutlich höheren Anzahl an veratmeten Zeitpunkten während der Messung entspricht. Daher könnten die niedrigeren Korrelationen schwerpunktmäßig durch die erhöhte Atmung entstanden sein. Ein ähnlicher Effekt war auch für den Espree-Scanner mit 33 mT-Gradientensystem zu erkennen.

Im nächsten Schritt wurden die berechneten Qualitätsscores für die Boluspassage betrachtet. Diese sind in Tabelle 5.17 dargestellt. Die Qualitätsscores für die Boluspassage zeigten, dass eine fehlende Bolusinjektion so gut wie nicht auftrat. Nur für den Espree mit 33 mT-Gradientensystem gab es überhaupt einen Fall einer fehlenden Bolusinjektion. Eine zu späte Bolusgabe oder eine fehlende bzw. unzureichende Chaser-Injektion schien auch eher selten ein Problem gewesen zu sein, da die Qualitätsscores aller Scanner oberhalb von 0,92 lagen. Das einzige potenzielle Problem schien eine zu frühe Gabe des Kontrastmittels zu sein. Jedoch waren auch hier alle Qualitätsscores oberhalb von 0,5. Möglicherweise spielte hier auch das Zuschneiden der Messung eine Rolle, da dies dazu führen kann, dass direkt zwei Zeitschritte vor der Kontrastmittelankunft abgeschnitten werden und daher die Zeit zwischen angepasstem Messstart und Bolusankunft sehr gering ausfällt.

Für die Untersuchung der Eignung des DCE-MRT Qualitätsscores zum Auffinden von Anomalien wurden jeweils die Histogramme der Median PBF-Werte verschiedener Zentren für ein Scanner-Modell untersucht. Für alle Ausreißer in

Tabelle 5.17: Vergleich der MRT-Qualitätsscores für die Bolusinjektion zwischen den einzelnen Scannern. Alle Angaben in Mittelwert ± Standardabweichung. QS_{nb}: Qualitätsscore für fehlende Bolusinjektion, QS_{eb}: Qualitätsscore für zu frühe Bolusinjektion, QS_{lb}: Qualitätsscore für zu späte Bolusinjektion, QS_{nc}: Qualitätsscore für fehlende Chaser-Injektion.

		n	QS_{nb}	QS_{eb}	QS_{lb}	QS_{nc}
Gesamt		504	0,99 ± 0,60	0,67 ± 0,15	0,99 ± 0,70	0,96 ± 0,10
Siemens Aera	**45mT D13**	117	1,00 ± 0,00	0,71 ± 0,13	0,99 ± 0,50	0,98 ± 0,80
	33mT E11	45	1,00 ± 0,00	0,69 ± 0,80	1,00 ± 0,00	0,99 ± 0,40
	45mT E11	34	1,00 ± 0,20	0,72 ± 0,16	0,97 ± 0,10	0,94 ± 0,16
Siemens Avanto	**33mT B17**	59	1,00 ± 0,00	0,79 ± 0,80	1,00 ± 0,00	0,98 ± 0,40
	45mT B17	62	1,00 ± 0,00	0,55 ± 0,16	1,00 ± 0,20	0,96 ± 0,70
	45mT D13	48	1,00 ± 0,00	0,67 ± 0,17	1,00 ± 0,20	0,97 ± 0,90
Siemens Espree	**45mT B17**	82	1,00 ± 0,00	0,64 ± 0,14	0,99 ± 0,11	0,93 ± 0,15
	33mT B19	35	0,92 ± 0,21	0,60 ± 0,18	0,97 ± 0,17	0,92 ± 0,16
	45mT B19	6	1,00 ± 0,00	0,66 ± 0,50	1,00 ± 0,00	0,99 ± 0,20
Siemens Trio	**45mT B17**	16	1,00 ± 0,00	0,66 ± 0,17	1,00 ± 0,00	0,99 ± 0,30

den einzelnen Histogrammen wurden jeweils die Qualitätsscores betrachtet. Abbildung 5.27 zeigt die Histogramme der einzelnen Zentren sowie die berechneten Qualitätsscores für die einzelnen Ausreißer. Man erkennt, dass alle Ausreißer mindestens einen niedrigen Qualitätsscore besaßen. Die zwei auftretenden Probleme waren eine sehr große Anzahl an veratmeten Zeitpunkten während der Messung oder aber eine zu frühe Bolusankunft, wodurch die Subtraktionsbilder nicht korrekt berechnet werden konnten.

Um zu überprüfen, ob das Herausfiltern von veratmeten Messungen zu einer Verbesserung der Quantifizierung führt, wurden die Korrelationen aufgeteilt nach Scanner nur für die PatientInnen berechnet, die während der Messung nicht atmeten. Die Ergebnisse sind in Tabelle 5.18 dargestellt und zeigten zunächst, dass sich hierdurch die Anzahl der PatientInnen durch das Filtern auf fast ein Drittel

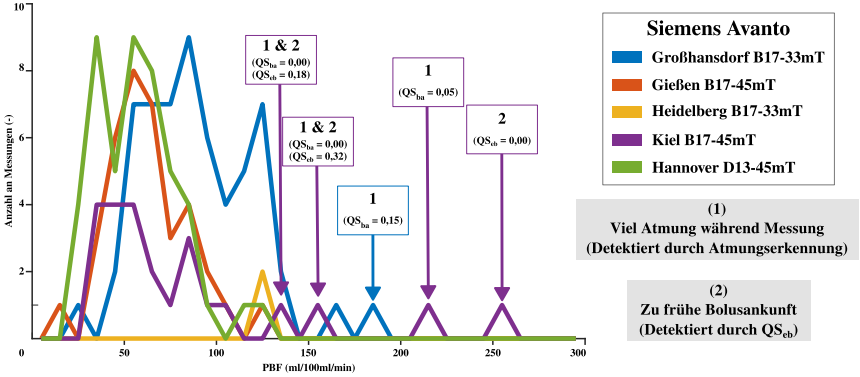


Abbildung 5.27: Histogramme der PBF-Mediane für alle PatientInnen eines Zentrums für den Siemens Avanto-Scanner. Für die jeweiligen Ausreißer wurden jeweils die Qualitätsscores berechnet und dargestellt.

reduziert hat. Die Korrelationen mit den QCT-Werten sind hierbei für alle Scanner, die noch eine statistisch signifikante RestpatientInnenanzahl besaßen, deutlich angestiegen. Insbesondere für den Aera 45mT-D13 stiegen die Korrelationen deutlich an, z. B. von 0,661 auf 0,756 für die Korrelation mit PRM_{emph}. Auch die Gesamtkorrelation für die komplette COSYCONET-Kohorte stiegen für alle QCT-Parameter durch das Entfernen der PatientInnen mit Atmung deutlich an. Die Bedeutung einer Methode zur Korrektur von veratmeten Zeitpunkten wird hierdurch erneut untermauert. Dies zeigt, dass die Qualitätsscores sehr gut geeignet sind, um automatisch fehlerhafte Quantifizierungen herauszufiltern. Die Histogramme zeigten ebenso, dass es bei den Avanto-Scannern zwei verschiedene Gruppen gab, die sich jeweils in eine Gruppe mit schwachem und eine Gruppe mit starkem Gradientensystem unterteilten.

5.12.9 Einordnung und Zusammenfassung der Ergebnisse

Die Ergebnisse dieses Unterkapitels haben das Potenzial der Pipeline zur Perfusionsquantifizierung aufgezeigt. Die Robustheit der Pipeline wurde durch die

Tabelle 5.18: Spearman-Korrelationen von QDP mit QCT-Werten ohne PatientInnen, die während der Messung atmeten.

		n	EI	PRM_{normal}	PRM_{fSAD}	PRM_{emph}
Gesamt		155	$0{,}705^{***}$	$-0{,}667^{***}$	$0{,}499^{***}$	$0{,}718^{***}$
Siemens Aera	**45mT D13**	47	$0{,}756^{***}$	$-0{,}680^{***}$	$0{,}452^{**}$	$0{,}756^{***}$
	33mT E11	28	$0{,}846^{***}$	$-0{,}705^{***}$	0,354	$0{,}839^{***}$
	45mT E11	9	0,117	−0,283	0,400	0,200
Siemens Avanto	**33mT B17**	15	$0{,}589^{*}$	$-0{,}564^{*}$	0,457	$0{,}596^{*}$
	45mT B17	19	$0{,}575^{**}$	$-0{,}656^{**}$	$0{,}570^{*}$	$0{,}570^{*}$
	45mT D13	13	$0{,}841^{***}$	$-0{,}852^{***}$	$0{,}841^{***}$	$0{,}885^{***}$
Siemens Espree	**45mT B17**	21	$0{,}800^{***}$	$-0{,}735^{***}$	$0{,}560^{**}$	$0{,}812^{***}$
	33mT B19	1	-	-	-	-
	45mT B19	2	1,000	−1,000	1,000	1,000

Validierung einzelner Methodenbestandteile untersucht. In Abbildung 5.28 ist die finale Version der entwickelten vollautomatischen und robusten Pipeline zur Perfusionsquantifizierung visualisiert. Die Qualiät der Segmentierung, die eine wichtige Rolle für die Quantifizierung spielt, zeigte eine sehr hohe Güte und erzeugte mindestens so gute Segmentierungen wie andere Lungensegmentierungsalgorithmen in der Literatur. Jedoch hat die Untersuchung der Registrierung gezeigt, dass die gute Segmentierung durch eine schlechte Registrierung teilweise wieder ruiniert wurde. Um dieses Problem zu beheben, könnte sowohl an der Registrierung selbst, als auch an einer direkten Segmentierung der Lungen im DCE-MRT gearbeitet werden. Aktuell erscheint eine direkte Segmentierung der Lunge im DCE-MRT jedoch schwierig, da die Qualität und das SNR des DCE-MRT relativ schlecht ist. Mögliche alternative Techniken wären die erwähnte ML-Segmentierung des DCE-MRT mittels der PBF-Karte aus Winther et al. [108], die jedoch Probleme bei der Segmentierung von Lungenbereichen mit Perfusionsdefekten besitzen könnte. Eine weitere Möglichkeit wäre eine Verbesserung der Bildqualität der DCE-MRT Messung, da dies die Verwendung der allgemeinen

algorithmischen Lungensegmentierung ermöglichen würde. Hierzu könnte beispielsweise die Super-Resolution-Rekonstruktion zum Einsatz kommen, die die Auflösung von medizinischen Bildern erhöhen kann [130, 179].

Die vollautomatische Atmungserkennung ist eine sinnvolle Ergänzung, da sie nicht nur die automatische Entfernung von veratmeten Zeitpunkten in der DCE-MRT Messung erlaubt, sondern auch Zeitpunkte identifizieren kann, die einer Korrektur bedürfen. Des Weiteren ist die Atmungserkennung ein zentraler Bestandteil des entwickelten DCE-MRT Qualitätsscores, der in der Lage war, Messungen zu identifizieren, die fehlerhafte Ergebnisse enthielten. Der Qualiätsscore könnte auch das Vertrauen von zukünftigen AnwenderInnen in die Pipeline erhöhen, da er auf fehlerhafte Quantifizierungen hinweisen könnte.

Die aktuelle Version der Perfusionsquantifizierungs-Pipeline besitzt jedoch noch ihre Limitationen. Das Problem fehlerhafter Registrierung wurde zuvor schon erwähnt und verhindert sicherlich noch stärkere Korrelationen. Außerdem zeigte sich eine deutliche Anfälligkeit der Perfusionsquantifizierung für veratmete DCE-MRT Messungen. Dieses Problem sollte in der Zukunft unbedingt näher

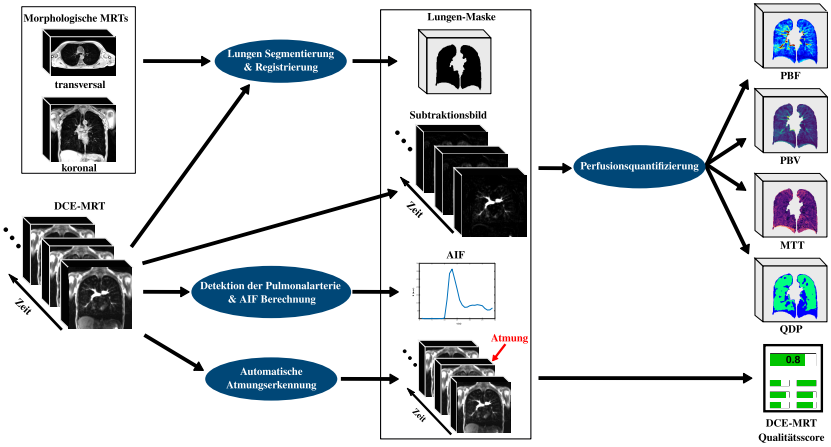


Abbildung 5.28: Darstellung der entwickelten vollautomatischen und robusten Pipeline zur Perfusionsquantifizierung für multizentrische Studien.

betrachtet und eine Methode zur Atmungskorrektur entwickelt werden. Des Weiteren unterschieden sich die Ergebnisse zwischen den einzelnen Zentren. Jedoch ist zu beachten, dass es sich um unterschiedliche PatientInnen handelt. Daher wäre eine Studie, bei der die gleichen PatientInnen bzw. ein richtiges DCE-MRT Lungenphantom an verschiedenen Scannern innerhalb kurzer Zeit gemessen werden, sehr hilfreich, um diesen Einfluss näher zu untersuchen.

Insgesamt lässt sich feststellen, dass die entwickelte Pipeline eine vollautomatische Perfusionsquantifizierung von DCE-MRT Messungen und damit die Auswertung an Zentren, an denen keine Expertise zur Bewertung von DCE-MRT Messungen vorliegt, ermöglicht. Zusätzlich könnte durch die Methode die zeitintensive Bewertung durch RadiologInnen, z. B. mit dem groben Eichinger-Score obsolet werden und RadiologInnen könnten sich auf die Bewertung der quantitativen Perfusionskarten beschränken. Die quantitativen Perfusionskarten ermöglichen aufgrund der sehr feingranularen Betrachtung einzelner Lungenbereiche die quantitative Verfolgung von Perfusionsdefekten und ermöglichen den Erkrankungsverlauf oder den Behandlungsfortschritt näher zu verfolgen.

6 Daten-Workflows für die Bildverarbeitung

Das Forschungsdatenmanagement (FDM) beschreibt ein Vorgehen, das Daten sowie deren zugehörige Metadaten strukturiert ablegt und diese so abspeichert, dass die Daten durchsucht und weiterverwendet werden können. Das Forschungsdatenmanagement spielt spätestens seit der Veröffentlichung der „Leitlinien zum Umgang mit Forschungsdaten“ durch die Deutsche Forschungsgemeinschaft eine wichtige Rolle in der Wissenschaft. Darin wird beschrieben, dass ein gutes FDM die Nachvollziehbarkeit von Forschungsdaten sowie deren weitere Nutzung in nachfolgenden Forschungsprojekten verbessert. [180] Auch für die Entwicklung von robusten und flexiblen Methoden spielt das FDM eine wichtige Rolle. Zum einen garantiert ein einheitliches, strukturiertes Ablegen von Daten, dass diese direkt mit passenden Methoden ausgewertet werden können, ohne das zuvor zusätzliche Datenkonvertierungsschritte durchgeführt werden müssen. Zum anderen fördert die Dokumentation der einzelnen Entwicklungsschritte einer Methode deren Weiterverwendung, da genau nachvollzogen werden kann, warum entsprechende Modellparameter gewählt wurden. In diesem Kapitel soll nun explizit auf das Forschungsgebiet der medizinischen Bildverarbeitung eingegangen werden und Konzepte vorgestellt werden, wie mithilfe von Daten-Workflows eine strukturierte Dokumentation von Forschungsdaten erreicht werden kann. Abbildung 6.1 zeigt die Einordnung der entwickelten Daten-Workflows in den vorgestellten informationsgetriebenen Bildanalysezyklus aus Kapitel 3. Hierbei stellen die Daten-Workflows die Verbindungen zwischen den entwickelten Methoden, den verwendeten und erzeugten Forschungsdaten sowie einem Repositorium dar.

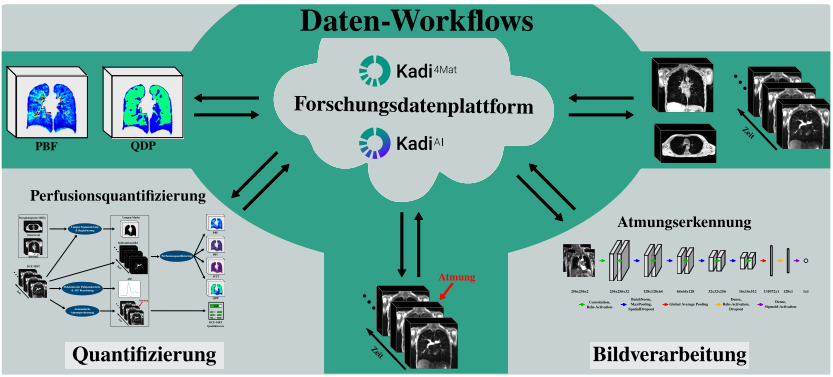


Abbildung 6.1: Einordnung der entwickelten Daten-Workflows in den Bildanalysezyklus.

6.1 Stand der Forschung

Die Grundlage des FDM stellen die sogenannten FAIR-Prinzipien (**F**indable, **A**ccessible, **I**nteroperable, **R**eusable) dar. Diese liefern eine Empfehlung wie Daten optimalerweise abgespeichert werden sollten. Die Auffindbarkeit (engl. findability) soll garantieren, dass die Daten anhand eines eindeutigen Identifikators und ihrer Metadaten aufgefunden werden können. Die Zugänglichkeit (engl. accessibility) beschreibt die Abrufbarkeit der Daten über standardisierte Verfahren oder Protokolle, die je nach Anwendungsfall an Zugangsvoraussetzungen geknüpft sein kann. Interoperabilität (engl. interoperability) erfordert die Verwendung von einheitlichen Datenformaten und Begriffen, die beispielsweise über eine Art semantisches Wörterbuch definiert sein können. Eine Wiederverwendung (engl. reusability) soll eine Weiterverwendung der Daten für ähnliche, aber auch andere Fragestellung ermöglichen. Hierzu sollten die Daten mit detaillierten Metadaten versehen sein und die Verknüpfung mit anderen Daten sowie deren Herkunft klar ersichtlich sein. [181]

Im klinischen Alltag werden medizinische Bilder in einem standardisierten Format, dem sogenannten DICOM-Format [182] abgespeichert. Die DICOM-Dateien

enthalten sowohl die eigentlichen Bilddaten, aber auch Metadaten, wie die genauen Aufnahmebedingungen oder Informationen zur Patientin oder zum Patienten. Dabei wird ein Bild nicht als einzelne Datei, sondern als einzelne DICOM-Dateien für jede Schicht und/oder Zeitpunkt des Bildes abgespeichert. Ein MRT-Bild mit 64 Schichten wird also in 64 einzelnen Dateien gespeichert. Für den klinischen Alltag ist dies unproblematisch, da die DICOM-Bilder in einem sogenannten PACS-System (vom Englischen Picture Archiving and Communication System) abgespeichert werden, das das Durchsuchen und Betrachten der Bilder ermöglicht. Mittlerweile besitzen die PACS-Systeme auch zusätzliche Softwaretools, die dann direkt auf die Bilder angewendet werden können. [183] Aus diesem Anlass werden die Bilder für den klinischen Alltag im PACS-System systematisch abgespeichert und erfüllen somit die Anforderungen eines FDM-Systems. Für die Entwicklung neuer Bildverarbeitungsmethoden müssen diese Bilddaten jedoch meist in ein alternatives Format überführt werden, da in Programmiersprachen meist mit Tensoren und Arrays gearbeitet wird und ein dreidimensionales Bild meistens als Ganzes betrachtet wird. Ein weiterer Grund für eine Konvertierung ist der geringere Speicherbedarf im Vergleich zu den vielen einzelnen DICOM-Dateien pro Bild. Daher werden die Bilder aus dem PACS-System exportiert, in ein Bildverarbeitungsformat wie z. B. NIFTI, MHD oder MAT konvertiert und meist lokal auf Festplatten abgespeichert, um sie dann weiterzuverarbeiten. Anschließend werden je nach Bedarf Methoden entwickelt und/oder angewendet um die Bilddaten näher zu analysieren.

Für die Verwendung wissenschaftlicher Methoden in einem interdisziplinären Team sind benutzerfreundliche akademische Frameworks wichtig, sodass auch Personen mit unterschiedlicher Expertise die Methoden zur Datenanalyse verwenden können [184]. Obwohl analytische und ML-Methoden eine wichtige Rolle in der Medizin spielen, fehlen oft passende Schnittstellen zu den genannten PACS-Systemen [185, 186]. Hauptprobleme hierbei sind die große Anzahl an medizinischen Softwaretools und deren spezifische Datenformate sowie die Nichteinhaltung von Datenstandards [187]. Es gibt eine Vielzahl an entwickelten Tools, die sich jeweils einem oder mehreren dieser Probleme angenommen haben. Die Software REDCap [188, 189] unterstützt dabei, klinische Daten zu sammeln und

diese in klinischen Forschungsprojekten zu verwalten. Hier werden jedoch keinerlei Bilddaten unterstützt. Die XNAT Plattform [190] beinhaltet Methoden zum Einlesen, Betrachten und Analysieren von Bildern und ermöglicht die Ausführung von automatischen Workflows. Die Tools KAAPANA [191] und MONAI [192] stellen Plattformen für ML-Tools und Workflows für die medizinische Bildverarbeitung zur Verfügung. Sie können einfach mit PACS-Systemen verbunden werden und besitzen integrierte Bildbetrachter.

Das in dieser Arbeit verwendete Tool ist die Open-Source Lösung Kadi4Mat [99], die ursprünglich für das FDM in den Materialwissenschaften entwickelt wurde. Dabei vereint Kadi4Mat ein elektronisches Laborbuch mit Workflow Funktionalitäten [193] mit einem strukturierten Datenrepositorium [99]. Mithilfe entsprechender Schnittstellen können spezialisierte Analysetechniken wie die Simulation oder ML-Methoden in die Workflows eingebunden werden [194, 195]. Für die Entwicklungen in dieser Arbeit spielt jedoch nur das Datenrepositorium von Kadi4Mat (KadiWeb) eine Rolle. KadiWeb besteht aus drei wesentlichen Hauptkomponenten, den Records, den Collections und den Templates. Records bieten die Möglichkeit, Daten mit ihren Metadaten zu verknüpfen. Hierzu können Dateien zum Record hinzugefügt werden und die dazugehörigen Metadaten innerhalb eines Wörterbuchs aus Key-Value-Paaren hinterlegt werden. Die Metadaten bestehen dabei aus den standardmäßigen Metadaten eines Records, wie dem Titel, dem einzigartigen Identifikator und dem Erstellungszeitpunkt sowie anwendungsspezifischen Extrawerten. Zwischen zwei Records können Verlinkungen hinzugefügt werden, die die Beziehung zwischen den beiden Records beschreiben. Die Verlinkungen besitzen jeweils eine Richtung und eine Bezeichnung. Die Collections werden in Kadi4Mat verwendet, um Records zu gruppieren. Dagegen dienen Templates als Blaupausen für Records, die alle einem gewissen Schema folgen sollen. Zusätzlich bietet Kadi4Mat ein feingranulares Zugangsmanagement, bei dem einzelne Records oder Collections für Nutzerinnen und Nutzer sowie Gruppen freigeschaltet werden können. [99]

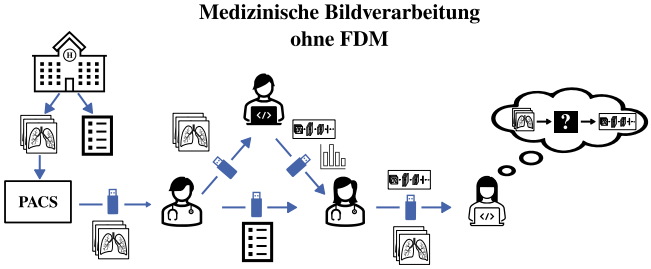


Abbildung 6.2: Darstellung der medizinischen Bildverarbeitung ohne FDM.

6.2 FDM für die medizinische Bildverarbeitung

Ein typischer Ablauf eines interdisziplinären Projektes im Bereich der medizinischen Bildverarbeitung ist in Abbildung 6.2 dargestellt. Die Bilder und weitere medizinische Daten werden in der Klinik erfasst. Das PACS-System dient der Speicherung der medizinischen Bilddaten. Soll nun eine Studie basierend auf bestimmten Bildern durchgeführt werden, so werden diese zumeist von RadiologInnen aus dem PACS exportiert und auf einem USB-Stick oder einer Festplatte gespeichert. Dieses Speichermedium wird dann zur Datenanalyse an entsprechende SoftwareentwicklerInnen weitergegeben. Nach der Analyse der Daten werden die Ergebnisse und gegebenenfalls auch die entwickelten Modelle per Speichermedium wieder für die klinische Beurteilung weitergereicht. Bei Bedarf kann die entwickelte Methode zu einem späteren Zeitpunkt erneut verwendet und optimiert werden, um andere Bilddaten auszuwerten. Falls die ursprünglich mit der Methodenentwicklung betraute Person nicht mehr greifbar ist, wird eine neue Person mit der Auswertung der Daten beauftragt. Da jedoch nur die Methode und nicht das Vorgehen zur Entwicklung der Methode auf dem Speichermedium gespeichert wurden, sind mögliche Erkenntnisse und Parameterstudien zur entwickelten Methode unbekannt und müssen gegebenenfalls erneut durchgeführt werden.

Dieser Ablauf bringt mehrere Nachteile mit sich. Einerseits werden die Daten regelhaft mit externen Speichermedien transferiert, was kompliziert ist, aber auch

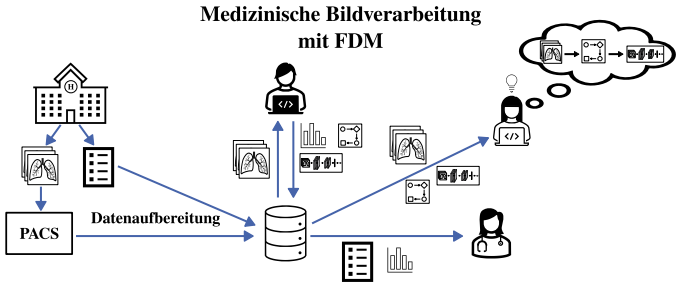


Abbildung 6.3: Darstellung der medizinischen Bildverarbeitung mit FDM.

die Gefahr birgt, dass die Daten verändert werden. Andererseits könnte es je nach Export vorkommen, dass die Daten nicht mehr alle nötigen Metadaten besitzen. Ein weiterer wesentlicher Nachteil ist die fehlende Dokumentation des Vorgehens zur Entwicklung und Verwendung der Methoden. Falls dieses nicht in Form einer Veröffentlichung publiziert wurde, kann dieses Wissen eventuell verloren gehen. Um diese Probleme zu beheben wird, wie in Abbildung 6.3 dargestellt, ein alternatives Vorgehen vorgeschlagen, welches die Abläufe durch den Einsatz von FDM optimiert.

In dieser Variante werden die Bilddaten aus dem PACS sowie weitere klinische Informationen aufbereitet und innerhalb einer Datenbank eines FDM-Tools zentral gespeichert. Sollen nun Daten analysiert werden, können die Daten direkt aus der Datenbank extrahiert werden. Anschließend findet die Entwicklung einer Methode zur Datenauswertung lokal statt. Nach der erfolgreichen Methodenentwicklung werden diesmal die Ergebnisse, das entwickelte Modell sowie die Entwicklungsschritte protokolliert und innerhalb der FDM-Datenbank abgespeichert. Aus dieser FDM-Datenbank können sowohl klinische MitarbeiterInnen die Auswertungsergebnisse beziehen, als auch zukünftige Forschende die entwickelten Modelle und deren Entwicklungsschritte herunterladen. Dieses Vorgehen macht das Herumreichen von Speichermedien obsolet und fördert die Nachvollziehbarkeit der Ergebnisse, da nun auch die Entwicklungsschritte klar dokumentiert sind.

Die nachfolgenden Unterkapitel beschreiben die Umsetzung des vorgeschlagenen Konzepts und wie dieses für die Dokumentation der entwickelten Methoden und

generierten Ergebnisse der Perfusionsquantifizierungs-Pipeline verwendet werden kann.

6.3 Aufbereitung und Exportieren der Bilddaten

Die MRT-Bilder der COSYCONET-Studie waren alle als DICOM-Dateien auf einer Festplatte gespeichert. Da dieses Dateiformat, wie zuvor beschrieben, nicht für die Bildverarbeitung geeignet ist, mussten diese zunächst in ein alternatives Dateiformat konvertiert werden. Da die Bilder neben den Bilddaten auch Metadaten besaßen, musste ein Dateiformat bzw. eine Kombination von Dateiformaten gefunden werden, die Bilder und Metadaten abspeichern können. Die Wahl fiel auf eine Kombination des weit verbreiteten NIFTI-Formats für die Bilddaten und dem JSON-Format für die Metadaten. Für das Einlesen der Bilder wurde ein vollautomatisches Python-Script geschrieben, dass zunächst einen Ordner mit DICOM-Dateien durchsucht und daraus ein geordnetes Python-Dictionary erstellt, um einen Überblick über die vorhandenen Bilder zu erhalten. Das Python-Dictionary besteht dabei aus fünf Ebenen: PatientInnen-ID, Aufnahmedatum, Studiennummer, Protokollname und Seriennummer. Alle diese Informationen können jeweils aus dem DICOM-Header einer DICOM-Datei extrahiert werden. Für die Zuordnung einer DICOM-Datei zu einem Bild wurden die DICOM-Metadaten Felder „SeriesInstanceUID“ einer jeden Datei verwendet. Jeder Seriennummer sind damit ein oder mehrere Bilder zugeordnet. Für ein 2D-Bild konnte diese Serie, die dann nur aus einem Bild besteht, direkt gespeichert werden. Bei 3D-Bildern wurden alle Einzelbilder einer Serie anschließend anhand ihrer Position geordnet und zu einem Volumen zusammengefügt. Im Sonderfall eines 4D-Bildes, wie bspw. dem TWIST-Bild, mussten mehrere Serien, die jeweils ein 3D-Bild eines Zeitpunkts repräsentieren, entlang der Zeitachse geordnet und ebenso zusammengefügt werden. Nachdem nun die Bilder als 2D, 3D oder 4D Volumen in Form eines Arrays vorlagen, wurden diese als NIFTI-Datei abgespeichert. Da viele der Metadaten-Felder der DICOM-Einzeldateien eines Bildes identisch sind, wurde

ein gemeinsames Python-Dictionary erstellt, das nur nicht redundante Daten aller zugehörigen DICOM-Header enthielt. Diese wurden anschließend in einer JSON-Datei gespeichert. Damit konnte die Anzahl der Dateien pro Bild von teilweise mehreren hundert Einzeldateien auf zwei Dateien reduziert werden.

Für jedes Bild wurde automatisch über die KadiWeb API-Schnittstelle ein Record generiert und die beiden Dateien pro Bild in den Record hochgeladen. Zusätzlich wurde jeweils ein Vorschaubild aus dem multidimensionalen Bild extrahiert und dieses zusätzlich hochgeladen. Mithilfe des JSON-Dictionary konnten die Metadaten des Records eines Bildes anschließend automatisch ausgefüllt werden. Zusätzlich wurde pro Bild ein Metadateneintrag in Form eines Wörterbuchs erzeugt, der die wichtigsten Informationen beinhaltet. Diese Informationen bestehen aus der PatientInnen-ID, dem Geschlecht, dem Alter, dem Messzeitpunkt, dem Namen der Klinik, der Modalität des Bildes, dem Protokollnamen und der Krankheitsgruppe der untersuchten Person. Ein beispielhafter Record eines MRT-Bildes ist in Abbildung 6.4 dargestellt. Man erkennt, dass sich der Name des Records aus der PatientInnen-ID, dem Messzeitpunkt und dem Protokoll zusammensetzt. Zudem ist in der Beschreibung des Records das zuvor hochgeladene Vorschaubild dargestellt. Die wichtigsten Informationen befinden sich im Metadatenfeld „General Information“.

Für den besseren Überblick über alle Studien und Bilder einer untersuchten Person wurde pro PatientIn zusätzlich ein Übersichts-Record erstellt. Dieser beinhaltete wichtige Metadaten zur untersuchten Person und den einzelnen Bildern. In Abbildung 6.5 ist ein Beispiel eines solchen Records dargestellt. Dieser Überblick-Record erleichtert die automatische Durchsuchung der Datenbank, da die Suche nun auf alle Records des Typs „medical-study-overview“ beschränkt werden kann. Damit lassen sich Statistiken über alle untersuchten PatientInnen sowie die durchgeführten Messungen erstellen. Alle Records der Bilder einer untersuchten Person sowie der Studienüberblick-Record wurden final in einer eigenen Collection gesammelt.

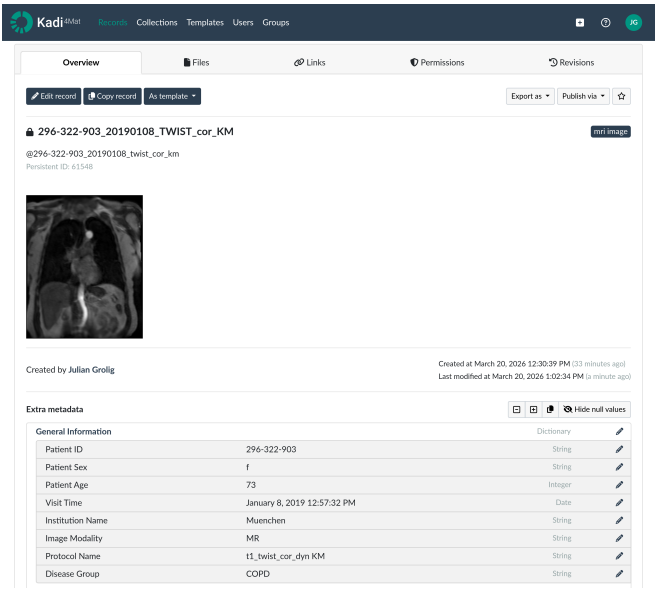


Abbildung 6.4: Kadi4Mat Record eines MRT-Bildes.

6.4 FDM für ML-Projekte

Die Dokumentation der einzelnen Schritte bei der Entwicklung von neuen Methoden wie z. B. ML-Modellen ist ein wichtiger Schritt, um die Weiterverwendung der entwickelten Modelle zu gewährleisten. Ein häufiges Problem hierbei ist, dass die strukturierte Ablage von Zwischenergebnissen oft ein sehr zeitintensives Unterfangen ist. Daher wurde ein Konzept entwickelt, das eine vollautomatische Dokumentation eines kompletten ML-Projekts durchführt. Für die Entwicklung der ML-Modelle in dieser Arbeit wurde das CIDS-Framework [196, 197] und die darin enthaltene KadiAI-Schnittstelle verwendet, die Funktionalitäten zur Kommunikation mit KadiWeb zur Verfügung stellt.

Zunächst wurde ein Ontologie-ähnliches Konzept entwickelt, das die Zusammenhänge zwischen den einzelnen Bestandteilen eines KadiAI Projektes definiert. Dieses Konzept ist in Abbildung 6.6 visualisiert.

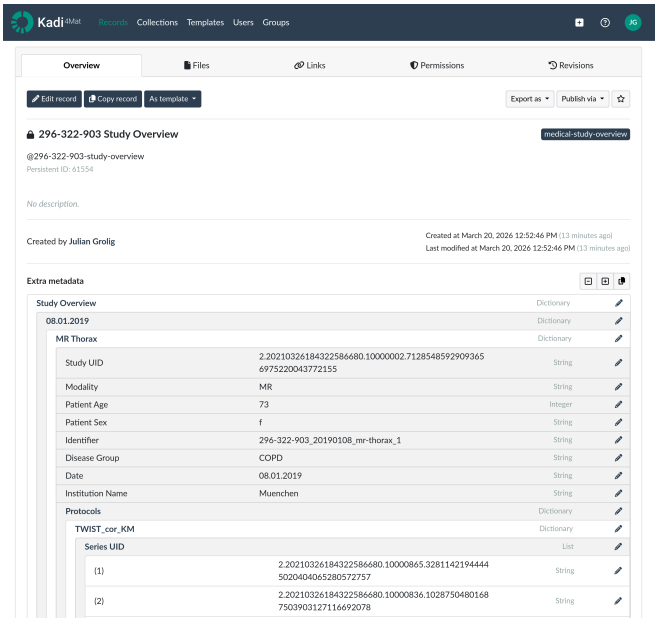


Abbildung 6.5: Kadi4Mat Record zum Überblick über alle durchgeführten Bildstudien einer Patientin.

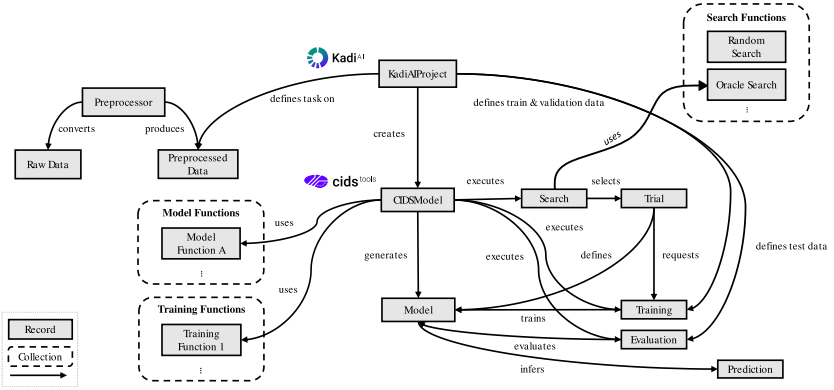


Abbildung 6.6: Konzept zur Beschreibung der Zusammenhänge eines KadiAI Projektes.

Hierbei spielt das KadiAI-Projekt (KadiAIProject) die zentrale Rolle. Das KadiAI-Projekt definiert die Aufgabe auf einem vorverarbeiteten Datensatz und teilt die Daten in einen Trainings-, Validierungs- und Testdatensatz ein. Außerdem erstellt es ein sogenanntes CIDSModel, welches ein Hypermodell repräsentiert, also eine ML-Architektur und deren dazugehörige Hyperparameter. Das CIDSModel verwendet eine Modellfunktion und eine Trainingsfunktion, die jeweils aus einem Pool von möglichen Funktionen gewählt werden und erzeugt damit dann ein ML-Modell (Model). Außerdem kann das CIDSModel mit einer der verschiedenen Suchfunktionen eines Pools an Suchfunktionen auch eine Hyperparametersuche (engl. Search) ausführen. Die Hyperparametersuche selbst wählt eine Kombination an Hyperparametern aus und definiert damit einen Versuch (engl. Trial), der ein Training eines spezifischen Modells anfragt. Das trainierte Modell kann dann entweder für eine Validierung (engl. Evaluation) auf dem Validierungsdatensatz oder aber für eine Vorhersage auf dem Testdatensatz verwendet werden.

Basierend auf diesem Konzept wurde eine Funktion in KadiAI erstellt, die ein Projekt automatisch anhand dieses Konzeptes in KadiWeb hochlädt und die einzelnen Komponenten entsprechend miteinander verbindet. Umgekehrt ermöglicht eine zweite Funktion das Herunterladen eines kompletten KadiAI-Projektes, sodass die originale Ordnerstruktur lokal wiederhergestellt wird. Die Funktion wurde anhand des Projektes zur Entwicklung eines Atmungserkennungs-Modells aus Abschnitt 5.4 validiert. Nach der Fertigstellung des ML-Projektes wurde das komplette Projekt mithilfe der entwickelten KadiAI-Funktion in KadiWeb hochgeladen. In Abbildung 6.7 ist der automatisch erzeugte Wissensgraph des kompletten Atmungserkennungsprojektes dargestellt. Man erkennt links unten sehr gut die Aufspaltung in drei Blätter, die jeweils eines der verwendeten Hypermodelle repräsentieren. Bei der Betrachtung nur eines dieser Blätter erkennt man die Vielzahl an einzelnen Records, die für das ML-Projekt erstellt wurden, die alle untereinander mittels Links verbunden sind. Auf der rechten Seite ist eine vereinfachte Form dargestellt, in der jede Record-Art nur einmal vorkommt. Die unterschiedlichen Record-Arten sind alle mit eigenen Farbcodierungen markiert.

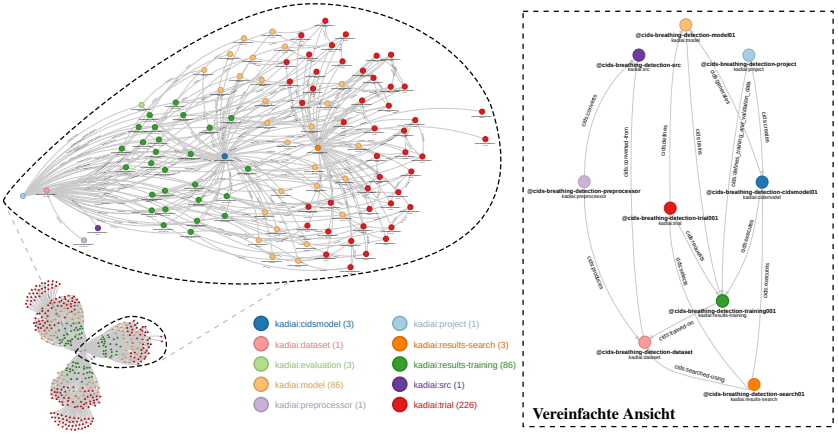


Abbildung 6.7: Kadi4Mat Wissensgraph des Atmungserkennungs-Projektes. Links unten ist der Wissensgraph des kompletten Projektes, links oben der Wissensgraph eines der drei Hypermodelle und rechts eine vereinfachte Ansicht dargestellt.

(a) (b)

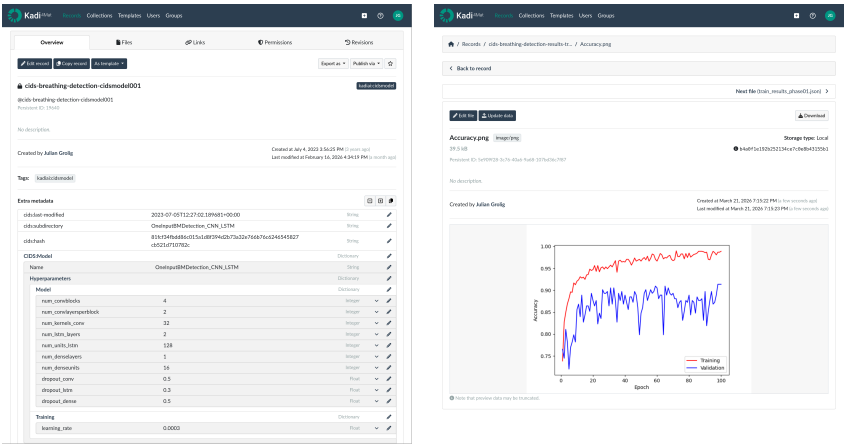


Abbildung 6.8: Beispiele von Records innerhalb des KadiAI-Projekts. (a) Record mit Metadaten eines Hypermodells (b) Record eines Trainings, mit dazugehörigem Trainingsplot.

Die erstellten Records enthielten alle Bestandteile des lokalen ML-Projektes. In Abbildung 6.8 (a) ist ein Record eines Hypermodells (CIDSModel) dargestellt. Die definierten Hyperparameter wurden innerhalb der Record-Metadaten definiert, wobei auch die erlaubten Bereiche/Werte für jeden Parameter hinterlegt wurden. Die beigefügten Dateien für jeden Record waren objektspezifisch. Für einen Trainings-Record wurden neben einer JSON-Datei, die die Werte aller Metriken enthält, unter anderem auch automatisch Plots der beobachteten Metriken während des Trainings erstellt. Ein solcher Plot der Genauigkeit (engl. Accuracy) ist in Abbildung 6.8 (b) zu sehen.

Der Upload des gesamten ML-Projektes inklusive der Generierung des kompletten Wissensgraphes geschah komplett automatisch durch die Ausführung einer Code-Zeile eines Python-Skriptes. Damit wurde das ML-Projekt nachhaltig archiviert und alle Zusammenhänge können dank der Verbindungen klar nachvollzogen werden. Ein manuelles Hochladen all dieser Einzelbestandteile wäre sehr zeitintensiv gewesen und hätte möglicherweise auch zu Inkonsistenzen führen können. Somit werden durch die Automatisierung dieses Schrittes zeitliche Ressourcen geschont, die wiederum zu Forschungszwecken genutzt werden können.

6.5 FDM für die Dokumentation der COSYCONET-Ergebnisse

Die einzelnen Zwischenergebnisse und die quantitativen Perfusionswerte, die bei der Auswertung der COSYCONET-Kohorte entstanden sind, wurden ebenfalls durch die Verwendung von Kadi4Mat automatisch dokumentiert. Zunächst wurde für jeden verwendeten MRT-Scanner in der COSYCONET-Studie ein eigener Scanner-Record angelegt, der die wichtigsten Metadaten des MRT-Scanners, wie das Modell oder das Gradientensystem beinhaltete. Ein beispielhafter Scanner-Record ist in Abbildung 6.9 dargestellt. Anschließend wurde pro PatientIn eine PatientInnen-Collection erstellt, zu der zunächst die drei Bild-Records der verwendeten MRT-Bilder hinzugefügt wurden. Die Bild-Records wurden dabei mit dem

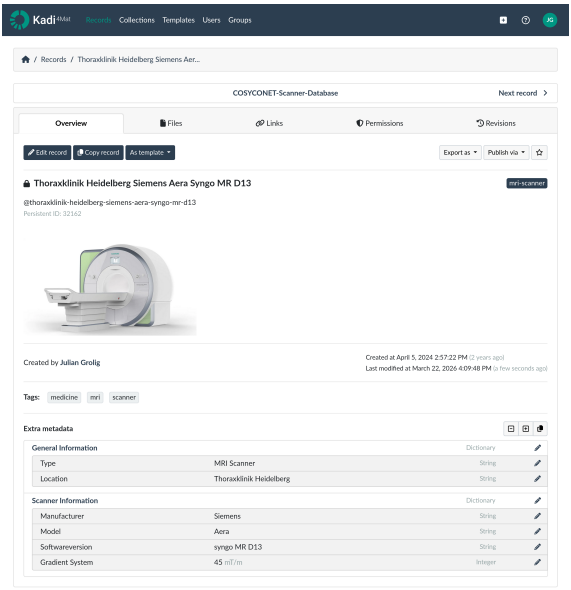


Abbildung 6.9: Kadi4Mat Record eines MRT-Scanners.

Scanner-Record des MRT-Scanners verknüpft, mit dem sie aufgenommen wurden. Auch die erfassten klinischen Daten sowie die quantitativen CT-Werte wurden jeweils in einem eigenen Record gespeichert und der PatientInnen-Collection hinzugefügt. Die beiden Records für eine beispielhafte Patientin sind in Abbildung 6.10 dargestellt. Die Daten wurden dabei in Form einer CSV-Tabelle innerhalb des Records gespeichert und zusätzlich alle Werte in die Metadaten des Records übertragen, um so eine direkte Visualisierung der gemessenen Werte zu erhalten.

Für die Ergebnisse der Perfusionsquantifizierungs-Pipeline wurden nicht nur die quantitativen Endergebnisse, sondern auch alle Zwischenergebnisse wie z. B. die Lungenmasken vor und nach der Registrierung oder die Atmungswahrscheinlichkeiten jeweils in eigenen Records abgespeichert. Dies verbessert die Nachvollziehbarkeit der durchgeführten Schritte und ermöglicht eine genaue Verlinkung der einzelnen Komponenten. Auch hier wurden sowohl die Daten in Form von

(a) (b)

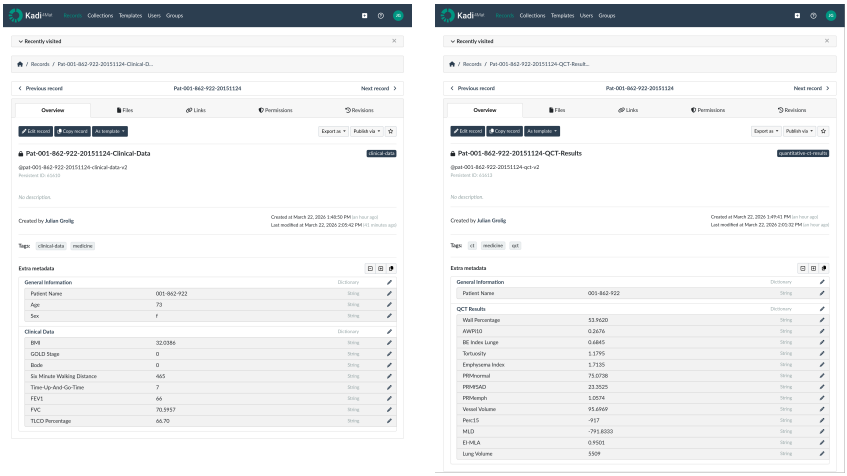


Abbildung 6.10: Darstellung zweier Kadi4Mat Records für (a) die klinischen Daten und (b) die QCT Daten einer beispielhaften Patientin.

Matlab-Dateien in die jeweiligen Records hochgeladen, als auch die Daten, sofern möglich, in die Metadaten eines jeden Records übertragen. Ebenso wurde ein eigener Record für die quantitativen Perfusionswerte erstellt, der neben den Matlab-Dateien auch jeweils die Histogramme der Perfusionswerte als Plots und die berechneten Mediane der Perfusionswerte als Teil der Metadaten enthielt. In Abbildung 6.11 ist ein solcher Record für die quantitativen Perfusionswerte einer Patientin dargestellt. Man erkennt sowohl die zusammengefassten Perfusionswerte innerhalb der Metadaten in Abbildung 6.11 (a), als auch ein geplottetes Histogramm für PBF in Abbildung 6.11 (b).

Aus allen erstellten Records einer untersuchten Person konnte nun automatisch ein Wissensgraph erzeugt werden, der alle (Teil-) Ergebnisse und deren Verknüpfungen darstellt. Ein beispielhafter Wissensgraph ist in Abbildung 6.12 visualisiert. Hier wird gezeigt, dass der Wissensgraph einerseits allgemeine Daten wie die klinischen Daten, die QCT-Daten, die MRT-Bilder, die Atmungswahrscheinlichkeiten und den Qualitätsscore enthielt. Andererseits wurden aber auch jeweils für jede der drei Registrierungsvarianten alle separaten Zwischenergebnisse und die

(a) (b)

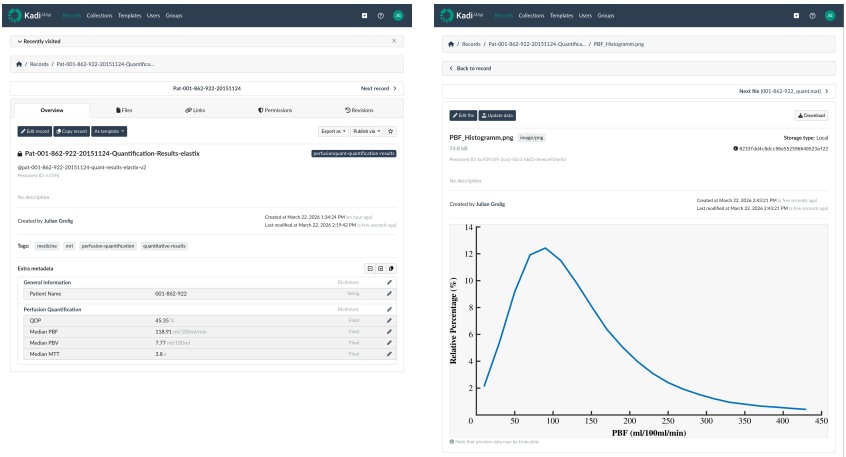


Abbildung 6.11: Kadi4Mat Record der quantitativen Perfusionswerte: (a) Metadaten mit Median-Werten und QDP (b) Beigefügte Histogramme der quantitativen Perfusionswerte.

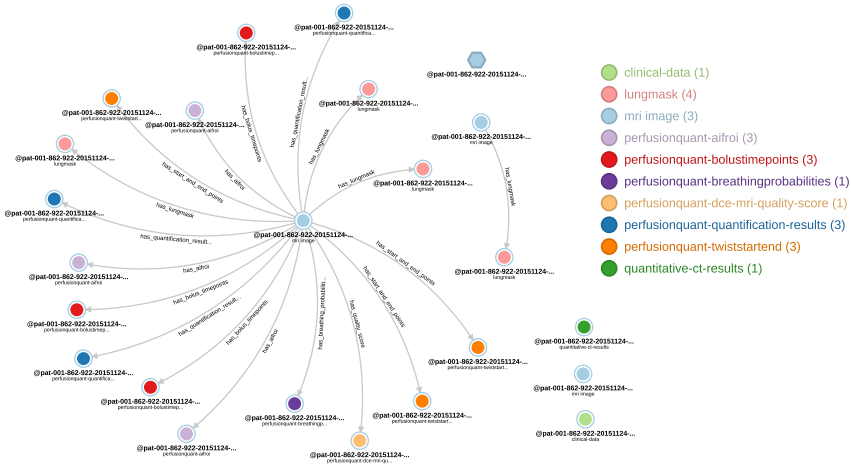


Abbildung 6.12: Darstellung eines Kadi4Mat Wissensgraphen für die gemessenen und bestimmten Daten einer Patientin.

finalen Perfusionswerte in Records abgespeichert. Ein solcher Wissensgraph kann für alle PatientInnen in der COSYCONET-Studie erzeugt werden und in Zukunft durch neue Ergebnisse ergänzt werden.

6.6 Einordnung des entwickelten FDM-Konzeptes

In dieser Arbeit wurde ein FDM-Konzept für die Verwaltung von medizinischen Bildverarbeitungsprojekten entwickelt. Dieses ermöglicht die strukturierte Ablage von medizinischen Bildern im DICOM-Format, in einem für die Bildverarbeitung geeigneten Format. Dabei bleiben alle relevanten Metadaten eines Bildes erhalten und werden als Teil des Bild-Records gespeichert. Dies ist ein zentraler Schritt, der eines der in Abschnitt 6.2 genannten Probleme löst und den Austausch von Bildern mittels portabler Speichermedien obsolet macht. Zusätzlich wird hierdurch garantiert, dass alle Forschenden mit den gleichen Daten arbeiten und das Importieren der Bilder nicht von allen Forschenden separat durchgeführt werden muss.

Zudem unterstützt das entwickelte Konzept die automatische Dokumentation von ML-Projekten. Durch das flexible Konzept zum vollautomatischen Exportieren und Importieren eines ML-Projektes nach/von Kadi4Mat, wird die häufig verwendete lokale Entwicklung eines ML-Projektes ermöglicht. Dabei wird beim Download des Projektes eine einheitliche Ordnerstruktur erzeugt, selbst wenn das Projekt von einer anderen Person heruntergeladen wird. Das vollautomatische Vorgehen erspart viel Arbeit, da das Hochladen der einzelnen Projekt-Bestandteile und deren Verlinkung keinerlei User-Interaktion benötigt. Hiermit wird auch ein einheitliches Konzept für das Speichern und Verknüpfen der Daten garantiert. Die einzige Limitation des Konzeptes zur Dokumentation von ML-Projekten ist, dass das Konzept aktuell an das Open-Source Framework *CIDS* gebunden ist und daher nur eine Verwendung mit in *CIDS* entwickelten Projekten möglich ist. Jedoch könnte das Konzept auch auf andere Entwicklungsumgebungen angewendet

werden und hierfür eine eigene automatische Upload- und Download-Funktion mithilfe der KadiWeb API-Schnittstelle entwickelt werden.

Der dritte Bestandteil des entwickelten Konzeptes ist die automatische Dokumentation der Ergebnisse des Projektes zur Perfusionsquantifizierung in der COSYCONET-Studie. Hier wurde gezeigt, wie alle Zwischenergebnisse, aber auch die finalen quantitativen Perfusionswerte automatisch in Kadi4Mat dokumentiert werden können. Dabei wurden die einzelnen Records anhand eines einheitlichen Konzeptes miteinander verbunden, was die Nachvollziehbarkeit der Ergebnisse für eine eventuelle spätere Weiternutzung der Daten verbessert.

Die Verwendung von FDM-Tools wie Kadi4Mat ist insbesondere in interdisziplinären Forschungsgruppen eine hilfreiche Ergänzung. Sie garantiert den unkomplizierten Austausch von Forschungsdaten und ermöglicht eine enge Zusammenarbeit aller Beteiligten. Zusätzlich wird hierdurch die Dokumentation der durchgeführten Schritte verbessert, was die Weiterverwendung der erzielten Forschungsergebnisse in nachfolgenden Projekten vereinfacht.

7 Zusammenfassung und Ausblick

7.1 Zusammenfassung

In dieser Arbeit wurden robuste und flexible Methoden zur quantitativen Analyse von multidimensionalen Bilddaten für die Verwendung in interdisziplinären Forschungsprojekten entwickelt. Hierbei lag der Fokus auf den beiden Disziplinen der interventionellen Radiologie und den Materialwissenschaften, da die Bildverarbeitung in beiden Fachgebieten eine wesentliche Rolle spielt. Als Grundlage für diese Arbeit wurde das Konzept des informationsgetriebenen Bildanalysezyklus erarbeitet. Der Gewinn und die strukturierte Aufbewahrung von Informationen in Form von Daten und Methoden waren Kernpunkte dieses Konzeptes. Ein weiterer wichtiger Bestandteil waren flexible und robuste Methoden, da diese den breiten Einsatz über die Disziplinen hinweg, aber auch für unterschiedliche Messbedingungen ermöglichen.

Für das in der Bildverarbeitung zentrale Thema der Segmentierung wurden in dieser Arbeit drei flexible Modelle entwickelt, die für die Segmentierung von Objekten in unterschiedlichen Bildarten verwendet werden können. Ein wesentliches Ziel war es, eine Methode zu finden, die keinerlei annotierte Groundtruth-Daten für das Training bzw. die Segmentierung benötigt. Für die Validierung der Modelle wurden drei verschiedene Datensätze, zwei PU-Schaum-Datensätze aus den Materialwissenschaften und ein Lungen-MRT-Datensatz aus der Radiologie, verwendet. Zwei der Modelle basierten auf verborgenen, zufälligen Markov-Feldern, die sowohl die Nachbarschaft eines Voxels als auch dessen Zugehörigkeitswahrscheinlichkeit zu einer Verteilung bewerten, um den Voxel einer Klasse zuzuordnen. Das erste Modell EHOHMRF verwendete zur Lösung des

HMRF-Optimierungsproblems einen Elefantenherden-Optimierungs-Algorithmus. Das Modell erzeugte für alle drei Datensätze gute Segmentierungsergebnisse und zeigte damit sein Potenzial als flexible Methode für die Segmentierung in verschiedenen Anwendungsgebieten. Jedoch besaß die Methode den großen Nachteil einer sehr langen Rechenzeit für jede neue Segmentierung, da sie für jedes neue Bild erneut trainiert werden musste.

Um das Problem der langen Rechenzeiten zu beheben, wurde das HMRF-Konzept mit einem neuronalen Netz zum HMRF-UNet Modell kombiniert. Durch die Verwendung einer unüberwachten HMRF-basierten Loss-Funktion in einem klassischen U-Net, konnten die Vorteile der HMRF-Methode und der neuronalen Netze kombiniert werden. Die generierten Segmentierungen besaßen eine hohe Genauigkeit und konnten sowohl für 2D-Bilddaten, als auch für 3D-Bilddaten innerhalb von wenigen Sekunden erzeugt werden. Zusätzlich konnte gezeigt werden, dass sich das HMRF-UNet sehr gut für das Vortrainieren von neuronalen Netzen eignet und damit die benötigte Anzahl an annotierten Groundtruth-Segmentierungen deutlich reduziert werden konnte. Das dritte Segmentierungsmodell wurde speziell für die Lungensegmentierung entwickelt. Durch die Verwendung eines anwendungsspezifischen Algorithmus konnte auf die Besonderheiten der Lungensegmentierung eingegangen werden. Insbesondere das Exkludieren der Trachea stellte sich als wichtige Maßnahme heraus, um die Segmentierung nachhaltig zu verbessern. Die entwickelte algorithmische Lungensegmentierung erzielte dabei bessere und insbesondere auch konstantere Ergebnisse für die Lungensegmentierung, als die beiden HMRF-basierten Methoden.

Die in dieser Arbeit entwickelten flexiblen Segmentierungsmethoden erlauben einen breiten Einsatz in den unterschiedlichsten Anwendungsgebieten ohne größere Anpassungen. Dabei erfüllen die Methoden das zu Beginn gesetzte Ziel einer flexiblen Anwendbarkeit auf unterschiedlichste mehrdimensionale Bilddaten. Dies wurde anhand verschiedener Anwendungsfälle aus Radiologie und Materialwissenschaften gezeigt. Die entwickelten Segmentierungsmethoden können die Entwicklung neuer Materialien beschleunigen und dabei helfen, noch schneller Materialien für zentrale Themen der klimafreundlichen Verkehrswende wie den

Leichtbau oder neue Batterietechniken zu entwickeln. Im Bereich der Radiologie könnten die Zeiteinsparungen neben der schnelleren Erforschung von Erkrankungen auch zu mehr Kapazitäten für die Entwicklung von Quantifizierungs- und Auswertungsmethoden und auf lange Sicht zu einer besseren Erforschung von Lungenerkrankungen führen.

Das zweite wesentliche Ziel dieser Arbeit lag in der Entwicklung einer robusten Methode zur vollautomatischen Perfusionsquantifizierung in DCE-MRT Bildern der Lunge. Hierzu wurde eine bestehende halbautomatische Methode, die bislang nur in monozentrischen Studien verwendet wurde, weiterentwickelt, sodass diese auch vollautomatisch und robust für multizentrische Studien eingesetzt werden kann. Die bestehende algorithmische Lungensegmentierung konnte durch eine optimierte Tracheasegmentierung verbessert werden. Die Segmentierungsergebnisse wurden anhand von registrierten CT-Lungenmasken validiert und es wurde gezeigt, dass die algorithmische Segmentierungsmethode zu besseren oder ähnlich guten Ergebnissen führt wie Referenzmethoden. Neben klassischen Segmentierungsmetriken wurde mit der EigenLung-Distanz eine eigene Metrik entwickelt, die sich als eine sinnvolle Ergänzung zu bestehenden Segmentierungsmetriken herausstellte.

Für das automatische Herausfiltern von veratmeten Zeitpunkten vor und nach der Boluspassage in der DCE-MRT Messung wurde ein einheitliches Vorgehen entworfen. Hierzu wurde ein Modell zur automatischen Atmungserkennung in DCE-MRT-Bildern der Lunge basierend auf einem CNN entwickelt und mit zwei entwickelten Algorithmen zur Boluscharakterisierung kombiniert. Somit konnte eine vollständige Automatisierung der Pipeline erreicht werden. Zusätzlich wurde ein DCE-MRT Qualitätsscore konzipiert, der die Qualität der DCE-MRT Bilder misst und damit eine Aussage darüber liefert, wie valide die quantitativen Perfusionswerte sind. Es konnte insgesamt gezeigt werden, dass der DCE-MRT Qualitätsscore in der Lage ist, automatisch Ausreißer zu identifizieren.

Die entwickelte Perfusionsquantifizierungs-Pipeline wurde abschließend anhand von zwei verschiedenen Anwendungsfällen getestet. Für den ersten Anwendungsfall, in dem der Zusammenhang zwischen Perfusionsdefekten und Perfusionsverzögerungen in CF untersucht werden sollte, wurden zwei zusätzliche quantitative Metriken entwickelt, die die Perfusionsverzögerung und die arterielle Korrelation quantifizieren. Es zeigte sich, dass damit eine verstärkte Perfusionsverzögerung innerhalb der Perfusionsdefekte nachgewiesen werden kann. Zusätzlich schien die arterielle Korrelation eine vielversprechende Metrik zu sein, um frühzeitig die Anwesenheit von dilatierten Bronchialarterien zu detektieren. Beim zweiten Anwendungsfall ging es um die Perfusionsquantifizierung in einer multizentrischen Studie mit COPD-PatientInnen. Hierbei konnten Einflüsse der MRT-Scanner auf die quantitativen Werte beobachtet werden. Dennoch wurden für alle Scanner hohe Korrelationen zwischen den berechneten quantitativen Perfusionswerten und klinischen Parametern und QCT-Werten erreicht, was das Potenzial der entwickelten Methode aufzeigt. Durch die Entwicklung einer vollautomatischen und robusten Perfusionsquantifizierungs-Pipeline wurde nun der Weg für einen breiten Einsatz der Methode geebnet. Hierdurch wird eine Verlaufs- und Therapiekontrolle von an COPD- und CF-Erkrankten möglich, die wichtige Erkenntnisse für die weitere Erforschung der beiden Erkrankungen, aber auch für die Erforschung weiterer Lungenerkrankungen liefern kann.

Für die nachhaltige Dokumentation der Forschungsergebnisse wurde ein Konzept zur Verwendung von FDM für die medizinische Bildverarbeitung entwickelt. Hierzu wurde Kadi4Mat verwendet. Für die Arbeit mit medizinischen Bildern wurde eine Methode entwickelt, die die Bilder automatisch im DICOM-Format einliest und diese inklusive der zugehörigen Metadaten strukturiert in Kadi4Mat ablegt. Des Weiteren wurde ein Konzept entwickelt, welches den automatischen Upload eines lokal durchgeführten ML-Projektes ermöglicht. Dabei wurden alle Bestandteile eines ML-Projektes, basierend auf einer Ontologie, in Kadi4Mat exportiert und deren Zusammenhänge mithilfe von Verlinkungen dokumentiert. Damit konnte eine einheitliche Abspeicherung von ML-Projekten erreicht werden. Abschließend wurde gezeigt, dass auch die Ergebnisse algorithmischer Methoden strukturiert in Kadi4Mat abgelegt werden können. Dies wurde anhand der

Perfusionsquantifizierungsergebnisse der multizentrischen Studie demonstriert, die in einzelnen Records mit entsprechenden Verlinkungen gespeichert wurden, wodurch ein großer Wissensgraph erstellt werden konnte. Durch die drei entwickelten FDM-Konzepte konnten die zu Beginn gesteckten Ziele nach einem strukturierten Vorgehen zur Dokumentation medizinischer Forschungsdaten und Forschungsmethoden erfüllt werden. Die entwickelten Konzepte erleichtern zukünftigen Forschenden die Weiterverwendung der entwickelten Methoden und Modelle, indem die einzelnen Zusammenhänge und Ergebnisse klar dokumentiert sind.

Die in dieser Arbeit entwickelten Methoden und Konzepte zur flexiblen Segmentierung, zur robusten automatischen Perfusionsquantifizierung und zur nachhaltigen Dokumentation von medizinischen Forschungsdaten fügen sich perfekt in das entworfene Konzept des informationsgetriebenen Bildanalysezyklus ein. Dabei vergrößern diese nicht nur das Repertoire der zur Verfügung stehenden Bildverarbeitungsmethoden, sondern sie können auch als Blaupause für ähnliche Entwicklungsprojekte in anderen Anwendungsbereichen der Bildverarbeitung genutzt werden.

7.2 Ausblick

In dieser Arbeit konnte gezeigt werden, wie sich flexible und robuste Methoden für die Bildverarbeitung entwickeln lassen und wie diese insbesondere in interdisziplinären Projekten im Bereich der Materialwissenschaften und der Radiologie eingesetzt werden können. Diese Arbeit bietet somit die Grundlage für viele weiterführende Forschungsthemen, die nachfolgend erläutert werden.

Im Blick auf die entwickelte EHOHMRF-Methode konnten sehr lange Segmentierungszeiten, insbesondere für 3D-Bilder beobachtet werden. Um dies zu beheben, könnte eine weitere Optimierung des Codes durch eine stärkere Parallelisierung stattfinden. Eine weitere Idee wäre, den EHOHMRF-Algorithmus ebenso auf

verschiedenen Eingangsbildern zu trainieren, sodass die bestimmten Eigenschaften der Verteilungen der einzelnen Klassen verallgemeinert werden. Dies könnte dazu führen, dass deutlich weniger Iterationen für neue Bilder notwendig sind, da die Klassen bereits adäquat initialisiert wurden. Ebenso wären weiterführende Studien sinnvoll, die die Segmentierungsgüte in Abhängigkeit der Herdenanzahl und Herdengröße des EHO untersucht.

Für die EHOHMRF-Methode wurde eine Verbesserung der Segmentierungen durch Verwendung der Stompnorm-Verteilung gegenüber der Gaußschen Normalverteilung beobachtet. Die Stompnorm-Verteilung könnte daher auch als alternativer Verteilungs-Loss im HMRF-UNet getestet werden, um zu ermitteln, ob hierdurch noch bessere Segmentierungsergebnisse erzeugt werden können. Die Verwendung des HMRF-UNet zum Vortrainieren von neuronalen Netzen zeigte vielversprechende Ergebnisse. Daher könnte auch für die Lungen-MRT Bilder untersucht werden, wie stark die Anzahl an Trainingsdaten durch die Verwendung des HMRF-UNet Pretraining reduziert werden kann. Des Weiteren könnte es auch interessant sein zu untersuchen, ob die Kombination eines klassischen überwachten DICE-Loss mit dem unüberwachten HMRF-Loss zu Verbesserungen führt.

Das große Potenzial der quantitativen Perfusionsauswertung von DCE-MRT Daten mit der entwickelten Pipeline wurde in dieser Arbeit deutlich ersichtlich. Nichtsdestotrotz zeigte die Methode auch einige Limitationen, die behoben werden sollten. Einerseits wurde beobachtet, dass die Genauigkeit der Ergebnisse stark von der Atmung während der Messung abhängt. Aktuell gibt es in der Pipeline keine Möglichkeit, für die Atmung während der Boluspassage zu korrigieren. Insbesondere im Hinblick darauf, dass die Methode bei Erkrankungen verwendet wird, bei denen ein Atemanhalt entweder aufgrund des jungen Alters der untersuchten Personen oder durch eine starke Lungenschädigung nicht möglich ist, sollte eine Methode entwickelt werden, die eine Korrektur veratmeter Zeitpunkte ermöglicht. Klassische Registrierungsalgorithmen sind hierfür zwar die naheliegendste Lösung, jedoch besteht dort die Gefahr, dass hierbei die gemessenen Signale künstlich verändert werden und damit nicht mehr der Realität entsprechen.

Daher sollten hier Methoden zum Einsatz kommen, die eine Registrierung sowohl unter Berücksichtigung der Perfusionsphysik, als auch der Atemphysiologie durchführen.

Eine intensivere Auseinandersetzung mit Registrierungsalgorithmen könnte zudem auch hinsichtlich der Registrierung der morphologischen Bilder auf die DCE-MRT Bilder wichtig sein. Die entwickelte algorithmische Segmentierungsmethode führte zwar zu sehr guten Ergebnissen, jedoch wurde die gute Segmentierung teilweise durch eine schlechte Registrierung der beiden Bilder wieder zunichtegemacht. Daher wären zusätzliche Untersuchungen sinnvoll, die sich mit einer weiteren Optimierung der Registrierung befassen.

Bei der quantitativen Auswertung der COSYCONET-Studie konnte festgestellt werden, dass die verwendeten MRT-Scanner vermutlich einen Einfluss auf die Quantifizierungsergebnisse besitzen. Daher sollte in einer Folgestudie untersucht werden, ob eine Korrektur für Scanner-Einflüsse während der quantitativen Perfusionsauswertung geschehen kann. Eine solche Korrektur wäre dank der algorithmischen Bauweise der hier dargestellten Methode gut umsetzbar. Jedoch konnte der Einfluss des MRT-Scanners nicht definitiv bewiesen werden, da die verschiedenen MRT-Scanner für unterschiedliche PatientInnen verwendet wurden. Daher wäre eine Untersuchung sinnvoll, bei der die unterschiedlichen Scanner innerhalb kurzer Zeit für die selben PatientInnen verwendet werden, sofern dies praktikabel und ethisch vertretbar ist. Alternativ könnte auch die Messung eines Perfusions-Phantoms an verschiedenen Scannern eine Möglichkeit sein, um dieses Problem näher zu untersuchen.

Da es bisher kein MRT-Phantom gibt, welches eine physiologische Repräsentation der Perfusionsprozesse innerhalb der menschlichen Lunge simuliert, wäre die Entwicklung eines solchen Phantoms für das Forschungsgebiet der Perfusionsquantifizierung von großem Interesse. Denn alle Methoden zur Perfusionsquantifizierung besitzen das Problem, dass keine Groundtruth-Daten vorhanden sind. Daher ist bislang immer nur eine Validierung anhand von alternativen Quantifizierungsmethoden, klinischen Parametern oder QCT-Werten möglich. Diese

ermitteln jedoch meist nicht exakt die selben physiologischen Prozesse und sind daher nur eingeschränkt für eine robuste Validierung geeignet.

Die entwickelten Workflows zur Integration von medizinischen Bildern in Kadi4Mat sind ein wichtiger Schritt zur Verwendung von Kadi4Mat in der radiologischen Forschung. Um die Anwendbarkeit weiter zu verbessern, könnten Bildbetrachter für multidimensionale Bilddaten in Kadi4Mat integriert werden, die eine Betrachtung der Bilder innerhalb von KadiWeb ermöglichen. Zusätzlich wäre die Entwicklung von spezifischen Dashboards sinnvoll, um einen Überblick über die gespeicherten medizinischen Daten zu erhalten und beispielsweise auf einen Blick sehen zu können, wie viele Personen mit einer spezifischen Erkrankung ein MRT erhalten haben.

Literaturverzeichnis

[1] WHO, *Global health estimates: Leading causes of death*, https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death.

[2] H. M. Tran, F.-J. Tsai, Y.-L. Lee, J.-H. Chang, L.-T. Chang, T.-Y. Chang, K. F. Chung, H.-P. Kuo, K.-Y. Lee, K.-J. Chuang und H.-C. Chuang, „The Impact of Air Pollution on Respiratory Diseases in an Era of Climate Change: A Review of the Current Evidence“, *Science of The Total Environment*, Jg. 898, S. 166 340, Nov. 2023. DOI: 10.1016/j.scitotenv.2023.166340.

[3] S. Ehrenberger, I. Dasgupta, M. Brost, L. Gebhardt und R. Seiffert, „Potentials of Light Electric Vehicles for Climate Protection by Substituting Passenger Car Trips“, *World Electric Vehicle Journal*, Jg. 13, Nr. 10, Okt. 2022. DOI: 10.3390/wevj13100183.

[4] J. Ralyté, „Requirements Definition for the Situational Method Engineering“, in *Engineering Information Systems in the Internet Context: IFIP TC8 / WG8.1 Working Conference on Engineering Information Systems in the Internet Context September 25–27, 2002, Kanazawa, Japan*, C. Rolland, S. Brinkkemper und M. Saeki, Hrsg., Boston, MA: Springer US, 2002, S. 127–152. DOI: 10.1007/978-0-387-35614-3_9.

[5] S. W. Aboelela, E. Larson, S. Bakken, O. Carrasquillo, A. Formicola, S. A. Glied, J. Haas und K. M. Gebbie, „Defining Interdisciplinary Research: Conclusions from a Critical Review of the Literature“, *Health*

Services Research, Jg. 42, Nr. 1p1, S. 329–346, 2007. DOI: 10.1111/j.1475-6773.2006.00621.x.

[6] D. Bösch, *Lunge Und Atemwege* (Springer-Lehrbuch), J. Steffel und T. Lüscher, Hrsg. Berlin, Heidelberg: Springer, 2014. DOI: 10.1007/978-3-642-28223-2.

[7] P. Lebiedz, „Anatomie und Physiologie des Atemsystems“, in *Neurologische Beatmungsmedizin: Auf der Intensivstation, in der Frührehabilitation, im Schlaflabor, zu Hause*, M. Groß, Hrsg., Berlin, Heidelberg: Springer, 2020, S. 15–25. DOI: 10.1007/978-3-662-59014-0_2.

[8] V. Subramaniyan, S. Fuloria, M. Sekar, S. Shanmugavelu, K. Vijeepallam, U. Kumari, K. Narain, D. U. Meenakshi, M. N. H. Maziz und N. K. Fuloria, „Introduction to Lung Disease“, in *Targeting Epigenetics in Inflammatory Lung Diseases*, G. Gupta, B. G. Oliver, K. Dua, M. K. Ali und P. Dave, Hrsg., Singapore: Springer Nature, 2023, S. 1–16. DOI: 10.1007/978-981-99-4780-5_1.

[9] Z. Syunyaeva, M. A. Mall und M. Stahl, „Mukoviszidose im Kindes- und im Erwachsenenalter“, *Die Innere Medizin*, Jg. 65, Nr. 6, S. 538–544, Juni 2024. DOI: 10.1007/s00108-024-01717-z.

[10] S. C. Bell, M. A. Mall, H. Gutierrez, M. Macek, S. Madge, J. C. Davies, P.-R. Burgel, E. Tullis, C. Castaños, C. Castellani, C. A. Byrnes, F. Cathcart, S. H. Chotirmall, R. Cosgriff, I. Eichler, I. Fajac, C. H. Goss, P. Drevinek, P. M. Farrell, A. M. Gravelle, T. Havermans, N. Mayer-Hamblett, N. Kashirskaya, E. Kerem, J. L. Mathew, E. F. McKone, L. Naehrlich, S. Z. Nasr, G. R. Oates, C. O'Neill, U. Pypops, K. S. Raraigh, S. M. Rowe, K. W. Southern, S. Sivam, A. L. Stephenson, M. Zampoli und F. Ratjen, „The future of cystic fibrosis care: A global perspective“, *The Lancet. Respiratory Medicine*, Jg. 8, Nr. 1, S. 65–124, Jan. 2020. DOI: 10.1016/S2213-2600(19)30337-6.

[11] M. A. Mall und D. Hartl, „CFTR: Cystic fibrosis and beyond“, *European Respiratory Journal*, Jg. 44, Nr. 4, S. 1042–1054, Sep. 2014. DOI: 10.1183/09031936.00228013.

[12] M. A. Mall, „Role of cilia, mucus, and airway surface liquid in mucociliary dysfunction: Lessons from mouse models“, *Journal of Aerosol Medicine and Pulmonary Drug Delivery*, Jg. 21, Nr. 1, S. 13–24, März 2008. DOI: 10.1089/jamp.2007.0659.

[13] P. A. Flume, J. R. Yankaskas, M. Ebeling, T. Hulsey und L. L. Clark, „Massive Hemoptysis in Cystic Fibrosis“, *Chest*, Jg. 128, Nr. 2, S. 729–738, Aug. 2005. DOI: 10.1378/chest.128.2.729.

[14] M. O. Wielpütz, O. von Stackelberg, M. Stahl, B. J. Jobst, M. Eichinger, M. U. Puderbach, L. Nährlich, S. Barth, C. Schneider, M. V. Kopp, I. Ricklefs, M. Buchholz, B. Tümmler, C. Dopfer, J. Vogel-Claussen, H.-U. Kauczor und M. A. Mall, „Multicentre Standardisation of Chest MRI as Radiation-Free Outcome Measure of Lung Disease in Young Children with Cystic Fibrosis“, *Journal of Cystic Fibrosis*, Jg. 17, Nr. 4, S. 518–527, Juli 2018. DOI: 10.1016/j.jcf.2018.05.003.

[15] M. M. de Oca, R. Perez-Padilla, B. Celli, S. D. Aaron, F. C. Wehrmeister, A. F. S. Amaral, D. Mannino, J. Zheng, S. Salvi, D. Obaseki, A. S. Buist und A. Menezes, „The global burden of COPD: Epidemiology and effect of prevention strategies“, *The Lancet Respiratory Medicine*, Jg. 13, Nr. 8, S. 709–724, Aug. 2025. DOI: 10.1016/S2213-2600(24)00339-4.

[16] A. Agustí, B. R. Celli, G. J. Criner, D. Halpin, A. Anzueto, P. Barnes, J. Bourbeau, M. K. Han, F. J. Martinez, M. Montes de Oca, K. Mortimer, A. Papi, I. Pavord, N. Roche, S. Salvi, D. D. Sin, D. Singh, R. Stockley, M. V. López Varela, J. A. Wedzicha und C. F. Vogelmeier, „Global Initiative for Chronic Obstructive Lung Disease 2023 Report: GOLD Executive Summary“, *American Journal of Respiratory and Critical Care Medicine*, Jg. 207, Nr. 7, S. 819–837, Apr. 2023. DOI: 10.1164/rccm.202301-0106PP.

[17] N. Diab, A. S. Gershon, D. D. Sin, W. C. Tan, J. Bourbeau, L.-P. Boulet und S. D. Aaron, „Underdiagnosis and Overdiagnosis of Chronic Obstructive Pulmonary Disease“, *American Journal of Respiratory and Critical Care Medicine*, Jg. 198, Nr. 9, S. 1130–1139, Nov. 2018. DOI: 10.1164/rccm.201804-0621CI.

[18] E. A. Hoffman, D. A. Lynch, R. G. Barr, E. J. van Beek, G. Parraga und I. Investigators, „Pulmonary CT and MRI phenotypes that help explain chronic pulmonary obstruction disease pathophysiology and outcomes“, *Journal of Magnetic Resonance Imaging*, Jg. 43, Nr. 3, S. 544–557, 2016. DOI: 10.1002/jmri.25010.

[19] J. Conder, C. Marino, P. Novák und C. Villevieille, „Do imaging techniques add real value to the development of better post-Li-ion batteries?“, *Journal of Materials Chemistry A*, Jg. 6, Nr. 8, S. 3304–3327, Feb. 2018. DOI: 10.1039/C7TA10622J.

[20] A. A. Umar und S. M. Atabo, „A review of imaging techniques in scientific research/clinical diagnosis“, *MOJ Anatomy & Physiology*, Jg. Volume 6, Nr. Issue 5, Okt. 2019. DOI: 10.15406/mojap.2019.06.00269.

[21] J. A. Seibert und J. M. Boone, „X-Ray Imaging Physics for Nuclear Medicine Technologists. Part 2: X-Ray Interactions and Image Formation“, *Journal of Nuclear Medicine Technology*, Jg. 33, Nr. 1, S. 3–18, März 2005, ISSN: 0091-4916, 1535-5675.

[22] A. Thompson und R. Leach, „Introduction to Industrial X-ray Computed Tomography“, in *Industrial X-Ray Computed Tomography*, S. Carmignato, W. Dewulf und R. Leach, Hrsg., Cham: Springer International Publishing, 2018, S. 1–23. DOI: 10.1007/978-3-319-59573-3_1.

[23] K. J. Treb, A. O. El Sadaney, A. Ferrero, F. I. Baffour, S. Leng, L. Yu, P. S. Rajiah, J. G. Fletcher, C. H. McCollough und K. Rajendran, „Quantitative CT Imaging: Where Are We, and What Is Missing?“,

British Journal of Radiology, Jg. 98, Nr. 1175, S. 1834–1842, Nov. 2025. DOI: 10.1093/bjr/tqaf075.

[24] H. Villarraga-Gómez, E. L. Herazo und S. T. Smith, „X-Ray Computed Tomography: From Medical Imaging to Dimensional Metrology“, *Precision Engineering*, Jg. 60, S. 544–569, Nov. 2019. DOI: 10.1016/j.precisioneng.2019.06.007.

[25] S. D. Rawson, J. Maksimcuka, P. J. Withers und S. H. Cartmell, „X-ray computed tomography in life sciences“, *BMC Biology*, Jg. 18, Nr. 1, S. 21, Feb. 2020. DOI: 10.1186/s12915-020-0753-2.

[26] C. J. Galbán, T. L. Chenevert, C. R. Meyer, C. Tsien, T. S. Lawrence, D. A. Hamstra, L. Junck, P. C. Sundgren, T. D. Johnson, D. J. Ross, A. Rehemtulla und B. D. Ross, „The Parametric Response Map: An Imaging Biomarker for Early Cancer Treatment Outcome“, *Nature Medicine*, Jg. 15, Nr. 5, S. 572–576, Mai 2009. DOI: 10.1038/nm.1919.

[27] C. J. Galbán, M. K. Han, J. L. Boes, K. A. Chughtai, C. R. Meyer, T. D. Johnson, S. Galbán, A. Rehemtulla, E. A. Kazerooni, F. J. Martinez und B. D. Ross, „Computed tomography–based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression“, *Nature Medicine*, Jg. 18, Nr. 11, S. 1711–1715, Nov. 2012. DOI: 10.1038/nm.2971.

[28] M. O. Wielpütz, M. Eichinger, O. Weinheimer, S. Ley, M. A. Mall, M. Wiebel, A. Bischoff, H.-U. Kauczor, C. P. Heussel und M. Puderbach, „Automatic Airway Analysis on Multidetector Computed Tomography in Cystic Fibrosis: Correlation With Pulmonary Function Testing“, *Journal of Thoracic Imaging*, Jg. 28, Nr. 2, S. 104, März 2013. DOI: 10.1097/RTI.0b013e3182765785.

[29] P. Konietzke, O. Weinheimer, M. O. Wielpütz, D. Savage, T. Ziyeh, C. Tu, B. Newman, C. J. Galbán, M. A. Mall, H.-U. Kauczor und T. E. Robinson, „Validation of Automated Lobe Segmentation on Paired Inspiratory-Expiratory Chest CT in 8-14 Year-Old Children with Cystic

Fibrosis“, *PLoS ONE*, Jg. 13, Nr. 4, e0194557, Apr. 2018. DOI: 10.1371/journal.pone.0194557.

[30] P. Konietzke, C. Brunner, M. Konietzke, W. L. Wagner, O. Weinheimer, C. P. Heußel, F. J. F. Herth, F. Trudzinski, H.-U. Kauczor und M. O. Wielpütz, „GOLD stage-specific phenotyping of emphysema and airway disease using quantitative computed tomography“, *Frontiers in Medicine*, Jg. 10, S. 1 184 784, 2023. DOI: 10.3389/fmed.2023.1184784.

[31] D. Weishaupt, V. D. Köchli und B. Marincek, *Wie funktioniert MRI?: Eine Einführung in Physik und Funktionsweise der Magnetresonanzbildgebung*. Berlin, Heidelberg: Springer, 2014. DOI: 10.1007/978-3-642-41616-3.

[32] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer und A. Haase, „Generalized autocalibrating partially parallel acquisitions (GRAPPA)“, *Magnetic Resonance in Medicine*, Jg. 47, Nr. 6, S. 1202–1210, 2002. DOI: 10.1002/mrm.10171.

[33] J. J. Van Vaals, M. E. Brummer, W. Thomas Dixon, H. H. Tuithof, H. Engels, R. C. Nelson, B. M. Gerety, J. L. Chezmar und J. A. Den Boer, „“Keyhole” method for accelerating imaging of contrast agent uptake“, *Journal of Magnetic Resonance Imaging*, Jg. 3, Nr. 4, S. 671–675, Juli 1993. DOI: 10.1002/jmri.1880030419.

[34] F. R. Korosec, R. Frayne, T. M. Grist und C. A. Mistretta, „Time-resolved contrast-enhanced 3D MR angiography“, *Magnetic Resonance in Medicine*, Jg. 36, Nr. 3, S. 345–351, 1996. DOI: 10.1002/mrm.1910360304.

[35] T. M. Grist, C. A. Mistretta, C. M. Strother und P. A. Turski, „Time-resolved angiography: Past, present, and future“, *Journal of Magnetic Resonance Imaging*, Jg. 36, Nr. 6, S. 1273–1286, 2012. DOI: 10.1002/jmri.23646.

[36] M. Wielpütz und H.-U. Kauczor, „MRI of the lung: State of the art“, *Diagnostic and Interventional Radiology*, Jg. 18, Nr. 4, S. 344–353, Juli 2012. DOI: 10.4261/1305-3825.DIR.5365-11.0.

[37] J. M. Wild, H. Marshall, M. Bock, L. R. Schad, P. M. Jakob, M. Puderbach, F. Molinari, E. J. R. Van Beek und J. Biederer, „MRI of the lung (1/3): Methods“, *Insights into Imaging*, Jg. 3, Nr. 4, S. 345–353, Aug. 2012. DOI: 10.1007/s13244-012-0176-x.

[38] J. Biederer, M. Beer, W. Hirsch, J. Wild, M. Fabel, M. Puderbach und E. J. R. Van Beek, „MRI of the Lung (2/3). Why … When … How?“, *Insights into Imaging*, Jg. 3, Nr. 4, S. 355–371, Feb. 2012. DOI: 10.1007/s13244-011-0146-8.

[39] U. S. v. Euler und G. Liljestrand, „Observations on the Pulmonary Arterial Blood Pressure in the Cat“, *Acta Physiologica Scandinavica*, Jg. 12, Nr. 4, S. 301–320, 1946. DOI: 10.1111/j.1748-1716.1946.tb00389.x.

[40] S. M. Triphan, G. Bauman, P. Konietzke, M. Konietzke, M. O. Wielpütz und t. I. W. f. P. F. Imaging (IWPFI), „Magnetic Resonance Imaging of Lung Perfusion“, *Journal of Magnetic Resonance Imaging*, Jg. 59, Nr. 3, S. 784–796, 2024. DOI: 10.1002/jmri.28912.

[41] A. Voskrebenzev, M. Gutberlet, F. Klimeš, T. F. Kaireit, C. Schönfeld, A. Rotärmel, F. Wacker und J. Vogel-Claussen, „Feasibility of quantitative regional ventilation and perfusion mapping with phase-resolved functional lung (PREFUL) MRI in healthy volunteers and COPD, CTEPH, and CF patients“, *Magnetic Resonance in Medicine*, Jg. 79, Nr. 4, S. 2306–2314, 2018. DOI: 10.1002/mrm.26893.

[42] G. Bauman und O. Bieri, „Matrix pencil decomposition of time-resolved proton MRI for robust and improved assessment of pulmonary ventilation and perfusion“, *Magnetic Resonance in Medicine*, Jg. 77, Nr. 1, S. 336–342, 2017. DOI: 10.1002/mrm.26096.

[43] J. Biederer, C. P. Heussel, M. Puderbach und M. O. Wielpuetz, „Functional Magnetic Resonance Imaging of the Lung“, *Seminars in Respiratory and Critical Care Medicine*, Jg. 35, Nr. 01, S. 074–082, Feb. 2014. DOI: 10.1055/s-0033-1363453.

[44] S. Ley und J. Ley-Zaporozhan, „Pulmonary perfusion imaging using MRI: Clinical application“, *Insights into Imaging*, Jg. 3, Nr. 1, S. 61–71, Feb. 2012. DOI: 10.1007/s13244-011-0140-1.

[45] M. Eichinger, D.-E. Optazaite, A. Kopp-Schneider, C. Hintze, J. Biederer, A. Niemann, M. A. Mall, M. O. Wielpütz, H.-U. Kauczor und M. Puderbach, „Morphologic and Functional Scoring of Cystic Fibrosis Lung Disease Using MRI“, *European Journal of Radiology*, Jg. 81, Nr. 6, S. 1321–1329, Juni 2012. DOI: 10.1016/j.ejrad.2011.02.045.

[46] M. O. Wielpütz, M. Eichinger, J. Biederer, S. Wege, M. Stahl, O. Sommerburg, M. A. Mall, H. U. Kauczor und M. Puderbach, „Imaging of Cystic Fibrosis Lung Disease and Clinical Interpretation“, *RoFo: Fortschritte Auf Dem Gebiete Der Rontgenstrahlen Und Der Nuklearmedizin*, Jg. 188, Nr. 9, S. 834–845, Sep. 2016. DOI: 10.1055/s-0042-104936.

[47] G. Bauman, M. Puderbach, T. Heimann, A. Kopp-Schneider, E. Fritzsching, M. A. Mall und M. Eichinger, „Validation of Fourier Decomposition MRI with Dynamic Contrast-Enhanced MRI Using Visual and Automated Scoring of Pulmonary Perfusion in Young Cystic Fibrosis Patients“, *European Journal of Radiology*, Jg. 82, Nr. 12, S. 2371–2377, Dez. 2013. DOI: 10.1016/j.ejrad.2013.08.018.

[48] T. F. Kaireit, A. Voskrebenzev, M. Gutberlet, J. Freise, B. Jobst, H.-U. Kauczor, T. Welte, F. Wacker und J. Vogel-Claussen, „Comparison of quantitative regional perfusion-weighted phase resolved functional lung (PREFUL) MRI with dynamic gadolinium-enhanced regional pulmonary perfusion MRI in COPD patients“, *Journal of Magnetic Resonance Imaging*, Jg. 49, Nr. 4, S. 1122–1132, 2019. DOI: 10.1002/jmri.26342.

[49] M. A. Neal, B. J. Pippard, A. J. Simpson und P. E. Thelwall, „Dynamic susceptibility contrast 19F-MRI of inhaled perfluoropropane: A novel approach to combined pulmonary ventilation and perfusion imaging“, *Magnetic Resonance in Medicine*, Jg. 83, Nr. 2, S. 452–461, 2020. DOI: 10.1002/mrm.27933.

[50] T. Nakagawa, H. Sakuma, S. Murashima, N. Ishida, K. Matsumura und K. Takeda, „Pulmonary ventilation-perfusion MR imaging in clinical patients“, *Journal of Magnetic Resonance Imaging*, Jg. 14, Nr. 4, S. 419–424, 2001. DOI: 10.1002/jmri.1202.

[51] P. Meier und K. L. Zierler, „On the Theory of the Indicator-Dilution Method for Measurement of Blood Flow and Volume“, *Journal of Applied Physiology*, Jg. 6, Nr. 12, S. 731–744, Juni 1954. DOI: 10.1152/jappl.1954.6.12.731.

[52] S. Sourbron, M. Dujardin, S. Makkat und R. Luypaert, „Pixel-by-pixel deconvolution of bolus-tracking data: Optimization and implementation“, *Physics in Medicine & Biology*, Jg. 52, Nr. 2, S. 429, Dez. 2006. DOI: 10.1088/0031-9155/52/2/009.

[53] M. Ingrisch, O. Dietrich, U. I. Attenberger, K. Nikolaou, S. Sourbron, M. F. Reiser und C. Fink, „Quantitative Pulmonary Perfusion Magnetic Resonance Imaging: Influence of Temporal Resolution and Signal-to-Noise Ratio“, *Investigative Radiology*, Jg. 45, Nr. 1, S. 7, Jan. 2010. DOI: 10.1097/RLI.0b013e3181bc2d0c.

[54] S. R. Hopkins, M. O. Wielpütz und H.-U. Kauczor, „Imaging Lung Perfusion“, *Journal of Applied Physiology*, Jg. 113, Nr. 2, S. 328–339, Juli 2012. DOI: 10.1152/japplphysiol.00320.2012.

[55] Y. Ohno, H. Hatabu, K. Murase, T. Higashino, H. Kawamitsu, H. Watanabe, D. Takenaka, M. Fujii und K. Sugimura, „Quantitative assessment of regional pulmonary perfusion in the entire lung using three-dimensional ultrafast dynamic contrast-enhanced magnetic resonance imaging: Preliminary experience in 40 subjects“, *Journal of*

Magnetic Resonance Imaging, Jg. 20, Nr. 3, S. 353–365, 2004. DOI: 10.1002/jmri.20137.

[56] M. Konietzke, S. M. F. Triphan, M. Eichinger, S. Bossert, H. Heller, S. Wege, R. Eberhardt, M. U. Puderbach, H.-U. Kauczor, G. Heußel, C. P. Heußel, F. Risse und M. O. Wielpütz, „Unsupervised clustering algorithms improve the reproducibility of dynamic contrast-enhanced magnetic resonance imaging pulmonary perfusion quantification in muco-obstructive lung diseases“, *Frontiers in Medicine*, Jg. 9, Okt. 2022. DOI: 10.3389/fmed.2022.1022981.

[57] M. Schiwek, S. M. F. Triphan, J. Biederer, O. Weinheimer, M. Eichinger, C. F. Vogelmeier, R. A. Jörres, H.-U. Kauczor, C. P. Heußel, P. Konietzke, O. von Stackelberg, F. Risse, B. J. Jobst, M. O. Wielpütz und on behalf of the COSYCONET study group, „Quantification of pulmonary perfusion abnormalities using DCE-MRI in COPD: Comparison with quantitative CT and pulmonary function“, *European Radiology*, Jg. 32, Nr. 3, S. 1879–1890, März 2022. DOI: 10.1007/s00330-021-08229-6.

[58] I. Goodfellow, Y. Bengio und A. Courville, *Deep Learning* (Adaptive Computation and Machine Learning). Cambridge, Massachusetts: The MIT Press, 2016. DOI: 10.1007/s10710-017-9314-z.

[59] S. Mirjalili, *Evolutionary Algorithms and Neural Networks* (Studies in Computational Intelligence). Cham: Springer International Publishing, 2019, Bd. 780. DOI: 10.1007/978-3-319-93025-1.

[60] D. Dasgupta und Z. Michalewicz, Hrsg., *Evolutionary Algorithms in Engineering Applications*. Berlin, Heidelberg: Springer, 1997. DOI: 10.1007/978-3-662-03423-1.

[61] S. Katoch, S. S. Chauhan und V. Kumar, „A review on genetic algorithm: Past, present, and future“, *Multimedia Tools and Applications*, Jg. 80, Nr. 5, S. 8091–8126, Feb. 2021. DOI: 10.1007/s11042-020-10139-6.

[62] S. R. Jino Ramson, K. Lova Raju, S. Vishnu und T. Anagnostopoulos, „Nature Inspired Optimization Techniques for Image Processing—A Short Review“, in *Nature Inspired Optimization Techniques for Image*

Processing Applications, J. Hemanth und V. E. Balas, Hrsg., Cham: Springer International Publishing, 2019, S. 113–145. DOI: 10.1007/978-3-319-96002-9_5.

[63] R. Eberhart und J. Kennedy, „A New Optimizer Using Particle Swarm Theory“, in *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, Okt. 1995, S. 39–43. DOI: 10.1109/MHS.1995.494215.

[64] M. Dorigo, M. Birattari und T. Stutzle, „Ant Colony Optimization“, *IEEE Computational Intelligence Magazine*, Jg. 1, Nr. 4, S. 28–39, Nov. 2006. DOI: 10.1109/MCI.2006.329691.

[65] X.-S. Yang und S. Deb, „Cuckoo Search via Lévy Flights“, in *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC)*, Dez. 2009, S. 210–214. DOI: 10.1109/NABIC.2009.5393690.

[66] J. Li, H. Lei, A. H. Alavi und G.-G. Wang, „Elephant Herding Optimization: Variants, Hybrids, and Applications“, *Mathematics*, Jg. 8, Nr. 9, S. 1415, Sep. 2020. DOI: 10.3390/math8091415.

[67] G.-G. Wang, S. Deb und L. d. S. Coelho, „Elephant Herding Optimization“, in *2015 3rd International Symposium on Computational and Business Intelligence (ISCBI)*, Dez. 2015, S. 1–5. DOI: 10.1109/ISCBI.2015.8.

[68] Y. Duan, C. Liu, S. Li, X. Guo und C. Yang, „Gradient-based elephant herding optimization for cluster analysis“, *Applied Intelligence*, Jg. 52, Nr. 10, S. 11 606–11 637, Aug. 2022. DOI: 10.1007/s10489-021-03020-y.

[69] H. Tjelmeland und J. Besag, „Markov Random Fields with Higher-order Interactions“, *Scandinavian Journal of Statistics*, Jg. 25, Nr. 3, S. 415–433, 1998. DOI: 10.1111/1467-9469.00113.

[70] S. Geman und D. Geman, „Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images“, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Jg. PAMI-6, Nr. 6, S. 721–741, Nov. 1984. DOI: 10.1109/TPAMI.1984.4767596.

[71] Y. Zhang, M. Brady und S. Smith, „Segmentation of Brain MR Images through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm“, *IEEE Transactions on Medical Imaging*, Jg. 20, Nr. 1, S. 45–57, Jan. 2001. DOI: 10.1109/42.906424.

[72] E.-H. Guerrout, R. Mahiou, D. Michelucci, B. Randa und O. Assia, *Hidden markov random fields and cuckoo search method for medical image segmentation*, Mai 2020. DOI: 10.48550/arXiv.2005.09377.

[73] S. Ait-Aoudia, R. Mahiou und E. Guerrout, „Evaluation of Volumetric Medical Images Segmentation Using Hidden Markov Random Field Model“, in *2011 15th International Conference on Information Visualisation*, Juli 2011, S. 513–518. DOI: 10.1109/IV.2011.83.

[74] A. Banerjee und P. Maji, „A Spatially Constrained Probabilistic Model for Robust Image Segmentation“, *IEEE Transactions on Image Processing*, Jg. 29, S. 4898–4910, 2020. DOI: 10.1109/TIP.2020.2975717.

[75] S. Bauer, P. Benner, T. Bereau, V. Blum, M. Boley, C. Carbogno, C. R. A. Catlow, G. Dehm, S. Eibl, R. Ernstorfer, Á. Fekete, L. Foppa, P. Fratzl, C. Freysoldt, B. Gault, L. M. Ghiringhelli, S. K. Giri, A. Gladyshev, P. Goyal, J. Hattrick-Simpers, L. Kabalan, P. Karpov, M. S. Khorrami, C. T. Koch, S. Kokott, T. Kosch, I. Kowalec, K. Kremer, A. Leitherer, Y. Li, C. H. Liebscher, A. J. Logsdail, Z. Lu, F. Luong, A. Marek, F. Merz, J. R. Mianroodi, J. Neugebauer, Z. Pei, T. A. R. Purcell, D. Raabe, M. Rampp, M. Rossi, J.-M. Rost, J. Saal, U. Saalmann, K. N. Sasidhar, A. Saxena, L. Sbailò, M. Scheidgen, M. Schloz, D. F. Schmidt, S. Teshuva, A. Trunschke, Y. Wei, G. Weikum, R. P. Xian, Y. Yao, J. Yin, M. Zhao und M. Scheffler, „Roadmap on data-centric materials science“, *Modelling and Simulation in Materials Science and Engineering*, Jg. 32, Nr. 6, S. 063 301, Juli 2024. DOI: 10.1088/1361-651X/ad4d0d.

[76] W. Maass, J. Parsons, S. Purao, V. Storey und C. Woo, „Data-Driven Meets Theory-Driven Research in the Era of Big Data: Opportunities and Challenges for Information Systems Research“, *Journal of the*

Association for Information Systems, Jg. 19, Nr. 12, Dez. 2018. DOI: 10.17705/1jais.00526.

[77] N. Gordillo, E. Montseny und P. Sobrevilla, „State of the Art Survey on MRI Brain Tumor Segmentation“, *Magnetic Resonance Imaging*, Jg. 31, Nr. 8, S. 1426–1438, Okt. 2013. DOI: 10.1016/j.mri.2013.05.002.

[78] T. Ridler und S. Calvard, „PICTURE THRESHOLDING USING AN ITERATIVE SLECTION METHOD.“, *IEEE Transactions on Systems, Man and Cybernetics*, Jg. SMC-8, Nr. 8, S. 630–632, 1978. DOI: 10.1109/tsmc.1978.4310039.

[79] N. Otsu, „A Threshold Selection Method from Gray-Level Histograms“, *IEEE Transactions on Systems, Man, and Cybernetics*, Jg. 9, Nr. 1, S. 62–66, Jan. 1979. DOI: 10.1109/TSMC.1979.4310076.

[80] R. Adams und L. Bischof, „Seeded Region Growing“, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Jg. 16, Nr. 6, S. 641–647, Juni 1994. DOI: 10.1109/34.295913.

[81] S. Hojjatoleslami und J. Kittler, „Region Growing: A New Approach“, *IEEE Transactions on Image Processing*, Jg. 7, Nr. 7, S. 1079–1084, Juli 1998. DOI: 10.1109/83.701170.

[82] L. Bogunia, S. Buchen und K. Weinberg, „Microstructure characterization and stochastic modeling of open-cell foam based on μCT-image analysis“, *GAMM-Mitteilungen*, Jg. 45, Nr. 3-4, e202200018, 2022. DOI: 10.1002/gamm.202200018.

[83] A. Neumann und C. Lorenz, „Statistical Shape Model Based Segmentation of Medical Images“, *Computerized Medical Imaging and Graphics*, Jg. 22, Nr. 2, S. 133–143, März 1998. DOI: 10.1016/S0895-6111(98)00015-9.

[84] A. Kelemen, G. Szekely und G. Gerig, „Elastic Model-Based Segmentation of 3-D Neuroradiological Data Sets“, *IEEE Transactions on Medical Imaging*, Jg. 18, Nr. 10, S. 828–839, Okt. 1999. DOI: 10.1109/42.811260.

[85] T. Kohonen, „The Self-Organizing Map“, *Proceedings of the IEEE*, Jg. 78, Nr. 9, S. 1464–1480, Sep. 1990. DOI: 10.1109/5.58325.

[86] B. Panić, M. Borovinšek, M. Vesenjak, S. Oman und M. Nagode, „A Guide to Unsupervised Image Segmentation of mCT-scanned Cellular Metals with Mixture Modelling and Markov Random Fields“, *Materials & Design*, Jg. 239, S. 112 750, März 2024. DOI: 10.1016/j.matdes.2024.112750.

[87] O. Ronneberger, P. Fischer und T. Brox, *U-Net: Convolutional Networks for Biomedical Image Segmentation*, Mai 2015. DOI: 10.48550/arXiv.1505.04597.

[88] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen und K. H. Maier-Hein, „nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation“, *Nature Methods*, Jg. 18, Nr. 2, S. 203–211, Feb. 2021. DOI: 10.1038/s41592-020-01008-z.

[89] N. Soboleva und A. Mushnikov, „Improving the Accuracy of Semantic Segmentation of Carbides in the Microstructure of Composite Coatings by the Neural Network“, *Materials Today Communications*, Jg. 38, S. 108 276, März 2024. DOI: 10.1016/j.mtcomm.2024.108276.

[90] L. R. Dice, „Measures of the Amount of Ecologic Association Between Species“, *Ecology*, Jg. 26, Nr. 3, S. 297–302, 1945. DOI: 10.2307/1932409.

[91] M. Beneš und B. Zitová, „Performance evaluation of image segmentation algorithms on microscopic image data“, *Journal of Microscopy*, Jg. 257, Nr. 1, S. 65–85, 2015. DOI: 10.1111/jmi.12186.

[92] S. Nikolov, S. Blackwell, A. Zverovitch, R. Mendes, M. Livne, J. D. Fauw, Y. Patel, C. Meyer, H. Askham, B. Romera-Paredes, C. Kelly, A. Karthikesalingam, C. Chu, D. Carnell, C. Boon, D. D'Souza, S. A. Moinuddin, B. Garie, Y. McQuinlan, S. Ireland, K. Hampton, K. Fuller, H. Montgomery, G. Rees, M. Suleyman, T. Back, C. O. Hughes, J. R. Ledsam und O. Ronneberger, „Clinically Applicable Segmentation of Head and Neck Anatomy for Radiotherapy: Deep Learning Algorithm

Development and Validation Study“, *Journal of Medical Internet Research*, Jg. 23, Nr. 7, e26151, Juli 2021. DOI: 10.2196/26151.

[93] P. Iassonov, T. Gebrenegus und M. Tuller, „Segmentation of X-ray computed tomography images of porous materials: A crucial step for characterization and quantitative analysis of pore structures“, *Water Resources Research*, Jg. 45, Nr. 9, 2009. DOI: 10.1029/2009WR008087.

[94] L. Griem, A. Koeppe, A. Greß, T. Feser und B. Nestler, „Synthetic Training Data for CT Image Segmentation of Microstructures“, *Acta Materialia*, Jg. 296, S. 121 220, Sep. 2025. DOI: 10.1016/j.actamat.2025.121220.

[95] P. Gargiulo, T. Helgason, C. Ramon, H. J. jr und U. Carraro, „CT and MRI Assessment and Characterization Using Segmentation and 3D Modeling Techniques: Applications to Muscle, Bone and Brain“, *European Journal of Translational Myology*, Jg. 24, Nr. 1, S. 3298, März 2014. DOI: 10.4081/ejtm.2014.3298.

[96] A. August, J. Ettrich, M. Rölle, S. Schmid, M. Berghoff, M. Selzer und B. Nestler, „Prediction of Heat Conduction in Open-Cell Foams via the Diffuse Interface Representation of the Phase-Field Method“, *International Journal of Heat and Mass Transfer*, Jg. 84, S. 800–808, Mai 2015. DOI: 10.1016/j.ijheatmasstransfer.2015.01.052.

[97] E. Saillard, M. Gardegaront, A. Levillain, F. Bermond, D. Mitton, J.-B. Pialat, C. Confavreux, T. Grenier und H. Follet, „Finite element models with automatic computed tomography bone segmentation for failure load computation“, *Scientific Reports*, Jg. 14, Nr. 1, S. 16 576, Juli 2024. DOI: 10.1038/s41598-024-66934-w.

[98] K. M. Hasan, F. G. Moeller und P. A. Narayana, „DTI-based Segmentation and Quantification of Human Brain Lateral Ventricular CSF Volumetry and Mean Diffusivity: Validation, Age, Gender Effects and Biophysical Implications“, *Magnetic Resonance Imaging*, Jg. 32, Nr. 5, S. 405–412, Juni 2014. DOI: 10.1016/j.mri.2014.01.014.

[99] N. Brandt, L. Griem, C. Herrmann, E. Schoof, G. Tosato, Y. Zhao, P. Zschumme und M. Selzer, „Kadi4Mat: A Research Data Infrastructure for Materials Science“, *Data Science Journal*, Jg. 20, Nr. 1, Feb. 2021. DOI: 10.5334/dsj-2021-008.

[100] O. Furat, M. Wang, M. Neumann, L. Petrich, M. Weber, C. E. Krill und V. Schmidt, „Machine Learning Techniques for the Segmentation of Tomographic Image Data of Functional Materials“, *Frontiers in Materials*, Jg. 6, Juni 2019. DOI: 10.3389/fmats.2019.00145.

[101] S. Medghalchi, J. Kortmann, S.-H. Lee, E. Karimi, U. Kerzel und S. Korte-Kerzel, „Automated Segmentation of Large Image Datasets Using Artificial Intelligence for Microstructure Characterisation and Damage Analysis“, *Materials & Design*, Jg. 243, S. 113 031, Juli 2024. DOI: 10.1016/j.matdes.2024.113031.

[102] W. D. Romero, Y. Gutierrez, L. M. Tami-Pimiento, S. Torres-Bermudez, A. M. Meléndez und F. Martínez, „Automatic Pore Shape Characterization in Metal Foams Templated by Hydrogen Bubbles from a Deep Learning Strategy“, *Materials Today Communications*, Jg. 41, S. 110 937, Dez. 2024. DOI: 10.1016/j.mtcomm.2024.110937.

[103] G. Bo, H. Zhou, C. Wang, C. Zhang, C. Deng, D. Jiang, W. Li, G. Mao, Y. Sun, F. Jiang und Z. Peng, „Automatic Si Phase Extraction from Microscopic Images of Al-Si Alloys by Unsupervised Machine Learning and Supervised Deep Learning“, *Materials Today Communications*, Jg. 42, S. 111 468, Jan. 2025. DOI: 10.1016/j.mtcomm.2024.111468.

[104] L. A. Jara-Lugo, J. Caro-Gutierrez, F. F. Gonzalez-Navarro, M. A. Curiel-Alvarez, A. Armenta-Garcia und O. M. Perez-Landeros, „Semantic and Instance Segmentation Deep Learning Methods for Nanoparticles Detection“, *Materials Today Communications*, Jg. 45, S. 112 074, Apr. 2025. DOI: 10.1016/j.mtcomm.2025.112074.

[105] J. R. Astley, A. M. Biancardi, P. J. C. Hughes, H. Marshall, G. J. Collier, H.-F. Chan, L. C. Saunders, L. J. Smith, M. L. Brook, R. Thompson, S. Rowland-Jones, S. Skeoch, S. M. Bianchi, M. Q. Hatton,

N. M. Rahman, L.-P. Ho, C. E. Brightling, L. V. Wain, A. Singapuri, R. A. Evans, A. J. Moss, G. P. McCann, S. Neubauer, B. Raman, C-MORE/PHOSP-COVID Collaborative Group, J. M. Wild und B. A. Tahir, „Implementable Deep Learning for Multi-sequence Proton MRI Lung Segmentation: A Multi-center, Multi-vendor, and Multi-disease Study“, *Journal of magnetic resonance imaging: JMRI*, Jg. 58, Nr. 4, S. 1030–1044, Okt. 2023. DOI: 10.1002/jmri.28643.

[106] F. G. Ringwald, L. Wucherpfennig, N. Hagen, J. Mücke, S. Kaletta, M. Eichinger, M. Stahl, S. M. F. Triphan, P. Leutz-Schmidt, S. Gestewitz, S. Y. Graeber, H.-U. Kauczor, A. Alrajab, J.-P. Schenk, O. Sommerburg, M. A. Mall, P. Knaup, M. O. Wielpütz und U. Eisenmann, „Automated Lung Segmentation on Chest MRI in Children with Cystic Fibrosis“, *Frontiers in Medicine*, Jg. 11, S. 1 401 473, Nov. 2024. DOI: 10.3389/fmed.2024.1401473.

[107] B. Mairhörmann, A. Castelblanco, F. Häfner, V. Koliogiannis, L. Haist, D. Winter, A. Flemmer, H. Ehrhardt, S. Stöcklein, O. Dietrich, K. Förster, A. Hilgendorff und B. Schubert, „Automated MRI Lung Segmentation and 3D Morphologic Features for Quantification of Neonatal Lung Disease“, *Radiology: Artificial Intelligence*, Jg. 5, Nr. 6, e220239, Nov. 2023. DOI: 10.1148/ryai.220239.

[108] H. B. Winther, M. Gutberlet, C. Hundt, T. F. Kaireit, T. M. Alsady, B. Schmidt, F. Wacker, Y. Sun, S. Dettmer, S. K. Maschke, J. B. Hinrichs, S. Jambawalikar, M. R. Prince, R. G. Barr und J. Vogel-Claussen, „Deep semantic lung segmentation for tracking potential pulmonary perfusion biomarkers in chronic obstructive pulmonary disease (COPD): The multi-ethnic study of atherosclerosis COPD study“, *Journal of Magnetic Resonance Imaging*, Jg. 51, Nr. 2, S. 571–579, 2020. DOI: 10.1002/jmri.26853.

[109] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollar und R. Girshick, „Segment Anything“, in *Proceedings of the IEEE/CVF International*

Conference on Computer Vision, 2023, S. 4015–4026. DOI: 10.48550/arXiv.2304.02643.

[110] J. Ma, Y. He, F. Li, L. Han, C. You und B. Wang, „Segment anything in medical images“, *Nature Communications*, Jg. 15, Nr. 1, S. 654, Jan. 2024. DOI: 10.1038/s41467-024-44824-z.

[111] P. Kohlmann, J. Strehlow, B. Jobst, S. Krass, J.-M. Kuhnigk, A. Anjorin, O. Sedlaczek, S. Ley, H.-U. Kauczor und M. O. Wielpütz, „Automatic lung segmentation method for MRI-based lung perfusion studies of patients with chronic obstructive pulmonary disease“, *International Journal of Computer Assisted Radiology and Surgery*, Jg. 10, Nr. 4, S. 403–417, Apr. 2015. DOI: 10.1007/s11548-014-1090-0.

[112] T. Ivanovska, K. Hegenscheid, R. Laqua, J.-P. Kühn, S. Gläser, R. Ewert, N. Hosten, R. Puls und H. Völzke, „A Fast and Accurate Automatic Lung Segmentation and Volumetry Method for MR Data Used in Epidemiological Studies“, *Computerized Medical Imaging and Graphics*, Jg. 36, Nr. 4, S. 281–293, Juni 2012. DOI: 10.1016/j.compmedimag.2011.10.001.

[113] A. Ortiz, J. M. Gorriz, J. Ramirez und D. Salas-Gonzalez, „Improving MR Brain Image Segmentation Using Self-Organising Maps and Entropy-Gradient Clustering“, *Information Sciences*, Jg. 262, S. 117–136, März 2014. DOI: 10.1016/j.ins.2013.10.002.

[114] I. Aganj, M. G. Harisinghani, R. Weissleder und B. Fischl, „Unsupervised Medical Image Segmentation Based on the Local Center of Mass“, *Scientific Reports*, Jg. 8, Nr. 1, S. 13 012, Aug. 2018. DOI: 10.1038/s41598-018-31333-5.

[115] R. Cohn und E. Holm, „Unsupervised Machine Learning Via Transfer Learning and k-Means Clustering to Classify Materials Image Data“, *Integrating Materials and Manufacturing Innovation*, Jg. 10, Nr. 2, S. 231–244, Juni 2021. DOI: 10.1007/s40192-021-00205-8.

[116] H. Kim, J. Inoue und T. Kasuya, „Unsupervised microstructure segmentation by mimicking metallurgists' approach to pattern recognition“, *Scientific Reports*, Jg. 10, Nr. 1, S. 17 835, Okt. 2020. DOI: 10.1038/s41598-020-74935-8.

[117] J. Su, S. Liu und J. Song, „A Segmentation Method Based on HMRF for the Aided Diagnosis of Acute Myeloid Leukemia“, *Computer Methods and Programs in Biomedicine*, Jg. 152, S. 115–123, Dez. 2017. DOI: 10.1016/j.cmpb.2017.09.011.

[118] A. Banerjee und P. Maji, „Rough Sets and Stomped Normal Distribution for Simultaneous Segmentation and Bias Field Correction in Brain MR Images“, *Trans. Img. Proc.*, Jg. 24, Nr. 12, S. 5764–5776, Dez. 2015. DOI: 10.1109/TIP.2015.2488900.

[119] R. Mahiou, E.-H. Guerrout und M. E. Sannef, „Taguchi Design for Setting EHO Variants Parameters: Application in Brain Image Segmentation Using HMRF“, *SN Computer Science*, Jg. 4, Nr. 6, S. 794, Okt. 2023. DOI: 10.1007/s42979-023-02197-y.

[120] J. Hötzer, A. Reiter, H. Hierl, P. Steinmetz, M. Selzer und B. Nestler, „The parallel multi-physics phase-field framework PACE3D“, *Journal of Computational Science*, Jg. 26, S. 1–12, 2018. DOI: 10.1016/j.jocs.2018.02.011.

[121] E. Tuba, R. Capor-Hrosik, A. Alihodzic, R. Jovanovic und M. Tuba, „Chaotic Elephant Herding Optimization Algorithm“, in *2018 IEEE 16th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, Feb. 2018, S. 000 213–000 216. DOI: 10.1109/SAMI.2018.8324842.

[122] F. Bação, V. Lobo und M. Painho, „The Self-Organizing Map, the Geo-SOM, and Relevant Variants for Geosciences“, *Computers & Geosciences*, Geospatial Research in Europe: AGILE 2003, Jg. 31, Nr. 2, S. 155–163, März 2005. DOI: 10.1016/j.cageo.2004.06.013.

[123] J. Grolig, L. Griem, M. Selzer, H.-U. Kauczor, S. M. F. Triphan, B. Nestler und A. Koeppe, „Unsupervised Segmentation of Micro-CT Scans of Polyurethane Structures by Combining Hidden Markov Random Fields and a U-Net“, *Materials Today Communications*, Jg. 51, S. 114 817, Feb. 2026. DOI: 10.1016/j.mtcomm.2026.114817.

[124] F. Milletari, N. Navab und S.-A. Ahmadi, „V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation“, in *2016 Fourth International Conference on 3D Vision (3DV)*, Okt. 2016, S. 565–571. DOI: 10.1109/3DV.2016.79.

[125] D. Erhan, Y. Bengio, A. Courville, P.-A. Manzagol, P. Vincent und S. Bengio, „Why Does Unsupervised Pre-training Help Deep Learning?“, *J. Mach. Learn. Res.*, Jg. 11, S. 625–660, März 2010. DOI: 10.1145/1756006.1756025.

[126] M. Caron, P. Bojanowski, J. Mairal und A. Joulin, „Unsupervised Pre-Training of Image Features on Non-Curated Data“, in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, S. 2959–2968. DOI: 10.1109/ICCV.2019.00305.

[127] A. Kalapos und B. Gyires-Tóth, „Self-supervised Pretraining for 2D Medical Image Segmentation“, in *Computer Vision – ECCV 2022 Workshops*, L. Karlinsky, T. Michaeli und K. Nishino, Hrsg., Cham: Springer Nature Switzerland, 2023, S. 472–484. DOI: 10.1007/978-3-031-25082-8_31.

[128] Y. Nan, P. Tang, G. Zhang, C. Zeng, Z. Liu, Z. Gao, H. Zhang und G. Yang, „Unsupervised Tissue Segmentation via Deep Constrained Gaussian Network“, *IEEE Transactions on Medical Imaging*, Jg. 41, Nr. 12, S. 3799–3811, Dez. 2022. DOI: 10.1109/TMI.2022.3195123.

[129] M. Konietzke, „Quantification of perfusion abnormalities using dynamic contrast-enhanced magnetic resonance imaging in muco-obstructive lung diseases“, Dissertation, Heidelberg, 2023. DOI: 10.11588/heidok.00033300.

[130] Y. Sui, O. Afacan, C. Jaimes, A. Gholipour und S. K. Warfield, „Gradient-Guided Isotropic MRI Reconstruction From Anisotropic Acquisitions“, en, *IEEE Transactions on Computational Imaging*, Jg. 7, S. 1240–1253, 2021. DOI: 10.1109/TCI.2021.3128745.

[131] L. Østergaard, R. M. Weisskoff, D. A. Chesler, C. Gyldensted und B. R. Rosen, „High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis“, *Magnetic Resonance in Medicine*, Jg. 36, Nr. 5, S. 715–725, 1996. DOI: 10.1002/mrm.1910360510.

[132] L. Østergaard, A. G. Sorensen, K. K. Kwong, R. M. Weisskoff, C. Gyldensted und B. R. Rosen, „High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part II: Experimental comparison and preliminary results“, *Magnetic Resonance in Medicine*, Jg. 36, Nr. 5, S. 726–736, 1996. DOI: 10.1002/mrm.1910360511.

[133] F. Calamante, D. L. Thomas, G. S. Pell, J. Wiersma und R. Turner, „Measuring Cerebral Blood Flow Using Magnetic Resonance Imaging Techniques“, *Journal of Cerebral Blood Flow & Metabolism*, Jg. 19, Nr. 7, S. 701–735, Juli 1999. DOI: 10.1097/00004647-199907000-00001.

[134] K. Murase, M. Shinohara und Y. Yamazaki, „Accuracy of deconvolution analysis based on singular value decomposition for quantification of cerebral blood flow using dynamic susceptibility contrast-enhanced magnetic resonance imaging“, *Physics in Medicine & Biology*, Jg. 46, Nr. 12, S. 3147, Nov. 2001. DOI: 10.1088/0031-9155/46/12/306.

[135] A. N. Tikhonov, „Solution of Incorrectly Formulated Problems and the Regularization Method“, *Soviet Math. Dokl.*, Jg. 5, S. 1035–1038, 1963.

[136] P. C. Hansen, „Analysis of Discrete Ill-Posed Problems by Means of the L-Curve“, *SIAM Review*, Jg. 34, Nr. 4, S. 561–580, 1992. DOI: 10.1137/1034115.

[137] M. Salehi Ravesh, G. Brix, F. B. Laun, T. A. Kuder, M. Puderbach, J. Ley-Zaporozhan, S. Ley, A. Fieselmann, M. F. Herrmann, W. Schranz, W. Semmler und F. Risse, „Quantification of pulmonary microcirculation by dynamic contrast-enhanced magnetic resonance imaging: Comparison of four regularization methods“, *Magnetic Resonance in Medicine*, Jg. 69, Nr. 1, S. 188–199, 2013. DOI: 10.1002/mrm.24220.

[138] F. Calamante, „Arterial Input Function in Perfusion MRI: A Comprehensive Review“, *Progress in Nuclear Magnetic Resonance Spectroscopy*, Jg. 74, S. 1–32, Okt. 2013. DOI: 10.1016/j.pnmrs.2013.04.002.

[139] O. S. Pianykh, „Perfusion Linearity and Its Applications in Perfusion Algorithm Analysis“, *Computerized Medical Imaging and Graphics*, Jg. 36, Nr. 3, S. 204–214, Apr. 2012. DOI: 10.1016/j.compmedimag.2011.08.001.

[140] F. Calamante, D. G. Gadian und A. Connelly, „Quantification of bolus-tracking MRI: Improved characterization of the tissue residue function using Tikhonov regularization“, *Magnetic Resonance in Medicine*, Jg. 50, Nr. 6, S. 1237–1247, 2003. DOI: 10.1002/mrm.10643.

[141] H. Hatabu, E. Tadamura, D. L. Levin, Q. Chen, W. Li, D. Kim, P. V. Prasad und R. R. Edelman, „Quantitative assessment of pulmonary perfusion with dynamic contrast-enhanced MRI“, *Magnetic Resonance in Medicine*, Jg. 42, Nr. 6, S. 1033–1038, 1999. DOI: 10.1002/(SICI)1522-2594(199912)42:6<1033::AID-MRM7>3.0.CO;2-7.

[142] Y. Ohno, K. Murase, T. Higashino, M. Nogami, H. Koyama, D. Takenaka, H. Kawamitu, S. Matsumoto, H. Hatabu und K. Sugimura, „Assessment of bolus injection protocol with appropriate concentration for quantitative assessment of pulmonary perfusion by dynamic contrast-enhanced MR imaging“, *Journal of Magnetic Resonance Imaging*, Jg. 25, Nr. 1, S. 55–65, 2007. DOI: 10.1002/jmri.20790.

[143] M. Ingrisch, D. Maxien, F. Schwab, M. F. Reiser, K. Nikolaou und O. Dietrich, „Assessment of Pulmonary Perfusion With Breath-Hold and Free-Breathing Dynamic Contrast-Enhanced Magnetic Resonance Imaging: Quantification and Reproducibility“, *Investigative Radiology*, Jg. 49, Nr. 6, S. 382, Juni 2014. DOI: 10.1097/RLI.0000000000000020.

[144] K. Hueper, J. Vogel-Claussen, M. A. Parikh, J. H. M. Austin, D. A. Bluemke, J. Carr, J. Choi, T. A. Goldstein, A. S. Gomes, E. A. Hoffman, S. M. Kawut, J. Lima, E. D. Michos, W. S. Post, M. J. Po, M. R. Prince, K. Liu, D. Rabinowitz, J. Skrok, B. M. Smith, K. Watson, Y. Yin, A. M. Zambeli-Ljepovic und R. G. Barr, „Pulmonary Microvascular Blood Flow in Mild Chronic Obstructive Pulmonary Disease and Emphysema. The MESA COPD Study“, *American Journal of Respiratory and Critical Care Medicine*, Jg. 192, Nr. 5, S. 570–580, Sep. 2015. DOI: 10.1164/rccm.201411-2120OC.

[145] A. Ter-Karapetyan, S. M. F. Triphan, B. J. Jobst, A. F. Anjorin, J. Ley-Zaporozhan, S. Ley, O. Sedlaczek, J. Biederer, H.-U. Kauczor, P. M. Jakob und M. O. Wielpütz, „Towards quantitative perfusion MRI of the lung in COPD: The problem of short-term repeatability“, *PLOS ONE*, Jg. 13, Nr. 12, e0208587, Dez. 2018. DOI: 10.1371/journal.pone.0208587.

[146] J. Glandorf, F. Klimeš, L. Behrendt, A. Voskrebenzev, T. F. Kaireit, M. Gutberlet, F. Wacker und J. Vogel-Claussen, „Perfusion quantification using voxel-wise proton density and median signal decay in PREFUL MRI“, *Magnetic Resonance in Medicine*, Jg. 86, Nr. 3, S. 1482–1493, 2021. DOI: 10.1002/mrm.28787.

[147] L. A. Torres, K. E. Lee, G. P. Barton, A. D. Hahn, N. Sandbo, M. L. Schiebler und S. B. Fain, „Dynamic Contrast Enhanced MRI for the Evaluation of Lung Perfusion in Idiopathic Pulmonary Fibrosis“, *The European respiratory journal*, Jg. 60, Nr. 4, S. 2 102 058, Okt. 2022. DOI: 10.1183/13993003.02058-2021.

[148] L. Behrendt, L. J. Smith, A. Voskrebenzev, F. Klimeš, T. F. Kaireit, G. H. Pöhler, A. L. Kern, C. C. Gonzalez, A.-M. Dittrich, H. Marshall, K. Schütz, P. J. C. Hughes, P. Ciet, H. A. W. M. Tiddens, J. M. Wild und J. Vogel-Claussen, „A dual center and dual vendor comparison study of automated perfusion-weighted phase-resolved functional lung magnetic resonance imaging with dynamic contrast-enhanced magnetic resonance imaging in patients with cystic fibrosis“, *Pulmonary Circulation*, Jg. 12, Nr. 2, e12054, 2022. DOI: 10.1002/pul2.12054.

[149] T. Moher Alsady, A. Voskrebenzev, L. Behrendt, K. Olsson, C. P. Heußel, E. Gruenig, H. Gall, A. Ghofrani, F. Roller, S. Harth, H. Marshall, P. J. Hughes, J. Wild, A. J. Swift, D. G. Kiely, J. Behr, J. Dinkel, D. Beitzke, I. M. Lang, K. H. Schmidt, K. F. Kreitner, T. Frauenfelder, S. Ulrich, O. W. Hamer und J. Vogel-Claussen, „Multicenter Standardization of Phase-Resolved Functional Lung MRI in Patients With Suspected Chronic Thromboembolic Pulmonary Hypertension“, *Journal of Magnetic Resonance Imaging*, Jg. 59, Nr. 6, S. 1953–1964, 2024. DOI: 10.1002/jmri.28995.

[150] N. Woodhouse, J. M. Wild, M. N. Paley, S. Fichele, Z. Said, A. J. Swift und E. J. van Beek, „Combined helium-3/proton magnetic resonance imaging measurement of ventilated lung volumes in smokers compared to never-smokers“, *Journal of Magnetic Resonance Imaging*, Jg. 21, Nr. 4, S. 365–369, 2005. DOI: 10.1002/jmri.20290.

[151] D. P. I. Capaldi, R. L. Eddy, S. Svenningsen, F. Guo, J. S. H. Baxter, A. J. McLeod, P. Nair, D. G. McCormack, G. Parraga und For the Canadian Respiratory Research Network, „Free-Breathing Pulmonary MR Imaging to Quantify Regional Ventilation“, *Radiology*, Jg. 287, Nr. 2, S. 693–704, Mai 2018. DOI: 10.1148/radiol.2018171993.

[152] H. Hatabu, Y. Ohno, W. B. Gefter, G. Parraga, B. Madore, K. S. Lee, T. A. Altes, D. A. Lynch, J. R. Mayo, J. B. Seo, J. M. Wild, E. J. R. van Beek, M. L. Schiebler, H.-U. Kauczor und Fleischner Society, „Expanding Applications of Pulmonary MRI in the Clinical Evaluation of

Lung Disorders: Fleischner Society Position Paper“, *Radiology*, Jg. 297, Nr. 2, S. 286–301, Nov. 2020. DOI: 10.1148/radiol.2020201138.

[153] D.-J. Kroon, „Multimodality Non-Rigid Demon Algorithm Image Registration“, *Matlab File Exchange*, Feb. 2026, URL: https://de.mathworks.com/matlabcentral/fileexchange/21451-multimodality-non-rigid-demon-algorithm-image-registration.

[154] J.-P. Thirion, „Non-Rigid Matching Using Demons“, in *Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Juni 1996, S. 245–251. DOI: 10.1109/CVPR.1996.517081.

[155] H. Wang, L. Dong, J. O'Daniel, R. Mohan, A. S. Garden, K. K. Ang, D. A. Kuban, M. Bonnen, J. Y. Chang und R. Cheung, „Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy“, *Physics in Medicine & Biology*, Jg. 50, Nr. 12, S. 2887, Juni 2005. DOI: 10.1088/0031-9155/50/12/011.

[156] D.-J. Kroon und C. H. Slump, „MRI Modality transformation in demon registration“, in *IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI '09*, IEEE, 2009, S. 963–966. DOI: 10.1109/ISBI.2009.5193214.

[157] S. Sourbron, R. Luypaert, P. Van Schuerbeek, M. Dujardin, T. Stadnik und M. Osteaux, „Deconvolution of dynamic contrast-enhanced MRI data by linear inversion: Choice of the regularization parameter“, *Magnetic Resonance in Medicine*, Jg. 52, Nr. 1, S. 209–213, 2004. DOI: 10.1002/mrm.20113.

[158] S. Klein, M. Staring, K. Murphy, M. A. Viergever und J. P. W. Pluim, „Elastix: A toolbox for intensity-based medical image registration“, *IEEE transactions on medical imaging*, Jg. 29, Nr. 1, S. 196–205, Jan. 2010. DOI: 10.1109/TMI.2009.2035616.

[159] D. P. Shamonin, E. E. Bron, B. P. F. Lelieveldt, M. Smits, S. Klein und M. Staring, „Fast Parallel Image Registration on CPU and GPU for

Diagnostic Classification of Alzheimer's Disease", *Frontiers in Neuroinformatics*, Jg. 7, Jan. 2014. DOI: 10.3389/fninf.2013.00050.

[160] N. J. Tustison, P. A. Cook, A. J. Holbrook, H. J. Johnson, J. Muschelli, G. A. Devenyi, J. T. Duda, S. R. Das, N. C. Cullen, D. L. Gillen, M. A. Yassa, J. R. Stone, J. C. Gee und B. B. Avants, „The ANTsX Ecosystem for Quantitative Biological and Medical Imaging", *Scientific Reports*, Jg. 11, Nr. 1, S. 9068, Apr. 2021. DOI: 10.1038/s41598-021-87564-6.

[161] T. Chan und L. Vese, „Active Contours without Edges", *IEEE Transactions on Image Processing*, Jg. 10, Nr. 2, S. 266–277, Feb. 2001. DOI: 10.1109/83.902291.

[162] T. O'Malley, E. Bursztein, J. Long, F. Chollet, H. Jin, L. Invernizzi et al., *Keras Tuner*, https://github.com/keras-team/keras-tuner, 2019.

[163] M. Kirby und L. Sirovich, „Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Jg. 12, Nr. 1, S. 103–108, Jan. 1990. DOI: 10.1109/34.41390.

[164] M. Turk und A. Pentland, „Face Recognition Using Eigenfaces", in *1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Proceedings*, Juni 1991, S. 586–591. DOI: 10.1109/CVPR.1991.139758.

[165] A. Reinke, M. D. Tizabi, M. Baumgartner, M. Eisenmann, D. Heckmann-Nötzel, A. E. Kavur, T. Rädsch, C. H. Sudre, L. Acion, M. Antonelli, T. Arbel, S. Bakas, A. Benis, F. Buettner, M. J. Cardoso, V. Cheplygina, J. Chen, E. Christodoulou, B. A. Cimini, K. Farahani, L. Ferrer, A. Galdran, B. van Ginneken, B. Glocker, P. Godau, D. A. Hashimoto, M. M. Hoffman, M. Huisman, F. Isensee, P. Jannin, C. E. Kahn, D. Kainmueller, B. Kainz, A. Karargyris, J. Kleesiek, F. Kofler, T. Kooi, A. Kopp-Schneider, M. Kozubek, A. Kreshuk, T. Kurc, B. A. Landman, G. Litjens, A. Madani, K. Maier-Hein, A. L. Martel,

E. Meijering, B. Menze, K. G. M. Moons, H. Müller, B. Nichyporuk, F. Nickel, J. Petersen, S. M. Rafelski, N. Rajpoot, M. Reyes, M. A. Riegler, N. Rieke, J. Saez-Rodriguez, C. I. Sánchez, S. Shetty, R. M. Summers, A. A. Taha, A. Tiulpin, S. A. Tsaftaris, B. Van Calster, G. Varoquaux, Z. R. Yaniv, P. F. Jäger und L. Maier-Hein, „Understanding metric-related pitfalls in image analysis validation“, *Nature Methods*, Jg. 21, Nr. 2, S. 182–194, Feb. 2024. DOI: 10.1038/s41592-023-02150-0.

[166] A. Karch, C. Vogelmeier, T. Welte, R. Bals, H.-U. Kauczor, J. Biederer, J. Heinrich, H. Schulz, S. Gläser, R. Holle, H. Watz, S. Korn, N. Adaskina, F. Biertz, C. Vogel, J. Vestbo, E. F. M. Wouters, K. F. Rabe, S. Söhler, A. Koch und R. A. Jörres, „The German COPD Cohort COSYCONET: Aims, Methods and Descriptive Analysis of the Study Population at Baseline“, *Respiratory Medicine*, Jg. 114, S. 27–37, Mai 2016. DOI: 10.1016/j.rmed.2016.03.008.

[167] University Hospital Heidelberg, *Image-Based Structural and Functional Phenotyping of the COSYCONET Cohort Using MRI and CT (MR-COPD)*, ClinicalTrials.gov, NCT Identifier: NCT02629432. Updated May 2017. Abgerufen am 19. März, 2026, 2017.

[168] O. Weinheimer, T. Achenbach, C. P. Heussel und C. Düber, „Automatic Lung Segmentation in MDCT Images“, in *Fourth International Workshop on Pulmonary Image Analysis*, Bd. Vol. 2011, Toronto, Canada, Sep. 2011, S. 241–255.

[169] P. Leutz-Schmidt, J. Grolig, L. Wucherpfennig, O. Sommerburg, M. Eichinger, S. Wege, S. Y. Graeber, J.-P. Schenk, A. Alrajab, H.-U. Kauczor, M. Stahl, M. A. Mall, A. Koeppe, B. Nestler, M. Selzer, S. M. F. Triphan und M. O. Wielpütz, „Quantitative MRI detects delayed perfusion and impact of bronchial artery dilatation on pulmonary circulation in patients with cystic fibrosis“, *European Radiology*, Jg. 35, Nr. 10, S. 6217–6228, Okt. 2025. DOI: 10.1007/s00330-025-11589-y.

[170] E. J. Monroe, D. B. Pierce, C. R. Ingraham, G. E. Johnson, G. M. Shivaram und K. Valji, „An Interventionalist's Guide to Hemoptysis in Cystic Fibrosis", *Radiographics: A Review Publication of the Radiological Society of North America, Inc*, Jg. 38, Nr. 2, S. 624–641, 2018. DOI: 10.1148/rg.2018170122.

[171] K.-M. Xu, „Using the Bootstrap Method for a Statistical Significance Test of Differences between Summary Histograms", *Monthly Weather Review*, Jg. 134, Nr. 5, S. 1442–1453, Mai 2006. DOI: 10.1175/MWR3133.1.

[172] H. C. Schwickert, H. U. Kauczor, F. Schweden und H. H. Schild, „[The anatomy of the bronchial arteries–their imaging with spiral CT]", *RoFo: Fortschritte Auf Dem Gebiete Der Rontgenstrahlen Und Der Nuklearmedizin*, Jg. 160, Nr. 6, S. 506–512, Juni 1994. DOI: 10.1055/s-2008-1032468.

[173] R. A. Jörres, T. Welte, R. Bals, A. Koch, M. Schnoor und C. Vogelmeier, „Einfluss systemischer Manifestationen und Komorbiditäten auf den klinischen Zustand und den Verlauf bei COPD", *DMW - Deutsche Medizinische Wochenschrift*, Jg. 135, Nr. 10, S. 446–449, März 2010. DOI: 10.1055/s-0030-1249185.

[174] T. Achenbach, O. Weinheimer, C. Buschsieweke, C. P. Heussel, M. Thelen und H. U. Kauczor, „[Fully automatic detection and quantification of emphysema on thin section MD-CT of the chest by a new and dedicated software]", *RoFo*, Jg. 176, Nr. 10, S. 1409–1415, Okt. 2004. DOI: 10.1055/s-2004-813530.

[175] O. Weinheimer, T. Achenbach, C. Bletz, C. Duber, H.-U. Kauczor und C. P. Heussel, „About Objective 3-D Analysis of Airway Geometry in Computerized Tomography", *IEEE Transactions on Medical Imaging*, Jg. 27, Nr. 1, S. 64–74, Jan. 2008. DOI: 10.1109/TMI.2007.902798.

[176] S. M. F. Triphan, J. Biederer, K. Burmester, I. Fellhauer, C. F. Vogelmeier, R. A. Jörres, H.-U. Kauczor, C. P. Heußel, M. O. Wielpütz und B. J. Jobst, „Design and application of an MR reference phantom for

multicentre lung imaging trials“, *PLOS ONE*, Jg. 13, Nr. 7, e0199148, Juli 2018. DOI: 10.1371/journal.pone.0199148.

[177] M. Eichinger, M. Puderbach, C. Fink, J. Gahr, S. Ley, C. Plathow, S. Tuengerthal, I. Zuna, F.-M. Müller und H.-U. Kauczor, „Contrast-enhanced 3D MRI of lung perfusion in children with cystic fibrosis—initial results“, *European Radiology*, Jg. 16, Nr. 10, S. 2147–2152, Okt. 2006. DOI: 10.1007/s00330-006-0257-7.

[178] A. M. Weng, J. F. Heidenreich, C. Metz, S. Veldhoen, T. A. Bley und T. Wech, „Deep learning-based segmentation of the lung in MR-images acquired by a stack-of-spirals trajectory at ultra-short echo-times“, *BMC Medical Imaging*, Jg. 21, Nr. 1, S. 79, Mai 2021. DOI: 10.1186/s12880-021-00608-1.

[179] K. Xuan, L. Si, L. Zhang, Z. Xue, Y. Jiao, W. Yao, D. Shen, D. Wu und Q. Wang, „Reducing magnetic resonance image spacing by learning without ground-truth“, *Pattern Recognition*, Jg. 120, S. 108 103, Dez. 2021. DOI: 10.1016/j.patcog.2021.108103.

[180] Deutsche Forschungsgemeinschaft, *Leitlinien zum Umgang mit Forschungsdaten*, Verabschiedet durch den Senat der DFG am 30. September 2015. https://www.dfg.de/resource/blob/172112/4ea861510ea369157afb499e96fb359a/leitlinien-forschungsdaten-data.pdf. Abgerufen am 15. März 2026, 2015.

[181] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. G. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. C. ’t Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao und B. Mons,

„The FAIR Guiding Principles for scientific data management and stewardship“, *Scientific Data*, Jg. 3, Nr. 1, S. 160 018, März 2016. DOI: 10.1038/sdata.2016.18.

[182] National Electrical Manufacturers Association (NEMA), *NEMA PS3 / ISO 12052: Digital Imaging and Communications in Medicine (DICOM) Standard*, (Frei verfügbar unter https://www.dicomstandard.org/), Rosslyn, VA, USA.

[183] L. Faggioni, E. Neri, C. Castellana, D. Caramella und C. Bartolozzi, „The Future of PACS in Healthcare Enterprises“, *European Journal of Radiology*, From PACS to the Clouds, Jg. 78, Nr. 2, S. 253–258, Mai 2011. DOI: 10.1016/j.ejrad.2010.06.043.

[184] Z. Ahmed, K. Mohamed, S. Zeeshan und X. Dong, „Artificial Intelligence with Multi-Functional Machine Learning Platform Development for Better Healthcare and Precision Medicine“, *Database*, Jg. 2020, baaa010, Jan. 2020. DOI: 10.1093/database/baaa010.

[185] I. S. Hofer, M. Burns, S. Kendale und J. P. Wanderer, „Realistically Integrating Machine Learning Into Clinical Practice: A Road Map of Opportunities, Challenges, and a Potential Future“, *Anesthesia & Analgesia*, Jg. 130, Nr. 5, S. 1115, Mai 2020. DOI: 10.1213/ANE.0000000000004575.

[186] G. Choy, O. Khalilzadeh, M. Michalski, S. Do, A. E. Samir, O. S. Pianykh, J. R. Geis, P. V. Pandharipande, J. A. Brink und K. J. Dreyer, „Current Applications and Future Impact of Machine Learning in Radiology“, *Radiology*, Jg. 288, Nr. 2, S. 318–328, Aug. 2018. DOI: 10.1148/radiol.2018171820.

[187] A. Qayyum, J. Qadir, M. Bilal und A. Al-Fuqaha, „Secure and Robust Machine Learning for Healthcare: A Survey“, *IEEE Reviews in Biomedical Engineering*, Jg. 14, S. 156–180, 2021. DOI: 10.1109/RBME.2020.3013489.

[188] P. A. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez und J. G. Conde, „Research Electronic Data Capture (REDCap)—A Metadata-Driven Methodology and Workflow Process for Providing Translational Research Informatics Support“, *Journal of Biomedical Informatics*, Jg. 42, Nr. 2, S. 377–381, Apr. 2009. DOI: 10.1016/j.jbi.2008.08.010.

[189] P. A. Harris, R. Taylor, B. L. Minor, V. Elliott, M. Fernandez, L. O'Neal, L. McLeod, G. Delacqua, F. Delacqua, J. Kirby und S. N. Duda, „The REDCap Consortium: Building an International Community of Software Platform Partners“, *Journal of Biomedical Informatics*, Jg. 95, S. 103 208, Juli 2019. DOI: 10.1016/j.jbi.2019.103208.

[190] D. S. Marcus, T. R. Olsen, M. Ramaratnam und R. L. Buckner, „The Extensible Neuroimaging Archive Toolkit: An informatics platform for managing, exploring, and sharing neuroimaging data“, *Neuroinformatics*, Jg. 5, Nr. 1, S. 11–34, 2007. DOI: 10.1385/ni:5:1:11.

[191] J. Scherer, M. Nolden, J. Kleesiek, J. Metzger, K. Kades, V. Schneider, M. Bach, O. Sedlaczek, A. M. Bucher, T. J. Vogl, F. Grünwald, J.-P. Kühn, R.-T. Hoffmann, J. Kotzerke, O. Bethge, L. Schimmöller, G. Antoch, H.-W. Müller, A. Daul, K. Nikolaou, C. la Fougère, W. G. Kunz, M. Ingrisch, B. Schachtner, J. Ricke, P. Bartenstein, F. Nensa, A. Radbruch, L. Umutlu, M. Forsting, R. Seifert, K. Herrmann, P. Mayer, H.-U. Kauczor, T. Penzkofer, B. Hamm, W. Brenner, R. Kloeckner, C. Düber, M. Schreckenberger, R. Braren, G. Kaissis, M. Makowski, M. Eiber, A. Gafita, R. Trager, W. A. Weber, J. Neubauer, M. Reisert, M. Bock, F. Bamberg, J. Hennig, P. T. Meyer, J. Ruf, U. Haberkorn, S. O. Schoenberg, T. Kuder, P. Neher, R. Floca, H.-P. Schlemmer und K. Maier-Hein, „Joint Imaging Platform for Federated Clinical Data Analytics“, *JCO Clinical Cancer Informatics*, Nr. 4, S. 1027–1038, Dez. 2020. DOI: 10.1200/CCI.20.00045.

[192] M. J. Cardoso, W. Li, R. Brown, N. Ma, E. Kerfoot, Y. Wang, B. Murrey, A. Myronenko, C. Zhao, D. Yang, V. Nath, Y. He, Z. Xu, A. Hatamizadeh, A. Myronenko, W. Zhu, Y. Liu, M. Zheng, Y. Tang, I. Yang, M. Zephyr, B. Hashemian, S. Alle, M. Z. Darestani, C. Budd, M. Modat,

T. Vercauteren, G. Wang, Y. Li, Y. Hu, Y. Fu, B. Gorman, H. Johnson, B. Genereaux, B. S. Erdal, V. Gupta, A. Diaz-Pinto, A. Dourson, L. Maier-Hein, P. F. Jaeger, M. Baumgartner, J. Kalpathy-Cramer, M. Flores, J. Kirby, L. A. D. Cooper, H. R. Roth, D. Xu, D. Bericat, R. Floca, S. K. Zhou, H. Shuaib, K. Farahani, K. H. Maier-Hein, S. Aylward, P. Dogra, S. Ourselin und A. Feng, *MONAI: An Open-Source Framework for Deep Learning in Healthcare*, Nov. 2022. DOI: 10.48550/arXiv.2211.02701.

[193] L. Griem, P. Zschumme, M. Laqua, N. Brandt, E. Schoof, P. Altschuh und M. Selzer, „KadiStudio: FAIR Modelling of Scientific Research Processes“, *Data Science Journal*, Jg. 21, Nr. 1, Sep. 2022. DOI: 10.5334/dsj-2022-016.

[194] Y. Zhao, P. Altschuh, J. Santoki, L. Griem, G. Tosato, M. Selzer, A. Koeppe und B. Nestler, „Characterization of Porous Membranes Using Artificial Neural Networks“, *Acta Materialia*, Jg. 253, S. 118 922, Juli 2023. DOI: 10.1016/j.actamat.2023.118922.

[195] A. Koeppe, F. Bamer, M. Selzer, B. Nestler und B. Markert, „Workflow concepts to model nonlinear mechanics with computational intelligence“, *PAMM*, Jg. 21, Nr. 1, e202100238, 2021. DOI: 10.1002/pamm.202100238.

[196] A. Koeppe, F. Bamer, M. Selzer, B. Nestler und B. Markert, „Explainable Artificial Intelligence for Mechanics: Physics-Explaining Neural Networks for Constitutive Models“, *Frontiers in Materials*, Jg. 8, Feb. 2022. DOI: 10.3389/fmats.2021.824958.

[197] A. Koeppe, „Deep Learning in the Finite Element Method“, Dissertation, RWTH Aachen University, Aachen, Germany, 2021. DOI: 10.18154/RWTH-2021-04990.

Abbildungsverzeichnis

2.1 Aufbau der Lunge und Atemweg 10
2.2 Darstellung pulmonaler Blutgefäße und des Gasaustausches 11
2.3 Erklärung der PRM-Methode 18
2.4 Koordinatensystem des MRT 19
2.5 Erklärung der T1- und T2-Relaxation 21
2.6 Vergleich zweier T1-gewichteter Bilder mit unterschiedlichen Flipwinkeln. 26
2.7 Verlauf des mittleren Lungensignals einer DCE-MRT Messung . . 31
2.8 Visualisierung der Faltungsoperation 36
2.9 Visualisierung des Pooling 37
2.10 Darstellung der einzelnen Schritte von populationsbasierten Metaheuristiken 38
2.11 Vergleich der Generierung neuer Individuen zwischen EAs und SIAs 40
2.12 Ablauf des EHO 42
2.13 Darstellung eines HMRF 48
3.1 Darstellung des informationsgetriebenen Bildanalysezyklus 52
4.1 Einordnung der Segmentierung in den informationsgetriebenen Bildanalysezyklus 64
4.2 Beispielhafte Segmentierungsergebnisse des EHOHMRF für verschiedene EHO Varianten 73
4.3 Beispielhafte Segmentierungsergebnisse für verschiedene HMRF Verteilungsfunktionen 75
4.4 Beispielhafte Segmentierung der realen PU-Struktur mit EHOHMRF-Modell für zwei Klassen 78
4.5 Beispielhafte Segmentierung der realen PU-Struktur für drei Klassen 80
4.6 Beispielhafte Lungensegmentierung für EHOHMRF-Methode . . . 83

4.7 Vergleich der EHOHMRF Lungensegmentierung mit Groundtruth . 85
4.8 Überblick über die verwendete U-Net Architektur für das HMRF-UNet 91
4.9 Visualisierung der Quader-basierten Datenaufspaltung, die für den realen PU-Schaum Datensatz verwendet wurde. 92
4.10 Vergleich der Segmentierungsergebnisse verschiedener HMRF-UNet Modelle für *ArtPUFoam*-Datensatz 96
4.11 Vergleich der Segmentierungsergebnisse von Modellen mit gewichteter Banerjee-Nachbarschaft und unterschiedlichem Schwellwerten 97
4.12 Visualisierung der Effekte des Vortrainierens mit dem HMRF-UNet auf ein überwachtes Training für verschiedene Mengen an Trainingsdaten 101
4.13 Boxplots der DICE-Scores für Modell mit und ohne Vortraining durch das HMRF-UNet für verschiedene Größen des Trainingsdatensatzes 102
4.14 Vergleich der Segmentierungsergebnisse verschiedener HMRF-UNet Modelle für ein beispielhaftes Bild des realen µCT-Datensatzes 104
4.15 Vergleich der Segmentierungsergebnisse von HMRF-UNet Modellen, die zwei oder drei Klassen im Bild suchen 105
4.16 Beispielhafte Lungensegmentierung für HMRF-UNet 108
4.17 DICE-Score pro Schicht für Lungen-MRT Segmentierung mit HMRF-UNet 110
4.18 Vergleich der HMRF-UNet Lungensegmentierung mit Groundtruth 111
4.19 Darstellung der einzelnen Schritte der algorithmischen Lungensegmentierung 114
4.20 Vergleich alte und optimierte Tracheasegmentierung 118
4.21 Beispielhafte Lungensegmentierung für algorithmische Lungensegmentierung 119
4.22 Vergleich der algorithmischen Lungensegmentierung mit Groundtruth 120
5.1 Einordnung der Perfusionsquantifizierung in den informationsgetriebenen Bildanalysezyklus 124
5.2 Darstellung der Herleitung der quantitativen Perfusionsparameter aus der Residuenfunktion 129

5.3 Überblick über die Perfusionsquantifizierungs-Pipeline von M. Konietzke . . . 132
5.4 Vergleich zwischen *elastix*-, Demon- und *ANTs*-Registrierung . . . 139
5.5 Atmung zwischen zwei Zeitpunkten einer DCE-MRT Messung . . 141
5.6 Hypermodelle für automatische Atemdetektion . . . 142
5.7 Aufbau des CNN zur automatischen Atemdetektion . . . 144
5.8 Mögliche Fehler während Injektion des Kontrastmittels im DCE-MRT . . . 148
5.9 Vorgehen zur automatischen Entfernung von Atmung im DCE-MRT 149
5.10 Vergleich Perfusionsquantifizierung mit und ohne entfernter Atmung 150
5.11 Darstellung des Vorgehens zur Berechnung der Perfusionsverzögerung 152
5.12 DCE-MRT Qualitätsscore . . . 153
5.13 Berechnung der EigenLung-Distanz . . . 156
5.14 EigenLung-Distanz für verschiedene Segmentierungsfehler . . . 157
5.15 Validierung der Lungensegmentierung . . . 159
5.16 Visualisierung einer dilatierten Bronchialarterie in der MR Angiographie . . . 161
5.17 Quantitative Perfusionskarten eines CF-Patienten ohne dilatierte Bronchialarterien . . . 163
5.18 Mittlere Histogramme der quantitativen Perfusionswerte für normale und defekte Perfusionsbereiche . . . 167
5.19 Quantitative Perfusionskarten eines CF-Patienten mit dilatierten Bronchialarterien . . . 169
5.20 Mittlere Histogramme der quantitativen Perfusionswerte für PatientInnen mit und ohne BAD jeweils für normale und defekte Perfusionsbereiche . . . 170
5.21 Boxplots verschiedener Segmentierungsmetriken aufgeteilt nach MRT-Scanner . . . 178
5.22 Boxplots der EigenLung-Distanz Metrik aufgeteilt nach MRT-Scanner 180
5.23 Korrelationsplot zwischen QDP und PRM_{emph} . . . 187
5.24 Quantitative Perfusionskarten für zwei verschiedene COPD-PatientInnen . . . 189
5.25 Boxplots verschiedener quantitativer Perfusionswerte nach MRT-Scanner . . . 191

5.26 Mittlere Histogramme verschiedener quantitativer Perfusionswerte nach MRT-Scanner 193
5.27 Untersuchung von Ausreißern innerhalb der PBF-Histogramme . . 198
5.28 Darstellung der entwickelten vollautomatischen und robusten Pipeline zur Perfusionsquantifizierung 200
6.1 Einordnung der entwickelten Daten-Workflows in den Bildanalysezyklus 204
6.2 Darstellung der medizinischen Bildverarbeitung ohne FDM 207
6.3 Darstellung der medizinischen Bildverarbeitung mit FDM 208
6.4 Kadi4Mat Record eines MRT-Bildes 211
6.5 Kadi4Mat Record für Studienüberblick 212
6.6 Konzept zur Beschreibung der Zusammenhänge eines KadiAI Projektes. 212
6.7 Kadi4Mat Wissensgraph des Atmungserkennungs-Projektes 214
6.8 Kadi4Mat Record Beispiele eines KadiAI-Projektes 214
6.9 Kadi4Mat Record eines MRT-Scanners 216
6.10 Kadi4Mat Records der klinischen und QCT Daten 217
6.11 Kadi4Mat Record der quantitativen Perfusionswerte 218
6.12 Kadi4Mat Wissensgraph einer Patientin 218
A.1 Boxplots verschiedener fuzzy Versionen der Segmentierungsmetriken aufgeteilt nach MRT-Scanner 288

Tabellenverzeichnis

4.1 Parameter des EHOHMRF-Algorithmus für alle durchgeführten Parameterstudien. 71
4.2 Vergleich der DICE-Scores für EHOHMRF mit normaler und gradientenbasierter EHO 72
4.3 Vergleich der DICE-Scores für EHOHMRF mit verschiedenen Verteilungsfunktionen für die Klassenmodellierung . 74
4.4 Vergleich der DICE-Scores für EHOHMRF mit verschiedenen Nachbarschaftsarten mit und ohne voxelweisem Gewichtungsfaktor 76
4.5 Vergleich der DICE-Scores für EHOHMRF mit verschiedenen Nachbarschaftsarten und -gewichten 77
4.6 Vergleich der DICE-Scores von EHOHMRF-Modellen mit verschiedenen Nachbarschaftsarten und -gewichten für drei beispielhafte Lungen-MRT Bilder. 84
4.7 Vergleich der mittleren DICE-Scores des HMRF-UNet für verschiedene Nachbarschaftsarten und Nachbarschaftsgewichte. . . 95
4.8 Analyse der Verwendung des HMRF-UNet zum Vortrainieren eines UNet als Ausgangspunkt für ein überwachtes Training 99
4.9 Vergleich der DICE-Scores von HMRF-UNet-Modellen mit verschiedenen Nachbarschaftsarten und -gewichten sowie Schwellwerten σ_{thresh} 109
4.10 Vergleich der DICE-Scores der algorithmischen Lungensegmentierung mit alter und neuer Tracheasegmentierung . 119
5.1 Vergleich der Performance der einzelnen Modellarchitekturen zur automatischen Atmungserkennung auf dem ungesehenen Testdatensatz. 144

5.2 Validierung des trainierten CNN zur Atmungserkennung auf einem künstlich erzeugten Testdatensatz 146
5.3 Überblick über die CF-BAD Studienkohorte 162
5.4 Spearman Korrelationen zwischen den quantitativen MRT-Werten und den MRT-Scores, ppFEV1 und LCI für CF-BAD-Studie. 165
5.5 Quantitative Perfusionsparameter aufgeteilt nach normalen und defekten Perfusionsbereichen für die Gesamtkohorte sowie die zwei Gruppen mit und ohne BAD 166
5.6 Überblick über die COSYCONET-Studienkohorte 174
5.7 Vergleiche zwischen algorithmischer, HMRF-UNet und AI-Segmentierung aus [106] 176
5.8 Korrelationen zwischen verschiedenen QDP-Varianten und klinischen Parametern 181
5.9 Korrelationen zwischen verschiedenen QDP-Varianten und QCT-Werten 182
5.10 Vergleich verschiedener Registrierungsvarianten durch Betrachtung der Korrelation mit klinischen Parametern 183
5.11 Vergleich verschiedener Registrierungsvarianten durch Betrachtung der Korrelation mit QCT-Werten 184
5.12 Spearman-Korrelationen von PBF und QDP mit allen klinischen Parametern sowie QCT-Werten 186
5.13 Quantitative Perfusionswerte für die gesamte COSYCONET-Kohorte sowie aufgeteilt nach GOLD-Klassen. . . . 188
5.14 Spearman-Korrelationen zwischen QDP und klinischen Parametern aufgeteilt nach MRT-Scanner 192
5.15 Spearman-Korrelationen zwischen QDP und QCT-Werten aufgeteilt nach MRT-Scanner 194
5.16 Vergleich der MRT-Qualitätsscores für Atmung zwischen den einzelnen Scannern 195
5.17 Vergleich der MRT-Qualitätsscores für die Bolusinjektion zwischen den einzelnen Scannern 197
5.18 Spearman-Korrelationen von QDP mit QCT-Werten ohne PatientInnen, die während der Messung atmeten. 199

A.1 Eigenschaften der einzelnen MRT-Scanner in der COSYCONET-Studie. 287
A.2 Parameter der einzelnen MRT-Sequenzen der COSYCONET-Studie. 288

Abkürzungs- und Symbolverzeichnis

Abkürzungen

μCT	Micro-CT
2D	Zweidimensional
3D	Dreidimensional
^{129}Xe-MRT	Xenon MRT
^{3}He-MRT	Helium MRT
ACO	Ameisenkolonie Optimierung
AFM	Rasterkraftmikroskopie
AIF	Arterielle Inputfunktion
API	Automatische Programmier Schnittstelle
ArtPUFoam	Datensatz mit künstlich erzeugten PU-Schäumen
AWPi10	Mittlere Wanddicke eines Atemweglumens von 10 mm
BAD	Dilatierte Bronchialarterien
BE-Index	Index der Verstärkung der Bronchialwände
CF	Mukoviszidose
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator
CNN	Convolutional Neural Network
CNR	Kontrast-Rausch-Verhältnis
CO_2	Kohlenstoffdioxid

COPD	Chronisch obstruktive Lungenerkrankung
CS	Kuckuck-Suchalgorithmus
CT	Computertomographie
CycleGAN	Cycle-Consistent Adversarial Networks
DCE-MRT	Dynamisches kontrastmittelunterstütztes MRT
DLCO	Kohlenmonoxid-Diffusionskapazität
DSC	DICE-Score
EA	Evolutionärer Algorithmus
EHO	Elefantenherden Optimierung
EHOHMRF	Segmentierungsmethode, die Elefantenherden Optimierung mit verborgenen zufälligen Markov-Feldern kombiniert
EI	Emphysem-Index
ELD	EigenLung-Distanz
EM	Expectation Maximization
F_2-MRT	MRT mit fluorierten Gasen
FAIR	Findable, Accessible, Interoperable, Reusable
FD-MRT	Fourier-Dekomposition MRT
FDM	Forschungsdatenmanagement
FDSC	Fuzzy DICE-Score
FELD	Fuzzy EigenLung-Distanz
FEV1	Forciertes exspiratorisches Volumen in einer Sekunde
FMASD	Fuzzy mittlerer Oberflächenabstand
FVC	Forcierte Vitalkapazität
GAN	Generatives neuronales Netz
GBCA	Gadolinium-basiertes Kontrastmittel
GBEHO	Gradientenbasierte Elefantenherden Optimierung
GCV	Generalisierte Kreuzvalidierung

GOLD	Global Initiative for Chronic Obstructive Lung Disease
GRAPPA	GeneRalized Autocalibrating Partial Parallel Acquisition
GRE	Gradientenecho
GT	Groundtruth
HF	Hochfrequenz
HMRF	Verborgenes zufälliges Markov-Feld
HMRF-UNet	Segmentierungsmethode, die verborgene zufällige Markov-Felder mit einem U-Net kombiniert
$\text{HMRF-UNet}_{\text{ban}}$	HMRF-UNet mit normaler Banerjee-Nachbarschaft
$\text{HMRF-UNet}_{\text{none}}$	HMRF-UNet ohne Nachbarschafts-Loss
$\text{HMRF-UNet}_{\text{pot}}$	HMRF-UNet mit normaler Potts-Nachbarschaft
$\text{HMRF-UNet}_{\text{wban}}$	HMRF-UNet mit gewichteter Banerjee-Nachbarschaft
$\text{HMRF-UNet}_{\text{wpot}}$	HMRF-UNet mit gewichteter Potts-Nachbarschaft
HPV	Hypoxische pulmonale Vasokonstriktion
HU	Hounsfield-Einheiten
ICM	Iterated Conditional Modes
Kadi4Mat	Karlsruhe Data Infrastructure for Materials Science
KI	Künstliche Intelligenz
LCC	L-Kurven Kriterium
LCI	Lungenbelüftung
LCI2%	Anzahl Lungenumsätze bis 98% des Markergases ausgespült wurden
LEO	Lokaler Ausbruch Operator
LOL	Linker Oberlappen
LSTM	Long Short-Term Memory
LUL	Linker Unterlappen
MASD	Mittlerer Oberflächenabstand

MBW	Gasauswaschverfahren
MIP	Maximumintensitätsprojektion
ML	Maschinelles Lernen
MLA	Mittlere Lungenabschwächung
MLD	Mittlere Lungendichte
MLP	Mulitlayer-Perzeptron
MP-MRT	Matrix-Pencil Decomposition MRT
MRF	Zufälliges Markov-Feld
MRT	Magnetresonanztomographie
MTT	Mittlere Transitzeit
NCC	Normalisierte Kreuzkorrelation
NN	Neuronales Netzwerk
NSD	Normalisierte Oberflächendistanz
O_2	Sauerstoff
O_2-MRT	Sauerstoffverstärktes MRT
Pace3D	Parallel Algorithms for Crystal Evolution in 3D
PACS	Picture Archiving and Communication System
PBF	Pulmonaler Blutfluss
PBV	Pulmonales Blutvolumen
PCA	Hauptkomponentenanalyse
Perc15	15. Perzentil des Lungen HU-Histogramms
ppFEV1	Forciertes exspiratorisches Volumen in einer Sekunde in Prozent vom Soll
ppFVC	Forcierte Vitalkapazität in Prozent vom Soll
PREFUL	Phase-Resolved Functional Lung
PRM	Parametric Response Map
PRM_{emph}	Parametric Response Map der Emphysem Voxel

PRM_{fSAD}	Parametric Response Map der Voxel von kranken, kleinen Atemwegen
PRM_{normal}	Parametric Response Map der normalen Voxel
PSO	Partikel-Schwarm-Optimierung
PU	Polyurethan
QCT	Quantitatives CT
QDP	Perfusionsdefekte in Prozent
RealPUFoam	Datensatz mit echten PU-Schäumen
ReLU	Rectified Linear Unit
REM	Rasterelektronenmikroskopie
RML	Rechter Mittellappen
ROL	Rechter Oberlappen
RUL	Rechter Unterlappen
SFTR	Tikhonov-Regularisierung in Standardform
SFTR-LCC	Tikhonov-Regularisierung in Standardform mit L-Kurven Kriterium
SIA	Schwarmintelligenz-Algorithmus
SNR	Signal-Rausch-Verhältnis
SOM	Selbstorganisierende Karte
SVD	Singulärwertzerlegung
T1	Spin-Gitter Relaxation oder Längsrelaxation
T2	Spin-Spin Relaxation oder Querrelaxation
T2*	Spin-Spin Relaxation inklusive Magnetfeldinhomogenitäten
TE	Echozeit
TEM	Transmissionselektronenmikroskopie
TR	Repetitionszeit
TRICKS	Time-Resolved Imaging of Contrast Kinetics

tSVD	Abgeschnittene Singulärwertzerlegung
TWIST	Time-resolved Angiography With Interleaved Stochastic Trajectories
VOI	Referenzvolumen
Voxel	Volumenelement
WHO	Weltgesundheitsorganisation
YACTA	Yet Another CT Analyzer

Operatoren

$(\,)^\top$	Transponieren einer Matrix
$*$	Faltungsoperator
arg max	Position des maximalen Wertes
arg min	Position des minimalen Wertes
$\bar{x}$	Mittelwert von Matrix x
$\cap$	Schnittmenge
$\cdot$	Vektorprodukt
$\int$	Integral
$\lfloor\,\rfloor$	Abrundungsfunktion
ln	Natürlicher Logarithmus
$\|\,\|$	Summe der Elemente einer Kante
$\mathcal{F}\{\,\}$	Fouriertransformation
$\mathcal{F}^{-1}\{\,\}$	Inverse Fouriertransformation
$\odot$	Hadamard-Produkt
$\sum$	Summe
$d(\,)$	Euklidische Distanz
$diag(\,)$	Diagonalmatrix
$P(\,\|\,)$	A-posteriori-Verteilung

Symbole

$\mathbf{1}$	Einsvektor
α	Konstantes Gewicht des Nachbarschafts-Strafterms
α_s	Voxelspezifisches Nachbarschaftsgewicht für Voxel s
$\widetilde{\alpha}_s$	Fuzzy voxelspezifisches Nachbarschaftsgewicht für Voxel s
β	Konstanter Gewichtungsfaktor für das Positionsupdate der Matriarchin
$\boldsymbol{\delta}_m$	10. Perzentil der Differenz aus globalen Minima und Maxima
$\delta_{x_s x_t}$	Kronecker Delta zwischen x_s und x_t
ε	Gewichtungsfaktor für Gradientenupdate des EHO
γ	Konstanter Gewichtungsfaktor für Positionsupdate der normalen Herdenmitglieder
λ_d	Verteilungs-Loss Gewicht
λ_n	Nachbarschafts-Loss Gewicht
$\mu_{C_{f,i}}$	Mittelwert des interpolierten hochaufgelösten tiefpassgefilterten Subtraktkionsbilds
μ_l	Mittelwert der Klasse l
$\widetilde{\mu}_l$	Fuzzy Mittelwert der Klasse l
$\mu_{\omega_{C_{a,i}}}$	Mittelwert der interpolierten hochaufgelösten AIF
$\widetilde{\mu}_{x_s}$	Fuzzy Mittelwert von Voxel s
$\omega_{\mathbf{C_{a,i}}}$	Bewegliches Fenster der mit am Rand mit Nullen aufgefüllten und interpolierten arteriellen Inputfunktion
ω_0	Larmorfrequenz
Φ	Funktion der Wahrscheinlichkeitsverteilung
ϕ	Funktion der Wahrscheinlichkeitsdichte
σ	Standardabweichung

$\sigma_{a,me}$	Standardabweichung des AIF-Signals zwischen Maximum und Ende der AIF
$\sigma_{C_{f,i}}$	Standardabweichung des interpolierten hochaufgelösten tiefpassgefilterten Subtraktkionsbilds
$\widetilde{\sigma}_l$	Fuzzy Standardabweichung der Klasse l
σ_l	Standardabweichung der Klasse l
$\sigma_{\max}$	Schwellwert für Singulärwerte der tSVD Regularisierung
$\sigma_{\mathcal{N}_s}^{(k)}$	Standardabweichung innerhalb der Nachbarschaft k-ten Ordnung eines Voxels s
$\sigma_{\omega_{C_{a,i}}}$	Standardabweichung der interpolierten hochaufgelösten AIF
$\sigma_{S_{h,\max}}$	Standardabweichung des maximalen Signals für jeden Zeitpunkt innerhalb einer Region, die das Herz umfasst
σ_{thresh}	Schwellwert für die Standardabweichung innerhalb der Nachbarschaft
$\widetilde{\sigma}_{x_s}$	Fuzzy Standardabweichung von Voxel s
τ	Schwellwert für Lunge
θ_l	Parameterset der Klasse l
A	Querschnitt
$a_{\max}$	Maximum der AIF
$\mathcal{B}_A^{(\tau)}$	Randregionen von Segmentierung A mit Toleranz τ
B_0	Permanentes Hauptmagnetfeld
C	Standard Cauchy-Verteilung
$c_{s,l}$	Wahrscheinlichkeit der Zugehörigkiet von Voxels s zu Klasse l
$\mathbf{C_a}$	Konzentration des arteriellen Inputs
$\mathbf{C_f}$	Tiefpassgefiltertes Subtraktkionsbild
$\mathbf{C_v}$	Konzentration des venösen Outputs
$\mathbf{C_{VOI}}$	Konzentration eines Tracers innerhalb eines Referenzvolumens
$\mathbf{C_{a,i}}$	Interpolierte hochaufgelöste AIF

$\mathbf{C_{f,i}}$	Interpoliertes hochaufgelöstes tiefpassgefiltertes Subtraktionsbild
$\mathbf{C}$	Kontrastmittelkonzentration des untersuchten Gewebes
$\mathbf{c}$	Wahrscheinlichkeitskarte der Klassenzugehörigkeiten
$\mathbf{c}_s$	Wahrscheinlichkeitsvektor von Voxel s
$\mathcal{C}$	Menge aller Elefantenherden
c_i	i-te Elefantenherde
D	Dimension des Suchraums
E	Nachbarschafts-Strafterm
$\widetilde{E}_{\text{ban}}$	Fuzzy Nachbarschafts-Strafterm der Banerjee-Nachbarschaft
$\widetilde{E}_{\text{potts}}$	Fuzzy Nachbarschafts-Strafterm der Potts-Nachbarschaft
E_{ban}	Nachbarschafts-Strafterm der Banerjee-Nachbarschaft
E_{potts}	Nachbarschafts-Strafterm der Potts-Nachbarschaft
F	Fitnessfunktion
f_1	Zufälliger reeller Wert aus Uniformverteilung
f_2, f_3	Zufällige reelle Zahlen
$\mathbf{h}$	Wahrscheinlichkeitsdichtefunktion der Transitzeiten
h_{ba}	Relative Atmungshäufigkeit während Kontrastmitteldurchlauf in DCE-MRT Messung
h_{ga}	Relative Atmungshäufigkeit während kompletter DCE-MRT Messung
k	Kreuzkorrelation
$\mathbf{K}$	Faltungs-Kernel
$\mathbf{k}$	Vektor mit Kreuzkorrelationen
$k_1 - k_4$	Zufällige ganze Zahlen
k_l	Plateaubreite der Stompnorm-Verteilung der Klasse l
l	Klasse innerhalb eines Bildes

$\mathcal{L}$	Menge aller möglichen Klassen eines Bildes
L	Loss-Funktion
L_d	Verteilungs-Loss
L_n	Nachbarschafts-Loss
l_{C_a}	Anzahl der Zeitpunkte der AIF
M	Magnetisches Moment
$\mathrm{M_{XY}}$	Transversalmagnetisierung
$\mathrm{M_Z}$	Längsmagnetisierung
$\mathbf{m}^{(t)}_{a,c_i}$	Position eines Elefanten a aus Herde c_i zum Zeitpunkt t
$\mathbf{m}^{(t)}_{best,c_i}$	Position der Matriarchin zum Zeitpunkt t
$\mathbf{m}^{(t)}_{center,c_i}$	Mittlere Position der Elefantenherde c_i zum Zeitpunkt t
$\mathbf{m}^{(t)}_{c_i,j}$	Position eines Elefanten j in Herde c_i zum Zeitpunkt t
$\mathbf{m}^{(t)}_{Gbest}$	Position der global besten Matriarchin zum Zeitpunkt t
$\mathbf{m}^{(t)}_{LEO}$	Position eines Elefanten zum Zeitpunkt t nach LEO-Update
$\mathbf{m}_{max}$	Globales Maximum des $\mathcal{D}$-dimensionalen Lösungsraums
$\mathbf{m}_{min}$	Globales Minimum des $\mathcal{D}$-dimensionalen Lösungsraums
$\mathbf{m}^{(t)}_{worst,c_i}$	Position des ausgewachsenen Elefantenbullen zum Zeitpunkt t
$\Delta\mathbf{m}$	Schrittweite der gradientenbasierten Elefantenherden Optimierung
N	Größe der Elefantenpopulation
$\mathcal{N}$	Nachbarsschaftssystem
$\mathcal{N}_s$	Nachbarsschafts-Clique des Voxels s
N_c	Größe einer Elefantenherde
p_{eb}	Wahrscheinlichkeit einer zu frühen Bolusinjektion
p_{lb}	Wahrscheinlichkeit einer zu späten Bolusinjektion
p_{nb}	Wahrscheinlichkeit einer nicht stattgefundenen Bolusinjektion
p_{nc}	Wahrscheinlichkeit einer falschen Chaserinjektion

PSR	Schwellwert für Matriarchinnen-Update
Q_{ba}	Qualitätsscore für Atmung während Kontrastmitteldurchlauf in DCE-MRT Messung
Q_{eb}	Qualitätsscore für zu frühe Bolusinjektion
Q_{ga}	Qualitätsscore für Gesamtatmung in DCE-MRT Messung
Q_{lb}	Qualitätsscore für zu späte Bolusinjektion
Q_{nb}	Qualitätsscore für nicht stattgefundene Bolusinjektion
Q_{nc}	Qualitätsscore für falsche Chaserinjektion
r	Zufälliger reeller Wert aus Uniformverteilung
$\mathbf{R}$	Residuenfunktion
$\mathbf{r}_a, \mathbf{r}_b$	Vektoren mit zufälligen uniformen Einträgen
$r_1 - r_6$	Zufällige reelle Werte aus Uniformverteilung
$r_{s_i}^{(k)}$	Relative Anzahl der i-häufigsten Nachbarn im Nachbarschaftssystem k-ter Ordnung
r_{art}	Arterielle Korrelation
$\mathbf{R}_{\max}$	Residuen-Karte zum Zeitpunkt der größten Kontrastverstärkung
$\mathbf{R_{norm}}$	Normierte Residuen-Karte
s	Voxel
$\mathbf{S_0}$	MRT-Signal des untersuchten Gewebes vor Kontrastmittelankunft
$\mathbf{S}$	MRT-Signal des untersuchten Gewebes
$\mathcal{S}$	Bildraster
S_A	Kanten von Segmentierung A
S_m	Bewegliches Bild
S_s	Statisches Bild
ΔS_s	Räumliche Voxelabständen im statischen Bild
T	Schwellwert für voxelspezifisches Nachbarschaftsgewicht

t	Zeitpunkt bzw. Epoche
t_{BA}	Zeitpunkt der Kontrastmittelankunft
t_{BC}	Zeitpunkt des vollständigen Kontrastmitteldurchlaufs
$t_{\max}$	Maximal Epochenanzahl
t_{delay}	Perfusionsverzögerung
$\mathbf{U}$	Orthogonale Matrix der Singulärwertzerlegung
$\mathbf{u}$	Verschiebungsfeld
$\mathcal{U}$	Uniformverteilung
$\mathbf{U_{ll}}$	Likelihood Energiefunktion
$\widetilde{\mathbf{U}}_{\mathbf{ll}}$	Fuzzy Likelihood Energiefunktion
$\mathbf{U_n}$	Nachbarschafts Energiefunktion
$\widetilde{\mathbf{U}}_{\mathbf{n}}$	Fuzzy Nachbarschafts Energiefunktion
v	Fließgeschwindigkeit
$\dot{V}$	Volumenstrom
$\mathbf{V}$	Orthogonale Matrix der Singulärwertzerlegung
$\mathbf{W}$	Diagonalmatrix der Singulärwertzerlegung
$\hat{\mathbf{x}}$	Optimale Segmentierung eines Objektes
$\mathbf{X}$	Hidden Markov Random Field einer Segmentierung
$\mathbf{x}$	Matrix der Segmentierung eines Bildes
x_s	Klasse des Voxels s
$\mathbf{Y}$	Markov Random Field eines Bildes
$\mathbf{y}$	Matrix eines Bildes
y_s	Intensität des Voxels s

A Anhang

A.1 Elastix Registrierungsparameter

Rigid Registrierung

```
(FixedInternalImagePixelType "float")
(MovingInternalImagePixelType "float")
(FixedImageDimension 3)
(MovingImageDimension 3)
(UseDirectionCosines "false")
(Registration "MultiResolutionRegistration")
(Interpolator "BSplineInterpolator")
(ResampleInterpolator "FinalBSplineInterpolator")
(Resampler "DefaultResampler")
(FixedImagePyramid "FixedRecursiveImagePyramid")
(MovingImagePyramid "MovingRecursiveImagePyramid")
(Optimizer "AdaptiveStochasticGradientDescent")
(Transform "EulerTransform")
(Metric "AdvancedMattesMutualInformation")
(AutomaticScalesEstimation "true")
(AutomaticTransformInitialization "false")
(AutomaticTransformInitializationMethod "CenterOfGravity")
(HowToCombineTransforms "Compose")
(NumberOfHistogramBins 32)
(ErodeMask "false")
(NumberOfResolutions 4)
(ImagePyramidSchedule 8 4 2 1)
(MaximumNumberOfIterations 1000)
(MaximumStepLength 0.4)
(NumberOfSpatialSamples 2048)
(NewSamplesEveryIteration "true")
(ImageSampler "RandomSparseMask")
(BSplineInterpolationOrder 1)
(FinalBSplineInterpolationOrder 3)
(DefaultPixelValue 0)
(WriteResultImage "false")
(ResultImagePixelType "float")
(ResultImageFormat "mhd")
(RequiredRatioOfValidSamples 0.001)
```

Affine Registrierung

```
(FixedInternalImagePixelType "float")
(MovingInternalImagePixelType "float")
(FixedImageDimension 3)
(MovingImageDimension 3)
(UseDirectionCosines "false")
(Registration "MultiResolutionRegistration")
(Interpolator "BSplineInterpolator")
(ResampleInterpolator "FinalBSplineInterpolator")
(Resampler "DefaultResampler")
(FixedImagePyramid "FixedRecursiveImagePyramid")
(MovingImagePyramid "MovingRecursiveImagePyramid")
(Optimizer "AdaptiveStochasticGradientDescent")
(Transform "AffineTransform")
(Metric "AdvancedMattesMutualInformation")
(AutomaticScalesEstimation "true")
(HowToCombineTransforms "Compose")
(NumberOfHistogramBins 32)
(ErodeMask "false")
(NumberOfResolutions 4)
(ImagePyramidSchedule 8 4 2 1)
(MaximumNumberOfIterations 2000)
(MaximumStepLength 0.4)
(NumberOfSpatialSamples 2048)
(NewSamplesEveryIteration "true")
(ImageSampler "RandomSparseMask")
(BSplineInterpolationOrder 1)
(FinalBSplineInterpolationOrder 3)
(DefaultPixelValue 0)
(WriteResultImage "false")
(ResultImagePixelType "float")
(ResultImageFormat "mhd")
(RequiredRatioOfValidSamples 0.001)
```

BSpline Registrierung

```
(FixedInternalImagePixelType "float")
(MovingInternalImagePixelType "float")
(FixedImageDimension 3)
(MovingImageDimension 3)
(UseDirectionCosines "false")
(Registration "MultiResolutionRegistration")
(Interpolator "BSplineInterpolator")
(ResampleInterpolator "FinalBSplineInterpolator")
(Resampler "DefaultResampler")
(FixedImagePyramid "FixedRecursiveImagePyramid")
(MovingImagePyramid "MovingRecursiveImagePyramid")
(Optimizer "AdaptiveStochasticGradientDescent")
(Transform "BSplineTransform")
(Metric "AdvancedMattesMutualInformation")
(FinalGridSpacingInPhysicalUnits 16 16 16)
(HowToCombineTransforms "Compose")
(NumberOfHistogramBins 32)
(ErodeMask "false")
(NumberOfResolutions 3)
(ImagePyramidSchedule 4 2 1)
(MaximumNumberOfIterations 2000 5000 10000)
(MaximumStepLength 0.05)
(NumberOfSpatialSamples 2048)
(NewSamplesEveryIteration "true")
(ImageSampler "Random")
(BSplineInterpolationOrder 1)
(FinalBSplineInterpolationOrder 3)
(DefaultPixelValue 0)
(WriteResultImage "false")
(ResultImagePixelType "float")
(ResultImageFormat "mhd")
```

A.2 ANTs Registrierungsparameter

Affine Registrierung

```
type_of_transform: "Affine"
grad_step: 0.2
aff_metric: "mattes"
aff_sampling: 64,
aff_random_sampling_rate: 0.4,
aff_iterations: (2000,1500,800,200)
aff_shrink_factors: (16,8,4,2)
aff_smoothing_sigmas: (4,3,2,1)
```

SyN Registrierung

```
type_of_transform: "SyN"
grad_step: 0.25
flow_sigma: 4
total_sigma: 0
syn_metric: "mattes"
syn_sampling: 64
reg_iterations: (500000,300000,200000,100000)
```

A.3 Untersuchte Hyperparameter Atmungserkennung

Hypermodell 1

- Lernrate: [0.003, 0.001, 0.0003, 0.0001, 3e-05, 1e-05]
- Anzahl Kernel: [16, 32, 64, 128]
- Anzahl Convolution-Blöcke: [1, 2, 3, 4, 5]
- Anzahl Convolutions pro Block: [1, 2, 3, 4]
- Anzahl Dense-Schichten: [1, 2, 3, 4]

- Anzahl der Dense-Units: [16, 32, 64, 128]
- Dropoutrate Dense-Schicht: {Min: 0,0 ; Max: 0,6 ; Schritt: 0,05}
- Dropoutrate Convolution-Block: {Min: 0,0 ; Max: 0,6 ; Schritt: 0,05}

Hypermodell 2

- Lernrate: [0.003, 0.001, 0.0003, 0.0001, 3e-05, 1e-05]
- Anzahl Kernel: [16, 32, 64, 128]
- Anzahl Convolution-Blöcke: [1, 2, 3, 4, 5]
- Anzahl Convolutions pro Block: [1, 2, 3, 4]
- Anzahl LSTM-Schichten: [1, 2, 3, 4, 5, 6, 7, 8]
- Anzahl LSTM Units pro Schicht: [8, 16, 32, 64, 128, 256]
- Anzahl Dense-Schichten: [1, 2, 3, 4]
- Anzahl der Dense-Units: [16, 32, 64, 128]
- Dropoutrate Dense-Schicht: {Min: 0,0 ; Max: 0,6 ; Schritt: 0,05}
- Dropoutrate Convolution-Block: {Min: 0,0 ; Max: 0,6 ; Schritt: 0,05}
- Dropoutrate LSTM-Block: {Min: 0,0 ; Max: 0,6 ; Schritt: 0,05}

Hypermodell 3

- Lernrate: [0.003, 0.001, 0.0003, 0.0001, 3e-05, 1e-05]
- Anzahl Kernel vor Verbindung: [16, 32, 64]
- Anzahl Kernel nach Verbindung: [16, 32, 64]
- Anzahl Convolution-Blöcke pro Zweig: [1, 2, 3]
- Anzahl Convolutions pro Block vor Verbindung: [1, 2, 3]
- Anzahl Convolution-Blöcke nach Verbindung: [1, 2, 3]
- Anzahl Convolutions pro Block nach Verbindung: [1, 2, 3]

- Anzahl Dense-Schichten: [1, 2, 3, 4]
- Anzahl der Dense-Units: [16, 32, 64]
- Dropoutrate Dense-Schicht: {Min: 0,0 ; Max: 0,6 ; Schritt: 0,05}
- Dropoutrate Convolution-Block: {Min: 0,0 ; Max: 0,6 ; Schritt: 0,05}

A.4 COSYCONET Studiendetails

Tabelle A.1: Eigenschaften der einzelnen MRT-Scanner in der COSYCONET-Studie.

Modell	Software Version	Gradienten-system	Anzahl Kanäle	B_0 Feldstärke	Bohrung Durchmesser	Bohrung Länge
Siemens Aera	D13	45 mT/T	48	1,5 T	70 cm	137 cm
Siemens Aera	E11	33 mT/T	48	1,5 T	70 cm	137 cm
Siemens Aera	E11	45 mT/T	48	1,5 T	70 cm	137 cm
Siemens Avanto	B17	33 mT/T	8	1,5 T	60 cm	150 cm
Siemens Avanto	B17	45 mT/T	32	1,5 T	60 cm	150 cm
Siemens Avanto	D13	45 mT/T	32	1,5 T	60 cm	150 cm
Siemens Espree	B17	45 mT/T	32	1,5 T	70 cm	120 cm
Siemens Espree	B19	33 mT/T	32	1,5 T	70 cm	120 cm
Siemens Espree	B19	45 mT/T	32	1,5 T	70 cm	120 cm
Siemens Trio	B17	45 mT/T	32	3,0 T	60 cm	198 cm

Tabelle A.2: Parameter der einzelnen MRT-Sequenzen der COSYCONET-Studie.

Sequenz	TE	TR	Flipwinkel	Schichtdicke	Δt
Koronale VIBE	0,78–1,85 ms	2,80–4,39 ms	5–8°	4 mm	-
Transversale VIBE	0,87–1,87 ms	3,00–3,77 ms	5–8°	4 mm	-
TWIST	0,70–0,93 ms	1,72–2,14 ms	10–20°	5 mm	1,16–2,49 s

A.5 Validierung Segmentierung

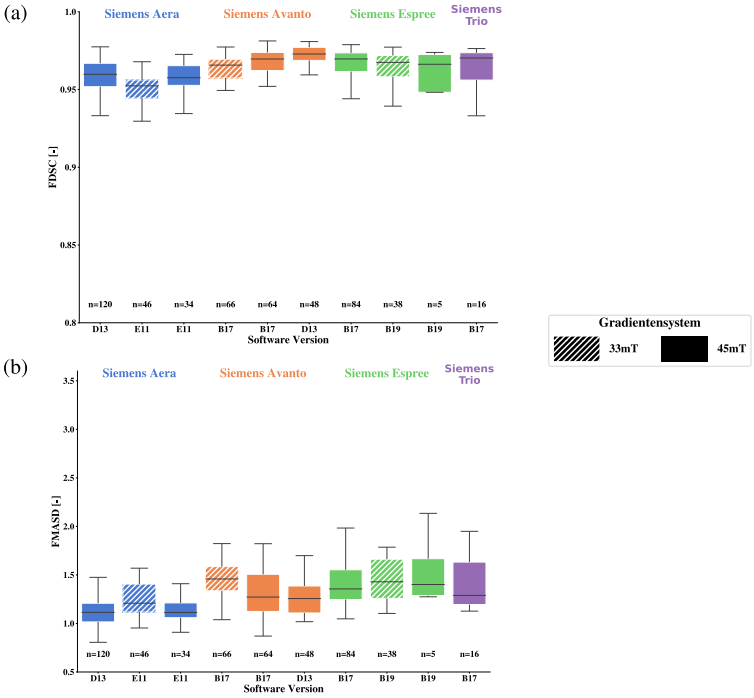


Abbildung A.1: Boxplots der fuzzy Versionen zweier Segmentierungsmetriken aufgeteilt nach MRT-Scanner. (a) FDSC und (b) FMASD. Die schraffierten Boxen entsprechen den MRT-Scannern mit 33 mT Gradientensystem.

B Veröffentlichungen

Beiträge in Printmedien

[1] **J. Grolig**, L. Griem, M. Selzer, H.-U. Kauczor, S. Triphan, B. Nestler, A. Koeppe. „Unsupervised segmentation of micro-CT scans of polyurethane structures by combining hidden markov random fields and a U-Net“. Materials Today Communications 51, (2026)
https://doi.org/10.1016/j.mtcomm.2026.114817

[2] P. Leutz-Schmidt*, **J. Grolig***, L. Wucherpfennig, D.-E. Optazaite, O. Sommerburg, M. Eichinger, S. Wege, S. Y. Graeber, M. Puderbach, J.-P. Schenk, A. Alrajab, H.-U. Kauczor, M. Stahl, M. Mall, A. Koeppe, B. Nestler, S. Triphan, M. O. Wielpütz, „Quantitative MRI detects delayed perfusion and impact of bronchial artery dilatation on pulmonary circulation in patients with cystic fibrosis“, Eur Radiol 35, 6217–6228 (2025).
https://doi.org/10.1007/s00330-025-11589-y
**Geteilte Erstautorenschaft*

[3] Y. Xu, S. Triphan, **J. Grolig**, H. Zhang, J. Biederer, C. J. Galban, H.-U. Kauczor, M. Wielpütz, O. Weinheimer, „A Learning Framework for Predicting CT-Based PRM Biomarker from MRI Sequences in COPD“, Medical Image Computing and Computer Assisted Intervention – MICCAI 2025. MICCAI 2025. Lecture Notes in Computer Science, vol 15967, 24-33 (2026)
https://doi.org/10.1007/978-3-032-04984-1_3

Konferenzbeiträge

[1] **J. Grolig**, „Anforderungen für Forschungsdatenmanagement und integriertes maschinelles Lernen zwischen Medizin und Ingenieurswissenschaften“, NFDI4Ing Conference (2021), Digital, Vortrag

[2] **J. Grolig**, S. Triphan, A. Koeppe, M. Selzer, B. Nestler, Mark O. Wielpütz, „Improving Image Resolution in 3-D Lung MRI Images by Combining Two Orthogonal Scans Using Super Resolution Reconstruction“, ESTI Annual Scientific Meeting (2022), Oxford, United Kingdom, Vortrag

[3] **J. Grolig**, S. Triphan, A. Koeppe, M. Selzer, B. Nestler, Mark O. Wielpütz, „Improving Image Resolution in 3-D Lung MRI Images by Combining Two Orthogonal Scans Using Super Resolution Reconstruction“, DZL Annual Meeting (2022), Hannover, Deutschland, Vortrag

[4] **J. Grolig**, P. Leutz-Schmidt, D.-E. Optazaite, O. Sommerburg, M. Eichinger, S. Wege, S. Y. Graeber, M. Puderbach, J.-P. Schenk, A. Alrajab, H.-U. Kauczor, M. Stahl, M. Mall, A. Koeppe, B. Nestler, S. Triphan, M. O. Wielpütz, „Magnetic Resonance Imaging Detects Delayed Perfusion Inside Perfusion Defects in Patients with Cystic Fibrosis“, DZL Annual Meeting (2023), Fürstenfeldbruck, Deutschland, Poster

[5] **J. Grolig**, A. Koeppe, M. Selzer, S. Triphan, M. O. Wielpütz, B. Nestler, „FAIR Development of Data-integrated AI to Detect Breathing Motion in Dynamic Lung MRI“, CORDI Conference on Research Data Infrastructure (2023), Karlsruhe, Deutschland, Poster

[6] **J. Grolig**, A. Koeppe, M. Selzer, S. Triphan, M. O. Wielpütz, B. Nestler, „FAIR Development of Data-integrated AI to Detect Breathing Motion in Dynamic Lung MRI“, AI InScide Out Unconference (2023), Heidelberg, Deutschland, Poster

[7] **J. Grolig**, S. Triphan, A. Koeppe, B. Nestler, M. Selzer, M. O. Wielpütz, „Automatic Evaluation of DCE-MRI Measurement Quality for Perfusion Quantification“, DZL Annual Meeting (2024), Bad Nauheim, Deutschland, Poster

[8] **J. Grolig**, Y. Xu, A. Koeppe, M. Selzer, B. Nestler, S. Triphan, M. O. Wielpütz, H.-U. Kauczor, „Quality Control for Perfusion Quantification in Lung DCE-MRI“, ISMRM (2025), Honolulu, USA, Digital Poster

[9] **J. Grolig**, Y. Xu, A. Koeppe, M. Selzer, B. Nestler, H.-U. Kauczor, M. O. Wielpütz, S. Triphan, „Quality Control for Perfusion Quantification in Lung DCE-MRI“, DZL Annual Meeting (2025), Heidelberg, Deutschland, Poster

[10] **J. Grolig**, A. Koeppe, L. Griem, S. Triphan, M. Selzer, B. Nestler, „Unsupervised Segmentation of Polyurethane Foam Structures By Combining Elephant Herding Optimization and Hidden Markov Random Fields (EHOHMRF)“, Helmholtz AI Conference (2025), Karlsruhe, Deutschland, Poster

Datensätze

[1] J. Grolig (2025). Data & Script HMRF-UNet. Zenodo.
`https://doi.org/10.5281/zenodo.17590658`

Betreute studentische Arbeiten

[1] J. Detroy, „Lungensegmentierung in morphologischen Lungen MRT Aufnahmen mit Self-Organizing-Maps“, Masterarbeit, KIT 2023.

[2] F. Habermann & P. Schepperle, „Identifikation der Bolusankunft in Perfusions-MRT mit Hilfe von CNN“, Hector-Seminar Projekt für hochbegabte Schüler, KIT 2023.

[3] M. Kollien, „Automatische Detektion des Zwerchfells in MRT-Perfusionsaufnahmen der Lunge“, Bachelorarbeit, KIT 2024.

[4] C-H. Lee, „Registration of Lung MRI with Machine Learning Methods“, Bachelorarbeit, KIT, 2024.

[5] K. L. Richter, „Patch-Wise Contrastive Learning for Self-Supervised Semantic Segmentation“, Bachelorarbeit, KIT, 2025.