

Aus dem Institut für Radiologie
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

KI-gestützte Segmentierung der Prostata

AI based segmentation of the prostate

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Günther Engel

aus Göttingen

Datum der Promotion: 25.11.2022

Inhaltsverzeichnis

Abbildungsverzeichnis	4
Abkürzungsverzeichnis	6
Zusammenfassung	7
1 Einleitung	11
1.1 Die Prostata	11
1.1.1 Anatomie und Funktion	11
1.1.2 Erkrankungen und klinische Bedeutung	14
1.2 Multiparametrische MRT der Prostata	19
1.2.1 Überblick.....	19
1.2.2 T2-gewichtete Bildgebung.....	20
1.2.3 Diffusionsgewichtete Bildgebung.....	21
1.2.4 DCE – dynamisches Kontrastmittelenhancement.....	22
1.2.5 Klinische Anwendung.....	24
1.3 Medizinische Bildanalyse	25
1.3.1 Grundlagen	25
1.3.2 Herausforderungen	26
1.4 Automatische Bildanalyse mittels Deep Learning	26
1.4.1 Allgemeines	26
1.4.2 Convolutional Neural Networks.....	27
1.4.3 U-Net.....	27
1.5 Zielsetzung der Arbeit	28
2 Materialien und Methoden	29
2.1 Erhebung der Datensätze	29
2.1.1 Auswahl der Studienpopulation	29
2.1.2 Untersuchungstechnik der MRT	30
2.1.3 Segmentierung	30
2.2 Automatisierte Vorverarbeitung der Bilddaten	32
2.2.1 Normalisierung und Entfernen des Biasfeldes	32
2.2.2 Resampling.....	33
2.2.3 Cropping	33
2.3 Training des Deep Learning Modells	36

2.3.1	Aufbau des Netzwerkes	36
2.3.2	Hyperparameter.....	37
2.4	Statistische Auswertung	37
3	Ergebnisse	39
3.1	Beschreibung des Datensatzes	39
3.2	Übereinstimmung zwischen den menschlichen Experten	39
3.3	Segmentierung der anatomischen Zonen	41
3.3.1	Verlauf des Trainings	41
3.3.2	Vergleich zu Beurteiler 1	42
3.3.3	Vergleich zu Beurteiler 2	43
3.4	Segmentierung des Karzinoms	43
3.4.1	Verlauf des Trainings	43
3.4.2	Vergleich zu den Beurteilern.....	45
4	Diskussion	48
4.1	Zusammenfassung der Ergebnisse	48
4.2	Diskussion der Methoden und Einordnung in die Literatur	48
4.3	Diskussion der Ergebnisse	58
4.4	Übersicht der Stärken und Limitationen	60
4.5	Schlussfolgerungen und Ausblick	62
	Literaturverzeichnis	65
	Lebenslauf	76
	Publikationsliste	78
	Danksagung	80
	Bestätigung eines akkreditierten Statistikers	81

Abbildungsverzeichnis

Abbildung 1: Schematische Darstellung der zonalen Gliederung der Prostata.....	10
Abbildung 2: Beispielsegmentierung einer Prostata mit karzinomsuspektem Areal...31	
Abbildung 3: Zusammenfassung der Vorverarbeitung.....	35
Abbildung 4: Aufbau eines Residual Blocks mit zwei Convolutional-Schichten.....	36
Abbildung 5: Boxplots für die Übereinstimmung der Segmentierungen zwischen den menschlichen Experten.....	40
Abbildung 6: Verlauf der Metriken über die verschiedenen Epochen des Trainings...41	
Abbildung 7: Exemplarische Segmentationen der anatomischen Zonen nach verschiedenen Trainingsepochen.....	42
Abbildung 8: Verlauf der Schlüsselmetrik über den Verlauf des Trainings.....	44
Abbildung 9: Beispielsegmentationen von Prostatakarzinomen durch das Modell nach verschiedenen Epochen des Trainings.....	45
Abbildung 10: Boxplots für die Übereinstimmung zwischen der Segmentierung des Modells und Beurteiler 1.....	46
Abbildung 11: Boxplots für die Übereinstimmung zwischen der Segmentierung des Modells und Beurteiler 2.....	47

Tabellenverzeichnis

Tabelle 1: Übereinstimmung der beiden menschlichen Experten bei der Segmentation der anatomischen Prostatazonen und tumorverdächtiger Areale.....	40
Tabelle 2: Übereinstimmung zwischen der Segmentation des Modells und der Beurteiler.....	47
Tabelle 3: Vergleich der Ergebnisse vorheriger Studien mit der hier vorliegenden Arbeit.....	60

Abkürzungsverzeichnis

ADC	apparent diffusion coefficient
AFMS	anteriores fibromuskuläres Stroma
AUC	area under the receiver operating characteristics curve
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften
BPH	benigne Prostatahyperplasie
bpMRT	biparametrische MRT
CNN	Convolutional Neural Networks
CT	Computertomographie
DCE	dynamic contrast enhanced
DICOM	Digital Imaging and Communications in Medicine
DRU	digital rektale Untersuchung
DWI	diffusion weighted imaging
ESUR	European Society of Urogenital Radiology
FSE	Fast-Spin-Echo
FOV	Field-of-View
IQR	Interquartilsabstand
ISUP	International Society of Urological Pathology
LUTS	lower urinary tract symptoms
mpMRT	multiparametrische Magnetresonanztomographie
MRT	Magnetresonanztomographie
PACS	Picture Archiving and Communication System
PCa	Prostatakarzinom
PI-RADS	Prostate Imaging Reporting and Data System
PReLU	Parametric Rectified Linear Unit
PSA	prostataspezifisches Antigen
PZ	periphere Zone
SI	Signalintensität
T	Tesla
T1W	T1 weighted imaging
T2W	T2-weighted imaging
TSE	Turbo-Spin-Echo
TZ	Transitionalzone
ZZ	zentrale Zone

Zusammenfassung

Hintergrund

Die Magnetresonanztomographie (MRT) ermöglicht eine zuverlässige Darstellung von Prostatakrebs (PCa) und kann die Erkennung von Läsionen und die Durchführung gezielter Biopsie verbessern. Die Segmentierung der Prostata im MRT Datensatz ist dabei für viele Aufgaben von entscheidender Bedeutung, u. a. für die Erstellung dreidimensionaler Modelle, z. B. zu Navigationszwecken bei der Planung von Biopsien oder interventionellen Therapien, für die Planung einer Strahlentherapie, für eine verbesserte Volumenschätzung zur Beurteilung des Krankheitsverlaufs und für die automatisierte Erkennung der Anatomie und von PCa. Eine Segmentierung von Hand ist jedoch zeitaufwändig, weshalb eine automatisierte maschinelle Lösung erstrebenswert ist.

Methoden

Es wurde ein Datensatz von insgesamt 158 MRT Untersuchungen der Prostata zusammengestellt, welche den technischen Anforderungen des PI-RADS V2.1 Standards entsprechen. Hierunter befanden sich 102 Patienten mit histologisch gesicherten Prostatakarzinomen und einem bildmorphologischen Befund von PI-RADS 4 oder höher. Die Untersuchungen wurden daraufhin auf einen Trainingsdatensatz und einen Testdatensatz aufgeteilt. Beide Datensätze wurden händisch durch zwei Experten mit mehrjähriger Erfahrung in urologischer Bildgebung segmentiert, wobei zum einen die zonale Anatomie und zum anderen die Tumorregionen annotiert wurden. Des Weiteren wurde ein Deep Learning Modell entwickelt und mit Hilfe des Trainingsdatensatzes auf die Segmentierung der Anatomie und der Tumorregion trainiert. Anschließend wurde am Testdatensatz die Übereinstimmung der Segmentierungen der Experten untereinander sowie die Übereinstimmung der Segmentierungen des Modells mit denen der Experten verglichen.

Ergebnisse

Die Übereinstimmung zwischen den Segmentierungen der beiden Experten war am höchsten für die zentrale Drüse, gefolgt von der peripheren Zone und am niedrigsten für die Tumorregion. Ein ähnliches Bild zeigte sich auch für die Segmentierungen des

Modells. Es bestand kein signifikanter Unterschied in der Übereinstimmung zwischen dem Modell und den jeweiligen Experten 1 und 2. Es konnte jedoch eine schlechtere Übereinstimmung zwischen dem Modell zu den Experten gegenüber der Interrater Übereinstimmung zwischen den Experten festgestellt werden.

Schlussfolgerung

Obgleich das verwendete Deep Learning Modells für die Segmentierung der Prostataanatomie sowie der Segmentierung der Tumorregion nicht ganz den menschlichen Expertenstandard erreichen konnte, lässt sich dennoch eine Perspektive und großes Potential für weitere Forschung und Fortschritte in diesem Bereich der medizinischen Bildanalyse erkennen. Automatisierte Segmentierungen und Tumordetektionen können den klinischen Arbeitsfluss erleichtern und beschleunigen sowie zukünftige Diagnostik und Therapien verbessern. Im Rahmen weiterer technischer Fortschritte ist eine ähnliche Qualität und Sicherheit wie langjährig antrainierte menschliche Experten erwartbar.

Abstract

Background

Magnetic resonance imaging (MRI) provides increasingly reliable imaging of prostate cancer (PCa) and can improve the detection of lesions and the performance of targeted biopsies. In this regard, segmentation of the prostate in the MRI dataset is critical for several tasks, including the creation of three-dimensional models, e.g., for navigational purposes when planning biopsies or interventional therapies, for planning radiotherapy, for improved volume estimation to assess disease progression, and for automated detection of prostate zones and PCa. However, segmentation by hand is very time-consuming, making an automated machine-based solution desirable.

Methods

For this project, a data set of 158 MRI examinations of the prostate was compiled, which meet the technical requirements of the PIRADS V2.1 standard. These included 102 patients with histologically confirmed prostate carcinoma and an image morphological finding of PIRADS 4 or higher. The examinations were then divided into a training data set and a test data set. Both datasets were manually segmented by two subject matter experts with several years of experience in urological imaging, firstly annotating the anatomy and zonal divisions and secondly annotating the tumor regions. Furthermore, a deep learning model was developed and trained on the segmentation of anatomy and tumor region using the training dataset. Subsequently, the agreement of the segmentations of the experts among themselves and the agreement of the segmentations of the model with those of the experts were compared on the test data set.

Results

The agreement between the segmentations of the two experts was highest for the central zone, followed by the peripheral zone and lowest for the tumor region. A similar picture was seen for the segmentations of the model. There was no significant difference in the agreement between the model and the respective experts 1 and 2. However, a worse agreement between the model to the experts compared to the interrater agreement between the experts could be observed.

Conclusion

Although the deep learning model used for this Thesis for prostate anatomy segmentation and tumor region detection and segmentation could not quite reach the human expert standard, a perspective and great potential for further research and progress in this area of medical image analysis can still be seen. Automated segmentations and tumor detection may facilitate and accelerate clinical workflow and improve future diagnostics and therapies. In the context of further technical advances, a similar quality and safety as long-time trained human experts can be expected.

1 Einleitung

1.1 Die Prostata

1.1.1 Anatomie und Funktion

Die Prostata des Menschen ist beim gesunden Erwachsenen ein etwa walnussgroßes Organ, welche sich im Subperitonealraum vom Blasen Hals bis zum Diaphragma pelvis um die Urethra legt (1). Das Volumen der adulten Prostata beträgt etwa 15 ml bei einem normalen Organgewicht von 15 – 20 g (2), welches typischerweise im Alter von 25 – 30 Jahren erreicht wird (3). Sie weist eine konische Form einer invertierten Pyramide auf. Dabei lassen sich eine Basis nahe der Harnblase sowie ein Apex nahe des Beckenbodens abgrenzen (1, 2). Die posteriore Fläche hat breitbasigen Kontakt zur Vorderwand des Rektums, während die anteriore Fläche unmittelbar hinter dem Os pubis liegt. Die inferolateralen Seitenflächen ruhen auf der Faszia des Musculus levator ani (2). Um die Prostata herum findet sich der periprostatice Venenplexus, welcher in die Iliakalvenen drainiert, sowie dorsal der präsakrale Venenplexus, welcher mit lumbisakralen Venenplexus der Wirbelsäule kommuniziert. Der laterale periprostatice Venenplexus wird zudem von Arterien und Nerven durchsetzt, welche gemeinsam das neurovaskuläre Bündel bilden. Dieses liegt posterolateral von kranial nach kaudal der Prostata an (1). Die Urethra tritt an der Basis der Prostata im anterioren Anteil ein und verläuft zunächst nahezu vertikal bis zum Colliculus seminalis, ab wo sie einen deutlich abgeflachten Verlauf nach anterior-inferior nimmt. Im proximalen intraprostatice Verlauf ist die Harnröhre hierbei von einem Sphinkter aus glatter Muskulatur umschieden, während sich an den Apex der Prostata ein hufeisenförmiger Sphinkter aus quergestreifter Skelettmuskulatur anschließt (2, 4). Die Urethra tritt an der Basis der Prostata im anterioren Anteil ein und verläuft zunächst nahezu vertikal bis zum Colliculus seminalis, ab wo sie einen deutlich abgeflachten Verlauf nach anterior nimmt (4). Die Samenbläschen oder auch Bläschendrüsen liegen posterior oberhalb der Basis der Prostata und schmiegen sich teilweise beidseitig posterolateral der Basis an. Ihre Ausführungsgänge (Ductus excretorius) vereinigen sich mit dem jeweiligen Ductus deferens zu den Ductus ejaculatorii, welche am Colliculus seminalis in die pars prostatica der Urethra münden (2).

Bei der Prostata handelt es sich um eine exokrine Drüse und um die größte akzessorische Drüse des Mannes. Sie zeigt einen für Säugetiere typischen gemischt

tubulär-alveolären Aufbau. Neben dem glandulären Epithelzellen besteht das Organ zudem aus einem fibromuskulären Stroma mit einem hohen Anteil an glatten Muskelzellen (5, 6). Eingefasst wird das Organ durch eine Art derber Pseudokapsel, deren Existenz jedoch nicht einheitlicher Konsens ist, da sie sich in manchen Teilen des Organs nicht vom normalen Stroma bzw. der Anheftung und Einbettung des Organs im kleinen Becken trennen lässt (2, 7). Innerhalb der Prostata finden sich etwa 50 einzelne Drüsen, deren Sekret sich mit dem aus den Nebenhoden und Samenbläschen, welches über die Ducti ejaculatorii transportiert wird, in der Harnröhre vermischt. Der Anteil des prostatistischen Sekrets am Gesamtejakulat beträgt in etwa 20 - 30% (1). Darin enthalten sind diverse Enzyme, welche den Samenzellen bei der Befruchtung der Eizelle helfen und dadurch zur Funktionalität des Spermas beitragen. Unter anderem das prostataspezifische Antigen (PSA), eine Serinprotease, welche das Sperma dünnflüssiger macht und damit die Motilität der Spermien unterstützt (6). Während der Ejakulation verhindert der proximale urethrale Sphinkter aus glatter Muskulatur das Auftreten einer retrograden Ejakulation in die Harnblase, während der distale, hufeisenförmige Sphinkter aus Skelettmuskulatur maßgeblich für die Harnkontinenz verantwortlich ist (2, 3, 6, 8). Die innere Struktur der Prostata wurde historisch auf unterschiedliche Weise betrachtet und in verschiedenen Modellen beschrieben. Zunächst wurde nach einer Publikation des amerikanischen Chirurgen Oswald Lowsley aus dem Jahre 1912 von einer Aufteilung in 5 Lappen ausgegangen. Nämlich in einen anterioren, einen mittleren, einen posterioren und zwei laterale Lappen (9). Die Studien, die von Lowsley zu diesem Modell durchgeführt wurden, enthielten jedoch lediglich fetale Prostaten und keine adulten Organe, an welchen sich der Aufbau in 5 Lappen histologisch nicht mehr nachvollziehen lässt (5, 8). Ein Weiterer und noch heute akzeptierter Ansatz war die Einteilung nach Zonen in ihrer jeweiligen Relation zur Urethra als zentralem Referenzpunkt. Diese wurde von dem Pathologen John E. McNeal im Jahr 1968 erstmalig beschrieben und im weiteren Verlauf über mehrere Jahrzehnte in einigen Publikationen überarbeitet. Nach diesem Modell ist die Prostata aufgebaut aus einer peripheren Zone (PZ), welche im gesunden etwa 70% des Drüsenvolumens ausmacht und wie eine Scheibe außen das restliche Organ ummantelt. Zudem aus einer zentralen Zone (ZZ), die sich unmittelbar um die Urethra legt, dabei die Ductus ejaculatorii umschließt und etwa 25% des Organvolumens einnimmt. Darüber hinaus aus einer sehr kleinen Transitionalzone (TZ), welche sich in Form von zwei kleinen Lappen zwischen

Harnblase und Colliculus seminalis ventral und lateral um die Urethra schließt und lediglich 5% der Drüse darstellt. Die Transitionalzone weist dabei einen geringeren Drüsenanteil und dafür ein sehr viel dichteres Stroma auf und stellt embryologisch eine Transition zwischen dem vesikosphinktären System der Urethra und dem glandulären System der zentralen und peripheren Prostatazonen dar. (8, 10-12). Zuletzt wurde darüber hinaus eine ventrale Verdickung der Organkapsel beschrieben, welche als anteriores fibromuskuläres Stroma (AFMS) bezeichnet wird. Dieses besteht in Richtung des Apex primär aus Skelettmuskulatur, welche sich mit dem Beckenbogen verbinden, während sich zur Basis hin vermehrt glatte Muskulatur findet, die in den Harnblasenhals übergeht (3, 13). Die zonale Gliederung der Prostata (Abb. 1) ist heutzutage weithin akzeptiert und wird im klinischen Alltag verwendet und findet ihre Relevanz unter anderem in der Einordnung und Therapie von Erkrankungen, da diese unterschiedliche Prädilektionsstellen innerhalb der Prostata aufweisen (14).

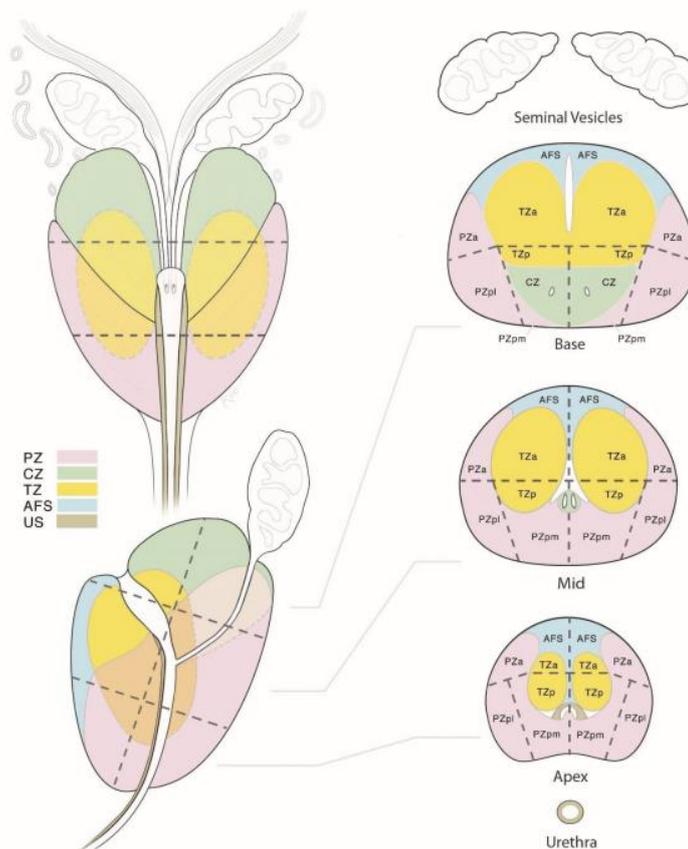


Abbildung 1. Schematische Darstellung der zonalen Gliederung der Prostata. PZ = periphere Zone, CZ = zentrale Zone, TZ = Transitionalzone, AFS = anteriores fibromuskuläres Stroma, US = urethraler Sphinkter. Entnommen aus American College of Radiology Committee on PI-RADS®. PI-RADS Assessment Categories version2.1. Abrufbar unter <https://www.acr.org/-/media/ACR/Files/RADS/PI-RADS/PI-RADS-V2-1.pdf>. Abgerufen am 24.09.2021

1.1.2 Erkrankungen und klinische Bedeutung

Die zonale Gliederung der Prostata nach McNeal bezieht sich neben der relativen Lage der einzelnen Zonen in Relation zur Urethra auch auf den histologischen Aufbau der Drüsenanteile und des Stromas sowie ihre embryologische Herkunft. Hiermit einhergehend wurde in den Arbeiten zusätzlich ein unterschiedliches Vorkommen der relevantesten prostatistischen Pathologien in den unterschiedlichen Zonen beschrieben (10-13). Glanduläre und stromale Anteile des periurethralen Gewebes und der TZ zeigen im Laufe des Lebens nahezu bei allen Männern eine Tendenz zur mikroskopischen Hyperplasie mit Entwicklung benigner knotiger Veränderungen, weshalb dieser Prozess als benigne (noduläre) Prostatahyperplasie (BPH) bezeichnet wurde (13). Dieses Geschehen beginnt in der Regel um das 40. Lebensjahr und schreitet stetig fort. In der vierten Lebensdekade weisen etwa 10% der Männer mikroskopische Anzeichen einer BPH auf, in der fünften Lebensdekade 20% und in der achten und neunten dann bis zu 90% (15). Durch diese Veränderung wächst die TZ im Verlauf der Zeit stetig und nimmt immer mehr Anteil des Drüsenvolumens in Anspruch. Hierdurch wird die ZZ nach dorsal und medial komprimiert. Ebenso wird auch die PZ immer weiter aufgespannt und verdrängt, was unter anderem die Abgrenzbarkeit in der Bildgebung erschwert (1, 16). Des Weiteren zeigt etwa die Hälfte der Männer irgendwann auch eine makroskopische BPH mit Vergrößerung des gesamten Organs. Dabei kommt es des Öfteren im Rahmen der Auftreibung der TZ zu einer Kompression der Urethra mit entsprechender Einschränkung des Harnflusses, was häufig langfristig zu Symptomen beim Urinieren führt. In diesem Kontext wird dann von einer Progression der makroskopischen BPH in eine klinische BPH gesprochen. Die Einengung der Urethra wird zum einen statisch durch die Verdickung des prostatistischen Gewebes und zum anderen dynamisch durch die vermehrte Spannung der glatten Muskulatur verursacht (17). Die entstehenden Symptome werden unter dem Überbegriff lower urinary tract symptoms (LUTS) zusammengefasst und beinhalten eine verzögerte Miktion, reduzierten Harnfluss während der Miktion mit möglicherweise Pausieren während der Miktion und Notwendigkeit eines Anspannens der Bauchmuskulatur sowie ein ebenso verzögertes Ende des Harnflusses mit Nachtröpfeln. Außerdem kann es zu einer Restharnbildung und zuletzt langfristig auch zu einem Harnverhalt und einer Überlaufinkontinenz kommen. Die Genese der BPH ist bisher nicht hinreichend geklärt und es gibt hierzu verschiedene

Theorien, wobei die Hypothese einer hormonellen Ursache unter Einfluss von Dihydrotestosteron (DHT) die prävalenteste zu sein scheint (15). Es wird davon ausgegangen, dass es im Alter zu einer Veränderung des Androgenstoffwechsels der Prostata kommt, was zu einer Akkumulation von DHT führt (18). Die Behandlung der BPH wird hauptsächlich von den klinischen Symptomen geleitet und kann von einer medikamentösen Behandlung mit Alpha-1-Rezeptorblocker zur Relaxation der glatten Muskelfasern oder 5-Alpha-Reduktasehemmern zur Blockade der Bildung von DHT bis hin zu lokal ablativen und chirurgischen Verfahren reichen (17, 19). Auf Grund der sehr hohen Prävalenz in der alternden, männlichen Bevölkerung stellt die klinische BPH nicht nur eine große Bürde für die Betroffenen, sondern auch für die Gesundheitssysteme dieser Welt dar. In den USA beispielsweise werden die jährlichen Kosten in Zusammenhang mit der Behandlung der BPH auf etwa 4 Milliarden Dollar geschätzt (20). Während die BPH ausschließlich in den stromalen und glandulären Anteilen des periurethralen Gewebes und der TZ stattfindet, zeigen sich andere Prozesse eher in der PZ. So ist die Prostatitis deutlich überwiegend eine Erkrankung der PZ, wobei vereinzelt auch die ZZ betroffen sein kann. Es wird jedoch teilweise davon ausgegangen, dass es sich hierbei um eine sekundäre Beteiligung bei Primärmanifestation in der PZ handelt (10, 14, 21). Bei der Prostatitis handelt es sich ebenfalls um eine häufig auftretende Erkrankung des Mannes, welche mit zunehmendem Alter immer öfter vorkommt. Sie macht etwa 25% der Arztbesuche im Zusammenhang mit urogenitalen Beschwerden aus und ist die häufigste urologische Diagnose bei Männern unter 50 Jahre sowie die dritthäufigste bei Männern über 50 Jahre (22). In einer Autopsiestudie von McNeal zeigten sich in 91 Autopsiepräparaten von Erwachsenen in 40 Fällen Anzeichen einer Prostatitis, wobei 21 als mild und 19 als schwergradig eingestuft wurden. Nur in zwei der 40 Fälle mit Prostatitis war die ZZ primär betroffen, wohingegen 24 Fälle allein die PZ betrafen. In 14 Fällen waren beide Zonen beteiligt, wobei hier die Entzündung in der ZZ deutlich geringer ausgeprägt war als in der PZ und mit dem affektierten Areal der PZ in unmittelbarer Verbindung stand, was eine sekundäre Begleitreaktion wahrscheinlich macht (10). Der Grund für dieses vermehrte Vorkommen der Prostatitis in der PZ lässt sich am ehesten mit anatomischen Faktoren in der Relation der Drüsenverläufe zur Urethra und damit verbundenen Flussdynamiken und Refluxmöglichkeiten erklären (21). Die zu Grunde liegenden Ursachen für die Prostatitis können unterschiedlich sein und es hat sich eine Einteilung in verschiedene Entitäten etabliert. Nach Meares et al. sind diese

die akute und die chronische bakterielle Prostatitis, die chronische nichtbakterielle Prostatitis und die Prostatodynie. Die Einteilung basiert auf dem Auftreten und der Schwere der klinischen Symptome sowie dem Nachweis oder der Abwesenheit von Leukozyten und Erregern in verschiedenen Sekreten wie Urin, Samenflüssigkeit oder exprimiertem Prostatasekret (23). Im Laufe der Zeit zeigten sich jedoch immer wieder Unstimmigkeiten zwischen der Einteilung nach Meares und alltäglichen klinischen Befunden wie z.B. atypischen Erregern in Kulturen sowie Nachweis von Biofilmen aus Bakterien in den Azini der Prostata in Biopsiepräparaten bei Patienten mit zuvor negativen Kulturen aus Sekreten. Daher wurde durch das NIH eine überarbeitete Einteilung herausgegeben, die neben der akuten und chronischen bakteriellen sowie der chronischen nichtbakteriellen Prostatitis noch ein inflammatorisches und nicht inflammatorisches chronisches Beckenschmerzsyndrom und eine asymptomatische chronische Prostatitis nennt (22). Die Therapie der akuten bakteriellen Prostatitis besteht meist aus Breitspektrumantibiotika und ggf. einer chirurgischen Intervention akuter Komplikationen, z.B. einer Spaltung oder Drainage bei Abszessbildung und ggf. einer suprapubischen Harnableitung. Die chronische bakterielle Prostatitis wird je nach klinischem Verlauf mittels kurzfristiger Antibiotikagabe bei initialer Phase bis hin zu langfristiger, prophylaktischer antibiotischer Abschirmung bei rekurrenten Verläufen behandelt. In schweren, therapierefraktären Fällen kann auch eine chirurgische Sanierung mittels Prostatektomie notwendig sein. Die chronischen Beckenschmerzsyndrome werden entweder mit einem initialem Antibiotika Therapieversuch angegangen oder mittels supportiver Behandlungsverfahren wie Antiphlogistika und Schmerzmedikation, Alpha-1-Rezeptorhemmer, Alpha-Dehydrogenase-Hemmer, Lifestyle Veränderungen sowie psychosomatischer Therapie, Biofeedback und Entspannungsübungen. Die asymptomatische Prostatitis bedarf zwar eigentlich keiner zwingenden Therapie, kann jedoch testweise mit einer Antibiotikagabe angegangen werden, wenn sich auf Grund der Inflammation ein erhöhter PSA-Wert im Blut zeigt (22). Die asymptomatische chronische Prostatitis ist dabei die häufigste benigne Ursache eines erhöhten PSA-Wertes, welche in Prostatabiopsien nachgewiesen wird, die zur Abklärung dieses Markers durchgeführt werden (24).

In Abgrenzung zum benignen Befund der Prostatitis steht das Prostatakarzinom als häufigster maligner Grund einer PSA-Wert Erhöhung im Blut. Das Prostatakarzinom ist in den meisten westlichen Ländern der häufigste maligne Tumor des Mannes und

zugleich die zweithäufigste Krebstodesursache. In Deutschland macht es bei Männern etwa 23% aller Krebserkrankungen und 12% der Sterbefälle in Zusammenhang mit Krebs aus (25, 26). Das Risiko an einem Prostatakarzinom zu erkranken, steigt mit zunehmendem Alter. Das durchschnittliche Erkrankungsalter liegt in Deutschland bei etwa 72 Jahren. Eine Diagnose vor dem 50. Lebensjahr ist äußerst selten. Mit einem Alter von 85 Jahren hingegen liegt das kumulative Risiko an Prostatakrebs zu erkranken weltweit zwischen 0,5 und 20%, wobei Autopsiestudien zeigen, dass sogar die Mehrheit der Männer dieses Alters ein histologisch nachweisbares Prostatakarzinom haben. Die Schwankungen sind hierbei geographisch bezogen sehr hoch, wobei die niedrigsten Inzidenzen im asiatischen Raum zu finden sind und die höchsten in Nordamerika auftreten, hierbei vor allem unter afroamerikanischen Männern. Neben dem Alter und der ethnischen Herkunft scheinen auch genetische Faktoren, Umwelteinflüsse, hormonelle Gegebenheiten und die Ernährung eine Rolle beim Erkrankungsrisiko zu spielen (27, 28). Auf Grund verbesserter Therapiemöglichkeiten und teils früherer Diagnose sowie Diagnosestellung in weniger fortgeschrittenen Stadien der Erkrankung steigt die durchschnittliche Überlebenszeit mit der Erkrankung. Diese Tatsachen und die zunehmende Alterung der Gesellschaft lassen in den nächsten Jahren einen Anstieg der Inzidenz und der Prävalenz des Prostatakarzinoms erwarten (26).

Die initiale Diagnostik des Prostatakarzinoms erfolgt in der Regel über eine digital rektale Untersuchung (DRU) sowie die Bestimmung des PSA-Wertes im Blut. Bei einem auffälligen Ergebnis dieser Untersuchungen sollte nach neuester S3 Leitlinie eine multiparametrische Magnetresonanztomographie (mpMRT) der Prostata erfolgen. Bei suspektem Befund sollte eine anschließende Biopsie erfolgen. Möglich sind hierbei eine transrektale, gezielte Ultraschall-gestützte Biopsie oder eine MRT-gestützte Biopsie, die jeweils durch eine systematische transperineale Biopsie mit 12 Stanzen ergänzt werden sollten (29).

Die Sinnhaftigkeit eines breitflächigen Screenings aller Männer mittels PSA-Kontrolle wird insgesamt in Frage gestellt. Gemäß einer Metaanalyse der Chochrane Library zeigte sich kein statistisch signifikanter Unterschied in der Prostatakarzinom-spezifischen Mortalität und der allgemeinen Mortalität zwischen der Screening-Gruppe und der Kontrollgruppe (30). Zudem müssen die Risiken der weiteren Diagnostik wie Biopsien sowie Überdiagnosen und Übertherapien mit in die Abwägung einbezogen werden (31). Gemäß den Empfehlungen der American Cancer Society sollten Patienten ab dem 50.

Lebensjahr bei einer Restlebenserwartung von über 10 Jahren über mögliche Risiken, Unsicherheiten und Vorteile eines Screenings auf Prostatakrebs aufgeklärt werden. Für besondere Risikogruppen mit hoher familiärer Vorbelastung sollte diese Diskussion ggf. etwas früher mit 40 oder 45 Jahren erfolgen (28, 32). Ebenso empfiehlt auch die European Urological Society eine individuelle, risikoadaptierte Strategie für gut informierte Patienten mit einer verbleibenden Mindestlebenserwartung von 10 bis 15 Jahren (33).

Unter den malignen Tumorentitäten, welche in der Prostata auftreten können, hat das azinäre Adenokarzinom den mit Abstand größten Anteil. Dem gegenüber stehen seltenere histologische Subtypen wie das ductale Adenokarzinom, das squamöse Adenokarzinom, das klarzellige Karzinom, das urotheliale Karzinom, das kleinzellige Karzinom und das sakomatöse Karzinom (34). Auch das Prostatakarzinom zeigt eine deutliche zonale Prädilektion und tritt zumeist in der PZ auf. Hier finden sich etwa 70% aller Prostatakarzinome, möglicherweise auf Grund des embryonalen Ursprungs aus den Zellen des Sinus urogenitalis. Währenddessen finden sich in der ZZ nur 5 % aller Prostatakarzinome, was neben der geringeren Größe im Vergleich zur PZ vermutlich auch am embryonalen Ursprung der ZZ aus Zellen des Wolff-Ganges liegt. Aus der TZ gehen etwa 25% der Prostatakarzinome hervor, was sich zum einen aus der initial sehr geringen Größe in Relation zur PZ und zum anderen aus dem relativ hohen Stromalen Anteil im Vergleich zum glandulären Gewebe erklärt (2, 35).

Histopathologisch werden Prostatakarzinome gemäß der International Society of Urologic Pathology (ISUP) nach dem Gleason-Score eingestuft, wobei das Gewebe auf seine Drüsenmorphologie hin untersucht wird und dabei gemäß dem Ausmaß an Entdifferenzierung in 5 Gleason-Grade eingeordnet. Der Gleason-Score wird anschließend berechnet, indem bei Biopsiepräparaten der am häufigsten vorkommende mit dem höchsten Gleason-Grad addiert wird, während bei Prostatektomiepräparaten der häufigste mit dem zweithäufigsten zusammengerechnet wird. Es ergeben sich daraus also Werte, die von einem Gleason-Score 2 (1 + 1) bis hin zu einem Gleason-Score 10 (5 + 5) reichen. Außerdem wird der ISUP-Grad für das Grading bestimmt, der den Tumor gemäß dem Gleason-Score in 5 mögliche Grade einteilt (33, 36). Das Grading entscheidet letztlich maßgeblich in Zusammenschau mit der Konstitution und generellen Lebenserwartung des Patienten über das weitere Prozedere im Hinblick auf Staging und Therapie (33, 37). Das Spektrum an möglichen Therapiestrategien reicht von watchful

waiting oder active surveillance über Prostatektomie, Strahlentherapie, Brachytherapie oder antihormoneller Therapie bis hin zu systemischen Chemotherapien bei metastasiertem Prostatakarzinom (33).

1.2 Multiparametrische MRT der Prostata

1.2.1 Überblick

Die Magnetresonanztomographie (MRT) spielt in der Diagnostik des Prostatakarzinoms eine immer mehr zunehmende Rolle. Dies ist vor allem technischen Weiterentwicklungen und steigender diagnostischer Genauigkeit und Qualität geschuldet. An modernen MRT-Scannern mit starken Gradienten und Mehrkanal-Oberflächenspulen kann die Untersuchung heutzutage bei den gängigen Feldstärken von 1,5T und 3T unter Erfüllung der gewünschten Qualitätsstandards erfolgen, wobei jedoch die Untersuchung bei 3T auf Grund es günstigeren Signal-zu-Rausch-Verhältnisses und der kürzeren Untersuchungsdauer zu bevorzugen ist (38). Gerade erst kürzlich wurde die S3 Leitlinie zum Prostatakarzinom in ihrer Auflage 6.1 überarbeitet und die multiparametrische MRT (mpMRT) der Prostata bei Verdacht auf ein Prostatakarzinom auf Grund eines erhöhten PSA-Wertes oder einer auffälligen DRU bereits vor Durchführung einer systematischen Biopsie als primäre Diagnostik empfohlen (29). Die Aufgabe der mpMRT bei der Diagnostik des Prostatakarzinoms erstreckt sich über die Beurteilung der Anatomie des Organs und der zonalen Gliederung über eine Untersuchung der Beschaffenheit des Gewebes und funktionaler Charakteristika wie Zelldichte und Durchblutung (38, 39). Hierbei kann die mpMRT insbesondere in der Krebsdiagnostik helfen bei der Detektion, der Angabe der Lokalisation, dem lokalen Tumorstaging und der Einschätzung der Aggressivität. Zudem können auch nichtmaligne Veränderungen wie die BPH und die Prostatitis abgebildet und differenziert werden (35, 40). Die mpMRT der Prostata besteht dabei für gewöhnlich aus einer hochaufgelösten T2-gewichteten Sequenz in axialer und sagittaler oder auch koronarer Schichtführung sowie zwei funktionalen MRT Techniken. Hierbei wird meist eine diffusionsgewichtete Sequenz in axialer Schichtführung sowie dynamische Kontrastmittel-gestützte Sequenzen ebenfalls in axialer Schichtführung durchgeführt (39). Zudem sind weitere funktionelle Bildgebungstechniken möglich, welche auch weitläufig eingesetzt werden, jedoch bisher nicht ihren Weg in die gängigen Qualitätsstandards und Leitlinien geschafft haben. Hierzu gehören unter anderem die MR-Spektroskopie und T1- bzw. T2-Mapping Sequenzen (41).

1.2.2 T2-gewichtete Bildgebung

Die T2-gewichtete Bildgebung (T2W) im Rahmen der mpMRT der Prostata erfolgt nach den meisten, geltenden Qualitätsstandards als hochaufgelöste Fast-Spin-Echo (FSE) oder Turbo-Spin-Echo (TSE) Sequenz mit kleinem Field-of-View (FOV) von 12 – 20 cm unter Einschluss der Prostata, der Samenbläschen und des externen Sphinkters mit einer Schichtdicke von ≤ 3 mm ohne Schichtlücke und einer in-plane Auflösung von maximal $0,7 * 0,4$ mm (35, 39, 41). Die axiale Schichtführung kann hierbei unterschiedlich gewinkelt werden. Gängig sind hierbei die Winkelung entlang der Urethra oder der Vorderwand des Rektums sowie eine streng axiale Schichtung (35). Die T2-gewichtete Bildgebung bietet eine exzellente Darstellung der Anatomie der Prostata sowie eine sehr gute Abgrenzbarkeit der zonalen Gliederung mit hervorragendem Weichteilkontrast. Ebenso lassen sich die kapsuläre Begrenzung und die umgebenden Strukturen wie das neurovaskuläre Bündel und die Samenbläschen beurteilen (39, 41, 42). Die Sequenz spiegelt den Wassergehalt des Gewebes und damit indirekt auch die Zellularität wider und ist somit gut geeignet für die Beurteilung der PZ, welche sich im Normalzustand wegen der hohen Drüsendichte und deren hohen Wassergehaltes hyperintens abbildet (41). In der PZ finden sich ein Großteil der Tumoren und diese können somit gut morphologisch in der T2-gewichteten Bildgebung erkannt und lokalisiert werden und es können Aussagen über kapselüberschreitendes Wachstum getroffen werden (39, 43, 44). Karzinome der PZ weisen einen geringeren Wassergehalt auf und stellen sich somit typischerweise hypointens in der T2-gewichteten Bildgebung dar und zeigen meist eine rundliche oder unscharf berandete Struktur. Das Ausmaß der Hypointensität kann hierbei auch gewisse Rückschlüsse auf die Aggressivität des Tumors erlauben (39, 41). Leider finden sich derartige Signalalterationen der PZ auch bei anderen, nicht malignen Prozessen, wie z.B. Narben, der Prostatitis, einer Einblutung nach Biopsie oder einer Atrophie, so dass eine Differenzierung allein Anhand der T2-Bildgebung erschwert ist. Auch für die Beurteilung der TZ sind einige herausfordernde Gegebenheiten zu berücksichtigen. So gewinnt die TZ im Laufe des Lebens zunehmend an Volumen, wie bereits im vorangegangenen Kapitel erläutert, und formt auf Grund der stromalen und glandulären Hyperplasie knotige, in ihrer Signalintensität heterogene Veränderungen. Diese BPH Knoten lassen sich in der T2-Bildgebung sehr gut abgrenzen, eine Detektion zwischengelagerter Tumore ist jedoch auf Grund der Heterogenität erschwert. Kriterien für den Verdacht eines Karzinoms sind eine umschriebene, homogen hypointense Läsion

der TZ mit einem oft verwascheneren, unschärferen Rand im Vergleich zu den klar umrandeten BPH Knoten (39, 42, 44). Tumore der TZ zeigen häufig bei Diagnose bereits eine Infiltration der Organkapsel oder des anterioren fibromuskulären Stromas. Die Infiltration der Kapsel kann anhand der T2-Bildgebung ebenfalls evaluiert werden und zeigt sich in Form einer Vorwölbung oder Irregularität der Kapsel sowie durch messbare knotige Tumorinvasion in das umgebende Gewebe, beispielsweise erkennbar durch eine Verlegung des rektoprostatistischen Winkels oder eine Verdickung des neurovaskulären Bündels sowie eine Auftreibung und Signalalteration der Samenbläschen (35, 39). Auch hier können jedoch andere Veränderungen wie z.B. ein prolabierter BPH Knoten oder narbige Veränderungen ein kapselüberschreitendes Wachstum vortäuschen (44). Insgesamt ist die T2-Bildgebung sehr gut zur anatomischen Beurteilung der Prostata und zur morphologischen Einordnung pathologischer Veränderungen geeignet und zeigt dabei z.B. eine hohe Sensitivität für die Detektion von Prostatakarzinomen. Die Spezifität ist dabei allerdings eher gering, weshalb die Verwendung zusätzlicher funktionaler Sequenzen zur näheren Charakterisierung der detektierten Veränderungen empfohlen wird (35, 39).

1.2.3 Diffusionsgewichtete Bildgebung

Die diffusionsgewichtete Bildgebung oder auch diffusion-weighted-imaging (DWI) Sequenz ist eine Technik, bei der die willkürliche Bewegung der Wassermoleküle, die sogenannte Brownsche Molekularbewegung, im Gewebe widergespiegelt wird. Hierfür wird eine Reihe magnetischer Gradienten angelegt, deren Amplitude und Dauer als B-Werte bezeichnet werden. Für die mpMRT der Prostata wird von den meisten Fachgesellschaften die Verwendung von mindestens vier B-Werten von 0 bis 1000 sec/mm^2 empfohlen mit der anschließenden Berechnung einer apparent diffusion coefficient (ADC) Map (35, 44). Des Weiteren wird dazu geraten, aus akquirierten B-Werten eine Serie mit einem errechneten hohen B-Wert von mindestens 1400 cm/mm^2 zu generieren.

Gemäß den technischen Anforderungen sollte die DWI-Sequenz als Spinecho EPI Sequenz kombiniert mit spektraler Fettsättigung gemessen werden. Die Vorgaben sehen hierbei eine möglichst kurze TE von ≤ 90 ms bei einer TR von ≥ 3000 ms vor sowie eine Schichtdicke von ≤ 4 mm mit lückenloser Akquisition bei gleicher Schichtführung wie in

der T2-Bildgebung. Das FOV sollte zwischen 16 und 22 cm liegen und die in plane Auflösung bei $\leq 2,5$ mm für Phase und Frequenz (35).

Die ADC-Map ist eine quantitative Abbildung der Verschiebung der Wassermoleküle innerhalb des Gewebes bzw. zwischen Geweben und berechnet sich aus zwei oder mehr B-Wert Bildern.

Eine Restriktion der freien Diffusion von Wassermolekülen im Gewebe ist hierbei umgekehrt proportional zu der Zelldichte des Gewebes und der Integrität der Zellmembranen (45). Diese stellt sich in der DWI als hyperintenses Signal dar, was jedoch teilweise schwer von anderen in T2W hyperintensiven Strukturen wie freiem Wasser zu unterscheiden ist, da die zu Grunde liegende Spinecho-Sequenz der DWI T2-gewichtet ist. Durch hohe B-Werte lässt sich das Hintergrundsignal der Prostata unterdrücken und in der ADC-Map lässt sich dieser „T2-shine-through“-Effekt von einer tatsächlichen Restriktion der freien Diffusion unterscheiden, da letztere hier zu einer Signalabsenkung führt (45, 46). Diverse Veränderungen der Prostata können zu einer Einschränkung der Diffusion im Gewebe führen. So können beispielsweise BPH-Knoten mit höherer Zelldichte eine Diffusionsrestriktion aufweisen. Ebenso akut entzündliche Veränderungen im Rahmen einer floriden Prostatitis, bei denen es zu Zellschwellung und zusätzlicher Infiltration durch Entzündungszellen kommt. Auch Abszesse können auf Grund der hohen Zelldichte und des Detritus im Eiter eine deutliche Einschränkung der Diffusion und damit eine Signalminderung in der ADC-Map zeigen. Darüber hinaus herrscht in der ZZ gegenüber den übrigen Zonen der Prostata physiologischerweise eine verminderte Diffusion (46). Prostatakarzinome weisen auf Grund ihrer hohen Zellularität eine Diffusionsrestriktion auf, wobei die ADC-Werte eine inverse Korrelation zum Gleason-Grad des Tumors haben und damit die Aggressivität des Karzinoms abbilden können (45, 47). Für die Detektion und Einordnung von karzinomsuspekten Läsionen ist die diffusionsgewichtete Bildgebung daher eine hilfreiche und wichtige Ergänzung zur T2W mit deutlicher Steigerung der Sensitivität und Spezifität bei biparametrischer MRT aus DWI und T2W im Vergleich zur T2W allein (48, 49).

1.2.4 DCE – dynamisches Kontrastmittelenhancement

Als weiterer Baustein der funktionalen Charakterisierung des Prostatagewebes in der mpMRT dient das dynamische Kontrastmittelenhancement (DCE). Das DCE zeigt die Vaskularisierung des Gewebes, indem es mittels schneller T1-gewichteter Sequenzen

nach Gabe von gadoliniumhaltigem Kontrastmittel die Anreicherung desselben als Signalanhebung darstellt. Dadurch kann die Perfusion und vaskuläre Permeabilität innerhalb der Prostata dargestellt werden (44). Laut gängigen technischen Empfehlungen wird die DCE als 2D oder 2D Gradientenecho-Sequenz mit einer TR von < 100 ms und einer TE von < 5 ms durchgeführt. Die Schichtdicke sollte hierbei ≤ 3 mm betragen bei lückenloser Akquisition in gleicher Schichtführung wie in der T2W und DWI. Die In-Plane-Auflösung sollte bei ≤ 2 mm für Phase und Frequenz liegen bei einem FOV, welches das gesamte Organ sowie die Samenbläschen einschließt. Für die zeitliche Auflösung wird ein Wert von ≤ 15 Sekunden mit einer Gesamtdauer von ≥ 2 Minuten empfohlen. Für eine optimale Anreicherung des Kontrastmittels sollte die applizierte Dosis des gadoliniumhaltigen Kontrastmittels etwa $0,1$ mmol/kg Körpergewicht betragen und mit einer Flussrate von $2-3$ ml/s injiziert werden (35, 39).

Das DCE kann qualitativ beurteilt werden und als Einzelbilder oder Cine-Modus betrachtet werden, oder es kann anhand der Signalintensität (SI) über die Zeit auch semi-quantitativ und quantitativ analysiert werden (50-53). Betrachtet werden hierbei die Geschwindigkeit und die Intensität der Kontrastmittelaufnahme sowie auch das Auswaschen des Kontrastmittels aus dem Gewebe. Darüber lässt sich beispielsweise die Neoangiogenese eines Tumors darstellen mit seiner veränderten Mikrovaskularisierung mit erhöhter Permeabilität der Basalmembranen. Karzinome weisen daher typischerweise eine sehr frühe und rasche Kontrastmittelaufnahme und ein ebenfalls schnelles Auswaschen auf, wobei es auch hier Ausnahmen von der Regel gibt. Es wurde jedoch gezeigt, dass die Dichte der Mikrovaskularisierung einen prognostischen Faktor für Prostatakarzinome darstellt und auch ein Hinweis auf die Aggressivität des Tumors bzw. seinen Gleason-Score sein kann (52-55). Dennoch gibt es wie auch bei den übrigen verwendeten Sequenzen Fallstricke durch benigne Veränderungen, welche ein ähnliches Kontrastmittelverhalten wie maligne Tumore zeigen. Dazu gehören beispielsweise entzündliche Läsionen im Rahmen einer Prostatitis wie auch manche fibromuskuläre und stromale BPH-Knoten. Aus diesem Grunde sollte die DCE immer in Zusammenschau mit der T2W und der DWI betrachtet werden (44, 51, 52). Die Sensitivität und Spezifität in der Detektion und Lokalisation werden durch die Zunahme der DCE gesteigert. Besonders ausgeprägt ist dieser Effekt bei Rezidiven nach erfolgter Operation und Bestrahlung, da fibrotische und narbige Veränderungen mit stark hypointenser Signalgebung in der T2W die Diagnostik hier erschweren können (51, 56).

1.2.5 Klinische Anwendung

Für die Beurteilung der mpMRT der Prostata hat die europäische urologische Gesellschaft (ESUR) gemeinsam mit dem American College of Radiology (ACR) ein strukturiertes Befundungsschema herausgegeben, welches seit 2019 inzwischen in seiner modifizierten zweiten Version vorliegt. Das Prostate Imaging – Reporting and Data System (PI-RADS) geht dabei zum einen auf technische Empfehlungen und Qualitätsstandards für die Untersuchung ein und zum anderen auf die Beschreibung und strukturierte Einordnung der identifizierten Befunde (35). Hierbei werden die Charakteristika in den Sequenzen T2W, DWI und DCE jeweils anhand einer 5-Stufigen Skala bewertet und anschließend gewichtet für eine finale Einordnung auf derselben Skala. Hierbei fallen die einzelnen Sequenzen unterschiedlich ins Gewicht, je nachdem, welche Zone der Prostata bewertet wird. In jeder Zone gibt es eine aussagekräftigste und damit führende Sequenz für die finale Einordnung. In der PZ ist die entscheidende Sequenz die DWI, während es in der TZ die T2W ist. Die 5-Stufige Skala bildet die Wahrscheinlichkeit für das Vorliegen eines signifikanten Karzinoms ab, welches definiert wird über einen Gleason-Score ≥ 7 (inklusive 3 +4), einem Volumen von $> 0,5$ ml oder einem extraprostatatischen Wachstum. Die DCE nimmt in der aktuellen Version 2.1 des PI-RADS nur eine untergeordnete Rolle ein, und zwar in der Differenzierung zwischen einer Stufe 3 und 4 Läsion der PZ. Sie bewirkt also im Falle eines suspekten Kontrastmittelverhaltens ein Upgrading von einer nicht eindeutigen Läsion mit intermediärem Risiko zu einer malignitätssuspekten Läsion mit hohem Risiko (35). Da Kontrastmittel mit relativ hohen Kosten und einem gewissen Risiko für unerwünschte Nebenwirkungen verbunden sind, wächst zunehmend die Tendenz, auf die Gabe gadoliniumhaltiger Kontrastmittel zu verzichten und stattdessen lediglich eine biparametrische MRT (bpMRT) aus T2W und DWI durchzuführen (35, 57). Die Studienlage hierzu ist heterogen und während einige Veröffentlichungen eine ähnliche Sensitivität und Spezifität für die bpMRT im Vergleich zur mpMRT zeigen (40, 58), zeigen andere weiterhin Vorteile für den Gebrauch der DCE (52, 59, 60). Zumindest in der primären Bildgebung vor Biopsie sowie in aktiver Überwachung bietet sich wohl insgesamt die Verwendung der bpMRT an, wohingegen nach vorangegangener negativer Biopsie oder nach Therapie zur Rezidivdiagnostik bzw. zur Beurteilung des Therapieerfolges die mpMRT weiterhin vorzuziehen ist. Des Weiteren kann bei multifokalen Geschehen ggf. die Läsion mit der höchsten Malignitätswahrscheinlichkeit

für eine bildgestützte, gezielte Biopsie ausgewählt werden (35, 52, 57). Insgesamt sollte hierbei nach erfolgter Biopsie oder Therapie jedoch 4-6 Wochen vor der Durchführung der Untersuchung abgewartet werden, da eventuelle Blutabbauprodukte zu Artefakten und entsprechenden Fehleinschätzungen führen können (35).

1.3 Medizinische Bildanalyse

1.3.1 Grundlagen

Die medizinische Bildanalyse kann als Teil der klassischen Bildverarbeitung sowie als Untersparte der Computer Vision bzw. des computergestützten Sehens angesehen werden. Dieses Feld hat in den letzten Jahrzehnten zunehmend an Bedeutung gewonnen, da durch die breitflächige Einführung digitaler Bildarchive (PACS – Picture Archiving and Communication System) die technischen Voraussetzungen ermöglicht wurden (61). Es können verschiedene Analyseverfahren unterschieden werden, wie beispielsweise die Segmentierung sowie die Mustererkennung und Klassifizierung von Objekten. Der typische Ablauf des Bildanalyseprozesses ist dabei in der Regel, dass zunächst eine Bildvorverarbeitung stattfindet, gefolgt von einer Bildsegmentierung sowie einer anschließenden Merkmalsextraktion und Klassifizierung, die schlussendlich eine Aussage über das Bild ermöglichen (62-64).

Die Segmentierung stellt hierbei einen essenziellen Schritt in der Verarbeitung dar und beinhaltet eine Zerlegung des Bildes in semantische Einheiten, wobei jedem Pixel ein Label bzw. eine Klasse zugeordnet wird. Die entstehenden Segmente sind die Bedeutungsträger der Strukturen innerhalb des Bildes und können in Form einer farblich codierten Segmentationsmaske dargestellt werden. In der medizinischen Bildanalyse im Speziellen findet meist eine Segmentierung in ein Vordergrundobjekt, welches von primärer Bedeutung ist, sowie in einen Hintergrund statt. Eine vollständige Zerlegung des Bildes in Segmente ist hierbei meist nicht gefordert, sondern lediglich das Labeln eines bestimmten Objektes von Interesse, wie z.B. eines zu vermessenden Organs oder eines Tumors (63, 65). Die Aufgaben im Rahmen der medizinischen Bildanalyse beinhalten häufig die Segmentierung von Organen oder Strukturen für eine anschließende Vermessung, dreidimensionale Oberflächendarstellung und ggf. Volumetrie oder auch Bewegungsanalysen. Auch die Planung von Interventionen oder von chirurgischen Verfahren kann Anhand von Bildanalyseverfahren mit Segmentierung bestimmter Organe oder Läsionen erfolgen. Des Weiteren spielen auch die Detektion, Lokalisation

und Bewertung von Läsionen wie z.B. Tumoren eine zunehmend größere Rolle (63, 65). Um dieses zu ermöglichen, werden die Läsionen zunächst segmentiert, um dann bestimmte Merkmale, oder auch features genannt, zu extrahieren, welche für die zu stellende Diagnose typische Ausprägungen haben. Dabei kann mit bereits diagnostizierten und gesicherten Trainingsdaten ein Klassifikationsalgorithmus angeleitet werden (65). Arbeiten mit neuronalen Netzwerken erzielen dabei inzwischen ähnliche oder gar höhere Genauigkeit als menschliche Experten (66).

1.3.2 Herausforderungen

Die medizinische Bildanalyse ist mehreren Herausforderungen unterworfen, welche den Prozess der Segmentierung oder auch der Klassifizierung deutlich erschweren können. Zum einen wäre da die oftmals komplexe Form der zu segmentierenden Objekte innerhalb eines medizinischen Bildes. So kann ein Organ oder auch ein Tumor eine komplexe geometrische Form aufweisen. Darüber hinaus kann ein Tumor auch unscharfe und stark variable Berandungen haben. Außerdem kann auch die interne Beschaffenheit des Organs sehr heterogen sein, sei es auf Grund multipler verteilter Gewebetypen oder wegen Veränderungen durch Pathologien. Alle diese Faktoren können eine Modellierung deutlich erschweren (63-65).

Ein weiterer wichtiger Punkt ist die Heterogenität der verschiedenen Bildgebungstechniken, welche auf verschiedenen physikalischen Prinzipien beruhen, wie z.B. der Abschwächung der Röntgenstrahlung oder der Magnetisierung von Wasser im Gewebe. Diese Heterogenität bringt Mehrdeutigkeiten mit sich, die sich z.B. in unterschiedlicher Skalierung der Pixel, einem unterschiedlichen Kontrastspektrum und verschiedenen Arten der Bildartefakte äußern können. Als störende Artefakte können sich beispielsweise ein Bildrauschen durch Patientenbewegung, atemabhängige Verlagerung der Organe sowie Fremdkörper (63, 65, 67).

1.4 Automatische Bildanalyse mittels Deep Learning

1.4.1 Allgemeines

Deep Learning ermöglicht die zielsichere Erkennung von Objekten im Bild, deren Klassifikation und Segmentierung. Durch Deep Learning und Verwendung neuronaler Netzwerke können Muster in Daten erkannt und erlernt werden. Als Deep Neural Network

bezeichnet man dabei ein neuronales Netzwerk, das aus drei oder mehr Schichten besteht (68, 69).

1.4.2 Convolutional Neural Networks

Eine vorrangig für die Bildanalyse entwickelte neuronale Netzwerkstruktur sind die Convolutional Neural Networks (CNN). CNN sind zur Standardmethode für die Klassifizierung, Bearbeitung und Erzeugung von Bildern avanciert. Das Grundprinzip basiert auf der Einbeziehung und Auswertung der Nachbarschaftsbeziehung räumlicher Merkmale zueinander. Der Name beruht auf der mathematischen Operation namens Faltung oder Konvolution, die in bestimmten Schichten des Netzes abläuft, die daher als Convolutional Neural Networks bezeichnet werden (70). Eine Konvolution besteht im Wesentlichen darin, einen Filter über den Input zu legen. Jeder Convolutional Layer enthält eine Reihe von Filtern, die als Faltungskerne bezeichnet werden. Der Filter ist eine Matrix ganzer Zahlen, die auf eine Teilmenge der Input-Pixelwerte angewendet wird, die die gleiche Größe wie der Kernel hat. Jeder Pixel wird mit dem entsprechenden Wert im Kernel multipliziert, dann wird das Ergebnis zu einem einzigen Wert summiert. Die Schichten werden verarbeitet, indem jedes Bildpixel in numerischer Form in einer Matrix abgebildet wird. Die wichtigsten Parameter sind die Anzahl der Kernel und die Größe der Kernel. Pooling Layers ähneln den Convolution Layers, führen aber eine bestimmte Funktion aus, wie z. B. Max-Pooling, das den Maximalwert in einer bestimmten Filterregion übernimmt, oder Average-Pooling, das den Durchschnittswert in einer Filterregion berücksichtigt. Sie werden in der Regel verwendet, um die Dimensionalität des Netzes zu reduzieren.

CNNs haben den Vorteil, dass sie sehr robust gegenüber visuellen Variationen ähnlichen Inhalts sind, wobei dies veränderte Kontraste, Größenskalierungen, Translationen und Rotationen miteinbezieht. Durch die Anwendung von CNNs gelingt es, eine verbesserte Realisierbarkeit und Automatisierbarkeit von Segmentationen zu bestreiten (71).

1.4.3 U-Net

Das U-Net ist eine Netzwerkarchitektur aus dem Kontext der CNNs, das den aktuellen State-of-the-Art für eine zuverlässige Segmentierung darstellt (72). Die U-Net-Struktur besteht aus zwei symmetrischen Pfaden, einem contracting und einem expanding path, die schematisch ein U-förmiges Aussehen nachahmen. Während der Kontraktion wird

die räumliche Information reduziert, während die Merkmalsinformation erhöht wird. Der expansive Pfad kombiniert die Merkmals- und Rauminformationen durch eine Abfolge von Aufwärtskonvolutionen und Verkettungen mit hochauflösenden Merkmalen aus dem kontrahierenden Pfad über Skip Connections (73). Bei der Entwicklung eines U-Net-basierten Netzwerkes für die Segmentierung der Prostata besteht die Herausforderung, dass die Prostata nur einen relativ kleinen Teil des Bildes einnimmt und es viele Voxel im Bild gibt, die für die Segmentierung nicht von Bedeutung sind. Demnach ergibt sich für den Algorithmus die Notwendigkeit, irrelevante Informationen zu umgehen und die Aufmerksamkeit auf die richtige Struktur zu lenken, um die segmentierten Regionen in entsprechender Lokalisation zu identifizieren. Darüber hinaus sind MRT-Aufnahmen nicht vollkommen reproduzierbar mit uneinheitlichen Intensitätswerten bei identischen Strukturen und sind artefaktanfällig (74, 75). Daraus ergibt sich die Notwendigkeit mehrerer Vorverarbeitungsschritte des Datensatzes, um eine zuverlässige Interpretation durch das Netzwerk zu ermöglichen.

1.5 Zielsetzung der Arbeit

Ziel dieser Arbeit ist die Implementierung eines Deep Learning Algorithmus zur automatisierten Segmentierung der Prostata. Dabei soll zudem sowohl die zonale Einteilung möglich sein wie auch die automatische Erkennung und Lokalisation karzinomsuspekter Areale.

Es soll ein neuronales Netz basierend auf der U-Net Architektur aufgebaut werden und mit einem Datensatz aus 79 von Experten annotierten Prostata MRT Untersuchungen trainiert werden. Anschließend soll mittels eines Testdatensatzes von 19 weiteren MRT Untersuchungen die Genauigkeit der Segmentation im Vergleich zum Expertengoldstandard ermittelt werden.

Es sollen hierbei folgende Fragen beantwortet werden:

1. Ist mittels eines neuronalen Netzwerkes basierend auf der U-Net Architektur eine robuste, automatisierte Segmentation der Prostata mit Hilfe eines verhältnismäßig kleinen Trainingsdatensatzes möglich?
2. Kann das neuronale Netzwerk eine ähnliche Genauigkeit erreichen wie ein menschlicher Expertenstandard?

3. Wie ist die Interrater Übereinstimmung zwischen zwei menschlichen Readern im Vergleich zum Algorithmus?
4. Ist eine verlässliche Detektion und Segmentation karzinomsuspekter Läsionen der Prostata möglich?

2 Materialien und Methoden

2.1 Erhebung der Datensätze

2.1.1 Auswahl der Studienpopulation

Für die vorliegende Arbeit wurden retrospektiv Patienten betrachtet, welche an der Charité Universitätsmedizin Berlin eine biparametrische MRT-Untersuchung erhalten hatten. Diese Patienten wurden von Februar 2016 bis Januar 2020 am Campus Mitte oder dem Campus Virchow Klinikum der Charité jeweils in einem 3T Scanner untersucht. Insgesamt lagen 269 MRT-Untersuchungen vor. Im nächsten Schritt wurden die Befundtexte nach der PI-RADS-Klassifikation durchsucht und alle Untersuchungen mit einem Befund von mindestens PI-RADS 4 herausgefiltert. Daraufhin wurden alle Patienten, welche im Anschluss entweder Prostatektomie oder in einigen Fällen eine Biopsie erhalten hatten, hinsichtlich ihres histopathologischen Befundes betrachtet. Zuletzt wurden alle Patienten mit einem positiven MRT-Befund eines Prostatakarzinoms, bei denen auch histopathologisch eines nachgewiesen werden konnte, in die Studie eingeschlossen. Im letzten Schritt wurden Untersuchungen aussortiert, bei denen die Bildqualität für eine Deep Learning basierte Auswertung nicht ausreichend war. Insgesamt konnten 102 Patienten identifiziert werden, welche ein histologisch gesichertes Prostatakarzinom mit positivem MRT-Befund und eine Untersuchung in hinreichender Bildqualität hatten. Diese wurden aufgeteilt und 83 für den Trainingsdatensatz verwendet, wohingegen 19 für den Testdatensatz verblieben. Als negative Kontrolle wurden 56 weitere Patientinnen mit einem negativen MRT (PI-RADS 1 oder PI-RADS 2) und mindestens einem Follow Up in welchem weiterhin kein Karzinom verdächtig wurde eingeschlossen. Für die Studie wurde ein Ethikantrag mit der Referenznummer EA4/062/20 im Voraus durch die Ethikkommission bewilligt.

2.1.2 Untersuchungstechnik der MRT

Sämtliche MRT-Untersuchungen der Prostata, welche in dieser Arbeit betrachtet wurden, wurden an einem von zwei hochmodernen 3T Kernspintomographen der Firma Siemens Healthineers aus Erlangen (Model Magnetom Skyra und Model Magnetom Vida) unter Verwendung einer 30-Kanal Körperoberflächenspule akquiriert. Für die Verbesserung der Bildqualität und zur Vermeidung von Artefakten wurden die Patienten vor der Untersuchung zum einen gebeten, ihr Rektum vollständig zu entleeren, und zum anderen durch den untersuchenden Arzt mit 40 mg Buscopan intramuskulär prämediziert. Das Protokoll der biparametrischen MRT mit einer axialen und coronaren T2W sowie einer axialen DWI mit errechneter ADC Map entsprach den gängigen Empfehlungen der ESUR. Die Untersuchungsparameter der T2W waren hierbei:

Standard Schichtanzahl 24, Schichtdicke 3 mm, In-Plane-Auflösung 0,52 x 0,47 mm, FOV-Größe 180 x 180 mm, TE / TR 116 ms / 4040 ms, Turbofaktor 25, Flip-Winkel 160°, Akquisitionszeit 3 Minuten und 56 Sekunden.

Für die DWI waren die Untersuchungsparameter wie folgt:

Standard Schichtanzahl 23, Schichtdicke 3 mm, In-Plane-Auflösung 1,4 x 1,4 mm, FOV 220 x 220 mm, TE / TR 61 ms / 4400 ms, b-Werte 5, 100, 500, 1000 s/mm³, Gesamtakquisitionszeit 5 Minuten und 5 Sekunden. Die auf dem MRT-Scanner vorinstallierte Software (Version VE11A) wurde verwendet, um eine ADC-Map aus den b-Werten von 50 bis 1000 s/mm³ sowie einen hohen b-Wert von 1400 s/mm³ zu berechnen. Bei sehr großen Drüsen mit einer besonders ausgeprägten kranio-kaudalen Ausdehnung, die nicht mit der standardisierten Anzahl an Schichten erfasst werden konnten, wurde die Schichtanzahl entsprechend erhöht.

2.1.3 Segmentierung

Die biparametrischen MRT-Untersuchungen wurden zunächst anonymisiert aus dem PACS der Charité exportiert und die DICOM-Bilder auf einem externen Server der Arbeitsgruppe gespeichert. Anschließend erfolgte die Annotation der Bilder durch zwei Fachärzt*Innen für Radiologie mit je 5 und 7 Jahren Erfahrung in uroonkologischer Bildgebung. Als ROI wurden karzinomsuspekte Areale mit einem PI-RADS Score von ≥ 4 festgelegt sowie die CG und PZ in einer separaten zonalen Segmentierung. Die Segmentierungen erfolgten von Hand und Schichtweise in den axialen Sequenzen. Die zonale Aufteilung in CG und PZ wurde in der axialen T2W identifiziert und pixelweise

segmentiert, während die karzinomsuspekten Areale in der axialen T2W sowie der ADC Map betrachtet wurden. Es wurde jeweils eine farbkodierte Maske für die CG, die PZ und die Karzinome verwendet. Bei mehr als einem Tumor innerhalb desselben Organs wurde dieselbe farbkodierte Maske für sämtliche Läsionen verwendet.

Für die Segmentierungen wurde die Open Source Software ITK-Snap (Version 3.8.0; www.itksnap.org) verwendet. Im Anschluss wurden die Annotationen als NIfTI-Dateien auf dem Server der Arbeitsgruppe zusammen mit den DICOM-Bildern der Untersuchungen gespeichert. Abbildung 2 zeigt eine Beispielsegmentierung.

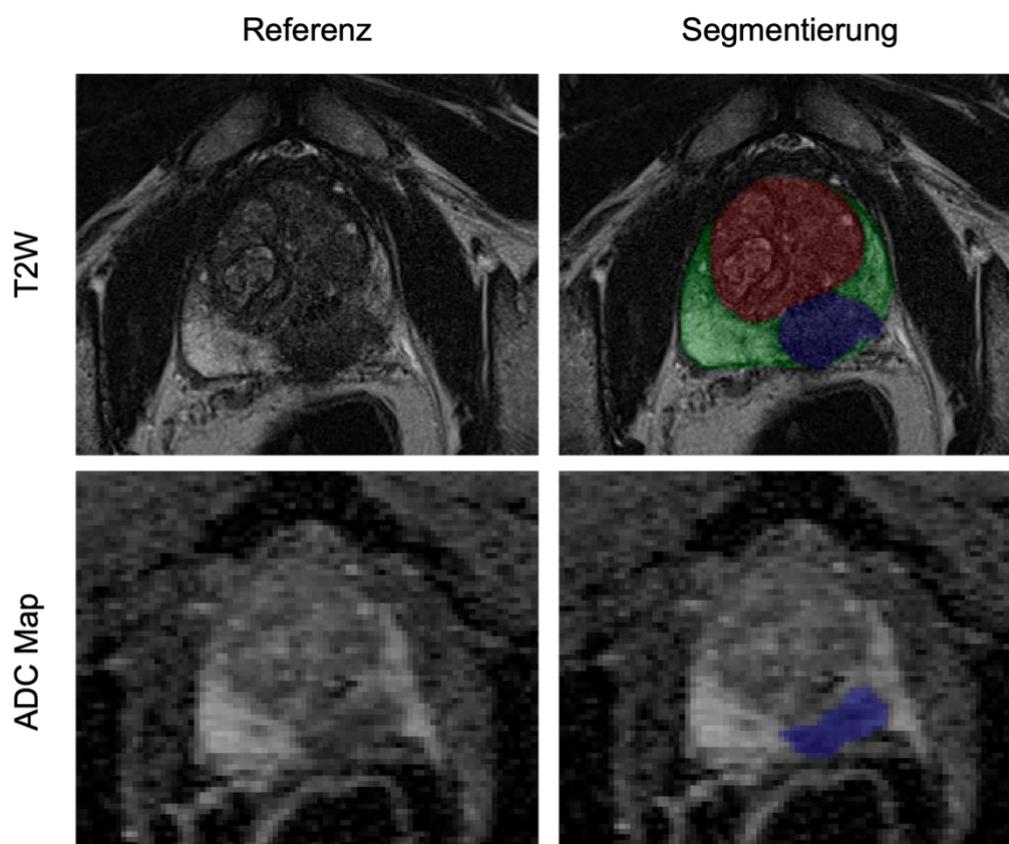


Abbildung 2. Beispielsegmentierung einer Prostata mit karzinomsuspektem Areal in der links posterolateralen PZ der mittleren Drüse. Dargestellt sind jeweils das Referenzbild und die Segmentierung in T2W und der ADC Map. Rote Maske = CG, grüne Maske = PZ, blaue Maske = Tumorareal.

2.2 Automatisierte Vorverarbeitung der Bilddaten

Die MRT Bilddaten lagen als DICOM Serien vor. Diese wurden mittels der Programmbibliothek SimpleITK (Version 2.0.2) eingelesen und in das NIFTI Format umgewandelt, um das Dateneinlesen zu vereinfachen und zu beschleunigen. Für die diffusionsgewichteten Sequenzen wurden jeweils nur die Serie mit dem höchsten B-Wert verwendet.

2.2.1 Normalisierung und Entfernen des Biasfeldes

MRT Aufnahmen können durch verschiedene Bildartefakte beeinflusst werden. Zum einen sind dies Bewegungsartefakte bedingt durch willkürliche Patientenbewegungen, z.B. Bewegung der Beine, oder unwillkürliche Bewegungen (Motilität des Rektums). Weitere mögliche Artefakte sind aufnahmebedingt, wie z.B. Ghosting bzw. Pulsationsartefakte. Hier wird durch die Bewegung von Protonen in Blutgefäßen die Phasenkodierung des MRT gestört, was zu schemenhaften Dopplungen der Gefäße im gesamten Bild führt. Ebenfalls durch fehlerhafte Phasenkodierung bedingt sind Phase wrapping oder Aliasing Artefakte, bei denen, bedingt durch ein zu klein gewähltes Field of View, Strukturen am Rand des Feldes auf der gegenüberliegenden Seite aufgetragen werden. Eine Korrektur dieser Artefakte ist jedoch in der Bildverarbeitung nicht möglich. Der Einfluss der genannten Artefakte auf das Training kann stattdessen durch gezielte Bildaugmentationen reduziert werden.

Inhomogenität des Magnetfeldes im Scanner führt zu dem Bildartefakt des Biasfeldes. Dies bezeichnet Intensitätsunterschiede innerhalb des MRT Bildes, welche mathematisch entfernt werden können. In der vorliegenden Arbeit wurde dies durch die Anwendung des N4ITK Algorithmus vollzogen (76).

Eine weitere Herausforderung bei der Analyse von MRT Bildern durch neuronale Netzwerke ist eine fehlende physikalische Korrelation der Voxelintensitätswerte. Während in anderen Untersuchungsmodalitäten wie der Computertomografie (CT) die Voxelwerte unmittelbar physikalischen Gewebeeigenschaften entsprechen und sich zwischen verschiedenen Geräten gleichen, ist dies bei der MRT nicht der Fall. Dieses erschwert das Training neuronaler Netzwerke. Deshalb erfolgte in der vorliegenden Arbeit eine Normalisierung der Voxelwerte auf einen Mittelwert von 0 und eine Standardabweichung von 1. Dieses wird durch folgende Formel erreicht:

$$\bar{x} = (x - x_{max})/\sigma$$

Hierbei entspricht x einem individuellen Wert, x_{max} dem Maximalwert der Sequenz und σ der Standardabweichung der Werte. Durch diese Transformation der Intensitätswerte kann eine annähernd gleiche Intensität aller Werte auf verschiedenen Bildern erreicht werden.

2.2.2 Resampling

MRT Aufnahmen bilden einen bestimmten Bereich des menschlichen Körpers ab. Im Falle der vorliegenden Arbeit zeigte jede Sequenz die Prostata. Jedoch unterscheiden sich Sequenzen voneinander in ihrer Auflösung, dem Field of View und auch der räumlichen Ausrichtung. Die Auflösung der T2w Sequenzen in dieser Arbeit betragen 768 x 768 Pixel oder 512 x 512 Pixel, die Auflösung der DWI/ADC hingegen lediglich 128 x 128 Pixel, obwohl alle Sequenzen ein ähnliches Field of View hatten. Diese Tatsache führt dazu, dass Voxel verschiedener Sequenzen, welche die selbe Körperregion abbilden, unterschiedliche Bildkoordinaten haben. Die maximale Koordinate eines Voxels der ADC kann somit nur (128, 128) betragen und zeigt den unteren rechten Rand der Aufnahme an. In der T2w würde die Koordinate (128, 128) jedoch einem Voxel im oberen rechten Bildquadranten entsprechen. Aus diesem Grund, enthalten DICOM Aufnahmen im Bildheader Informationen über die Größe eines Voxels in Millimeter sowie die Lokalisation des Bildes in einem globalen Koordinatensystem. Wenn diese Informationen zur Verfügung stehen, kann errechnet werden welche Voxel in einem Bild zueinander korrespondieren, obwohl sich ihr lokales Koordinatensystem unterscheidet. Beim Training neuronaler Netzwerke können diese Headerinformationen jedoch nicht genutzt werden. Aus diesem Grund erfolgte in dieser Arbeit ein Resampling der Sequenzen, sodass die Voxelgröße und die Ausrichtung der Sequenzen identisch waren. Nach dem Resampling hatten alle Sequenzen isotrope Voxel mit einer Größe von 0,5 x 0,5 x 0,5 mm und eine identische Ausrichtung (RAS = Right, Anterior, Superior).

2.2.3 Cropping

Der Rechenaufwand neuronaler Netzwerke ist direkt proportional zur Größe des Inputs. Durch technische Limitationen der MRT Bildgebung kann die Größe des Field of Views jedoch nicht beliebig klein gewählt werden, da sonst das Signal zu Rausch Verhältnis

ungünstig beeinflusst wird. Hierdurch bildet jede Aufnahme der Prostata auch Teile des umliegenden Beckens ab. Für die Segmentation der Prostata und Detektion des Prostatakarzinoms bringt dieses jedoch keinen Mehrwert. Die Aufnahmen wurden deshalb systematisch prozentual beschnitten, um die Größe zu reduzieren, ohne dabei relevante Bildinformationen zu verlieren. Eine Zusammenfassung der Vorverarbeitungsschritte bietet Abbildung 3.

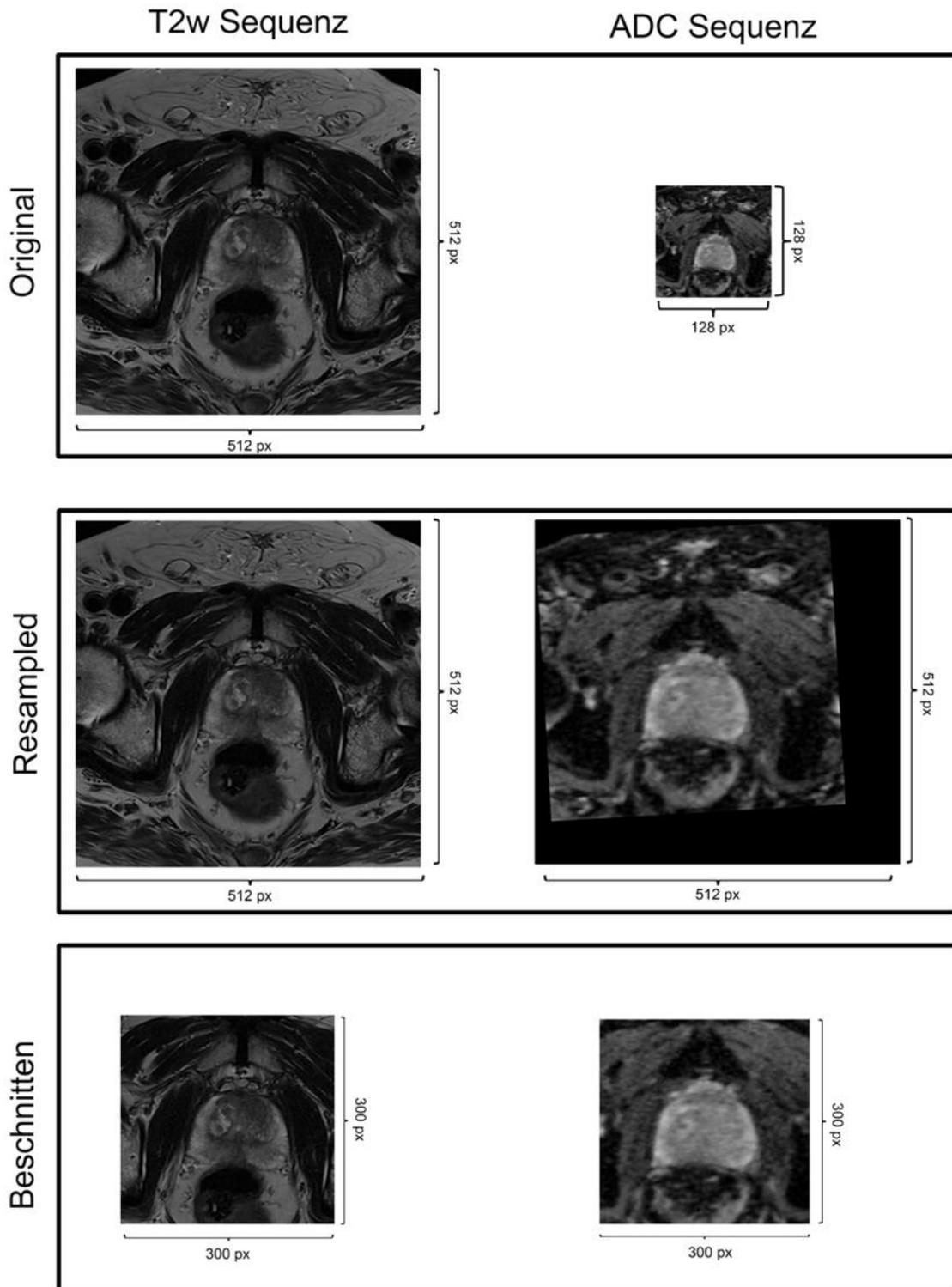


Abbildung 3. Zusammenfassung der Vorverarbeitung. Im Original unterscheiden sich die Aufnahmen in der Auflösung sowie im Field of View (FOV). Nach Durchführung des Resampling sind sowohl FOV als auch Auflösung der Aufnahmen angeglichen. Die Randbereiche der Bilder enthalten keine relevanten Informationen für die Segmentierung der Prostata. Durch die unterschiedlichen FOV's der Aufnahmen entstehen zudem schwarze Areale im Randbereich. Durch eine Beschneidung der Bilder werden die Randbereiche entfernt, ohne relevante Informationen für die Segmentierung zu verlieren.

2.3 Training des Deep Learning Modells

2.3.1 Aufbau des Netzwerkes

Ein U-ResNet mit sechs Blöcken wurde für die Segmentation der Anatomie und des Tumors genutzt. Jeder Block des Netzwerkes bestand hierbei aus einem Residual Block mit zwei (bzw. vier für die Tumorsegmentation) Convolutional Blöcken, deren Aufbau in Abbildung 3 schematisch dargestellt ist. Nach jedem Block erhöhte sich die Anzahl der Bildkanäle von initial 1 (Anatomiesegmentierung) bzw. 3 (Karzinomsegmentierung) auf 16, 32, 64, 128, 256 und 512.

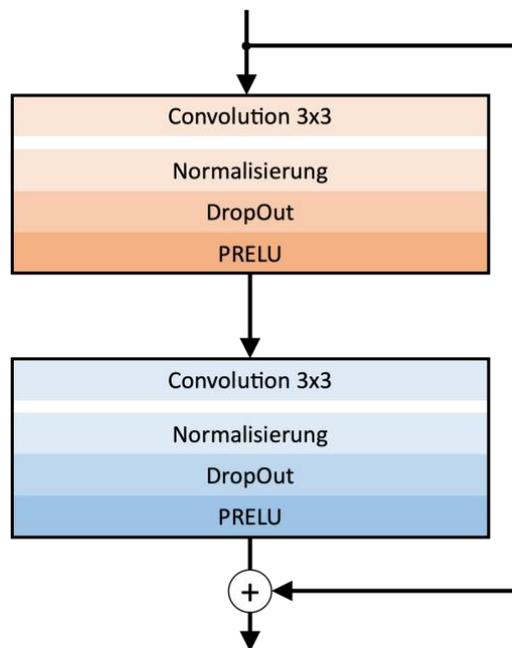


Abbildung 4 zeigt den Aufbau eines Residual Blocks mit zwei Convolutional-Schichten. Der Input durchläuft die erste Convolution, es folgt eine Normalisierung, DropOut und schließlich die Aktivierungsfunktion. DropOut betrug in dieser Arbeit bei allen Modellen 15%. Als Aktivierungsfunktion wurde die Parametric Rectified Linear Unit (PReLU) gewählt (77). Nachdem beide Convolutional-Schichten durchlaufen wurden, erfolgt die Zusammenführung des Outputs mit dem ursprünglichen Input (Residual bzw. Skip-connection).

Da die Größe der Aufnahmen unterschiedlich war, wurde eine systematische Bildbeschneidung während des Trainings vorgenommen. Hierbei wurden automatisiert Subregionen der Aufnahme mit einer Größe von 96 x 96 x 96 Voxel erstellt, welche nur einen Teil der Aufnahme enthielten. Hierdurch konnte die Größe des Inputs für das Netzwerk konstant gehalten werden, ohne die Originalaufnahmen skalieren zu müssen.

2.3.2 Hyperparameter

Das Training wurde für 500 Epochen durchgeführt unter Nutzung von Early Stopping. Das bedeutet, dass der durchschnittliche Dice Koeffizient nach jeder Epoche als Schlüsselmetrik errechnet wurde. Zeigte sich keine Verbesserung der Schlüsselmetrik innerhalb von 100 Epochen, wurde das Training verfrüht beendet. Der Novograd Optimizier wurde genutzt mit einem Learning-Rate Planer. Dieser Planer folgte der One-Cycle Policy, bei welcher die Learnrate zuerst langsam bis zu einem definierten Maximum von 0.001 ansteigt, um dann langsam bis zum Ende des Trainings wieder zu fallen. Hierdurch werden zu große Anpassungen der Gewichte am Anfang bzw. Ende des Trainings vermieden. Ein Strafterm (Weight Decay) von 0.001 wurde gewählt für die Regularisierung des Netzwerkes. Eine Kombination aus Cross-Entropy und Soft-Dice diente als Loss-Funktion.

2.4 Statistische Auswertung

Die statistische Analyse wurde mit Microsoft Excel, „Python“ (Version 3.8) und "R" (Version 4.1.2) unter Verwendung des "tidyverse"-Paketes sowie des MONAI Softwarepaketes (Medical Open Network for Artificial Intelligence, Version 0.8) durchgeführt.

Die Performance der Segmentierung wurde durch den Sørensen-Dice Koeffizient, die Hausdorff Distanz und die durchschnittliche Oberflächendistanz evaluiert. Der Dice Koeffizient misst die Übereinstimmung von zwei Volumina (A und B) und wird errechnet als:

$$\frac{2 |A \cap B|}{|A| + |B|}$$

Die Schwäche des Dice Koeffizienten liegt in der Evaluation der Oberfläche eines Volumens. Deshalb wurde zusätzlich die Hausdorff Distanz bestimmt. Diese gibt an, wie groß die maximale Distanz zwischen einem Punkt auf Oberfläche A und einem Punkt auf Oberfläche B sein kann und wird in Voxeln angegeben.

Im Gegensatz zur Hausdorff Distanz gibt die durchschnittliche Oberflächendistanz den Mittelwert des Abstandes zwischen zwei Oberflächen an.

Neben der Bewertung der Modellperformance wurden Dice Koeffizient, Hausdorff Distanz und durchschnittliche Oberflächendistanz auch für die Bestimmung der Interraterübereinstimmung verwendet. In Ermangelung eines echten Referenzstandards wurde davon ausgegangen, dass sich die Leistung des Modells nicht signifikant vom Menschen unterscheidet, wenn die Übereinstimmungen zwischen Modell und Beurteiler 1 und zwischen Modell und Beurteiler 2 nicht signifikant von der Übereinstimmung zwischen den Beurteilern abweichen. Für alle Werte wurden der Median und der Interquartilsbereich (IQR) bestimmt, und die Werte wurden zudem als Boxplots aufgetragen. Der Wilcoxon Test wurde verwendet, um die Performance zwischen Menschen und Modell sowie das Alter in den Trainings- und Testsätzen zu vergleichen. Aufgrund multipler Testungen erfolgte eine Bonferroni Korrektur für Tests zur Segmentationsperformance, was zu einem Signifikanzniveau von p von $< 0,0055$ führte. Für alle anderen Tests wurde ein p -Wert von $< 0,05$ als signifikant angesehen.

3 Ergebnisse

3.1 Beschreibung des Datensatzes

Der finale Datensatz umfasste 158 Patienten mit einem Durchschnittsalter von 69 Jahren (Standardabweichung (SD) 9 Jahre, Spanne 35-84 Jahre). 139 Patienten mit einem Durchschnittsalter von 66 Jahren (SD 9 Jahre, Spanne 35-84 Jahre) wurden dem Trainings-/Validierungsdatsatz zugeordnet. Das Alter der Patienten im Testdatensatz war mit 71 Jahren (SD 10 Jahre, Spanne 53-83 Jahre) um 5 Jahre höher als im Trainingsdatensatz. Dieser Unterschied war jedoch statistisch nicht signifikant ($p=0,06$). 83/139 (60 %) der Patienten in der Trainingsgruppe hatten verifizierte PCas mit einem in der ADC Map gemessenen Tumolvolumen von $3,0 \pm 4,7$ ml (Spanne 0,1-24,2 ml). Das in T2w-Bildern gemessene mittlere Prostatavolumen betrug $56,0 \pm 30,7$ ml (Spanne 14,8-197,4 ml), wobei die zentrale Drüse (CG) $41,2 \pm 29,6$ ml (Spanne 8,1-177,4 ml) und die periphere Zone (PZ) $14,8 \pm 6,3$ ml (Spanne 3,4-41,3 ml) umfasste. 20 MRTs aus dem Trainingsset wurden nach dem Zufallsprinzip als Validierungsset zugewiesen. Alle Patienten in der Testgruppe hatten ein nachgewiesenes PCa mit einer Tumogröße von $3,5 \pm 5,1$ ml (Spanne 0,4-21,2 ml). Das mittlere Volumen der gesamten Drüse betrug $54,5 \pm 17,5$ ml (Spanne 28,1-93,8 ml), das der zentralen Drüse $35,2 \pm 15,1$ ml (Spanne 13,6-77,7 ml) und das der peripheren Zone $19,3 \pm 7,6$ ml (Spanne 11,0-42,4 ml).

3.2 Übereinstimmung zwischen den menschlichen Experten

Die Interrater-Übereinstimmung war am höchsten für die zentrale Drüse mit einem medianen Dice Koeffizienten von 0,87, einer Hausdorff Distanz von 11,1 und einer durchschnittlichen Oberflächendistanz von 1,0. Für die periphere Zone war die Übereinstimmung geringer mit einem medianen Dice-Koeffizienten von 0,75 und einer Hausdorff Distanz von 15,8. Die durchschnittliche Oberflächendistanz von 0,74 war jedoch geringer als in der zentralen Drüse, was auf eine bessere Übereinstimmung hindeutet. Darüber hinaus war eine größere Variation zwischen den Metriken zu erkennen, was durch den größeren IQR für Dice Koeffizient und Hausdorff Distanz angezeigt wird.

Die Übereinstimmung für die Tumore war geringer, mit einem Dice-Koeffizienten von 0,6, einer Hausdorff-Distanz von 18,8 und einer durchschnittlichen Oberflächendistanz von 5,3. Auch hier war die Abweichung zwischen den Metriken höher mit einem größeren IQR für alle Metriken.

Alle Metriken der Interrater-Übereinstimmung sind in Tabelle 1 enthalten und in Abbildung 5 als Boxplots visualisiert.

	<i>Dice Koeffizient</i>	<i>Hausdorff Distanz</i>	<i>Durchschnittliche Oberflächendistanz</i>
<i>Zentrale Drüse</i>	0,87 IQR 0,05	11,1 IQR 5,1	1,0 IQR 0,4
<i>Periphere Zone</i>	0,75 IQR 0,08	15,8 IQR 6,4	0,74 IQR 0,25
<i>Tumor</i>	0,6 IQR 0,27	18,8 IQR 11,6	5,3 IQR 3,8

Tabelle 1 zeigt die Übereinstimmung der beiden menschlichen Experten bei der Segmentierung der anatomischen Prostatazonen und tumorverdächtiger Areale.

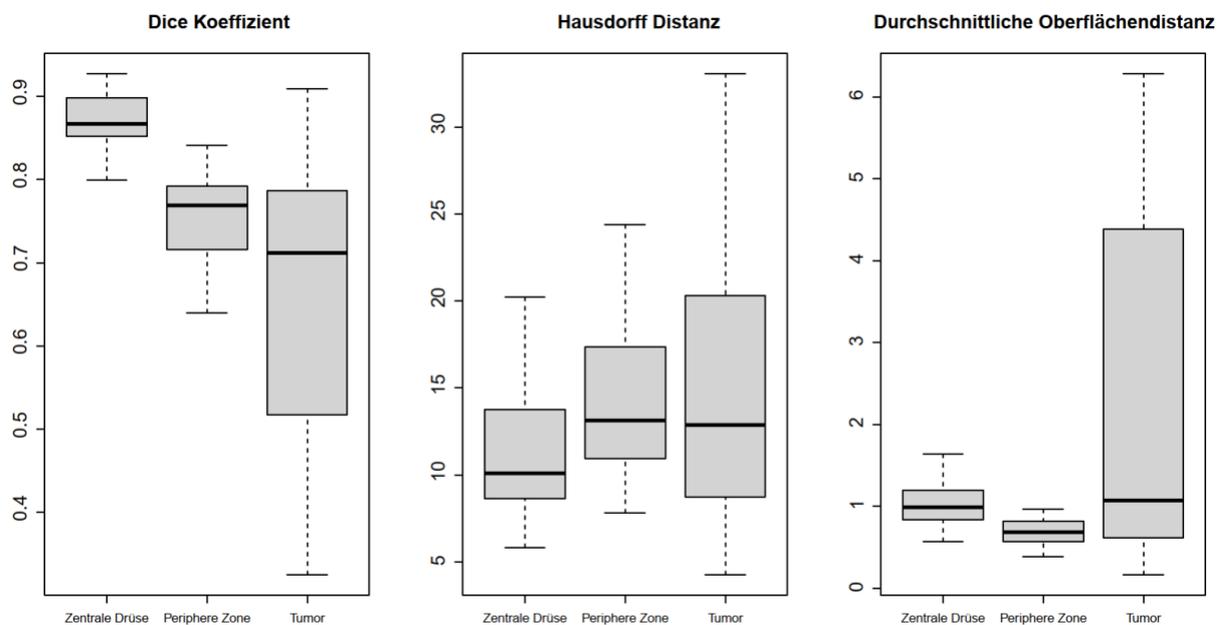


Abbildung 5. Boxplots für die Übereinstimmung der Segmentierungen zwischen den menschlichen Experten.

3.3 Segmentierung der anatomischen Zonen

3.3.1 Verlauf des Trainings

Das Training für die Segmentierung der anatomischen Zonen umfasste 500 Epochen, wobei der höchste durchschnittliche Dice Koeffizient ($(\text{Dice}_{\text{periphäre Zone}} + \text{Dice}_{\text{central gland}}) \cdot 0,5$) nach Epoche 313 erreicht wurde. Der größte Zuwachs an Genauigkeit entstand in den ersten 100 Epochen, in welchen sich der durchschnittliche Dice Koeffizient von 0,39 (Epoche 1) auf 0,72 (Epoche 100) verbesserte. Es zeigten sich jedoch nach 100 Epochen noch starke Schwankungen in den Ergebnissen zwischen den Epochen, z.B. 0,75 in Epoche 106, 0,61 in Epoche 107 und 0,74 in Epoche 108. Nach 200 Epochen waren die Verbesserungen des Dice Koeffizienten nur noch gering und auch die Schwankungen zwischen den Epochen geringer. Ab Epoche 300 kam es schließlich zu keinen nennenswerten Verbesserungen mehr. Abbildung 6 zeigt den zeitlichen Verlauf der Metriken über das Training an, Abbildung 7 zeigt exemplarische Segmentationen zu verschiedenen Trainingszeitpunkten.

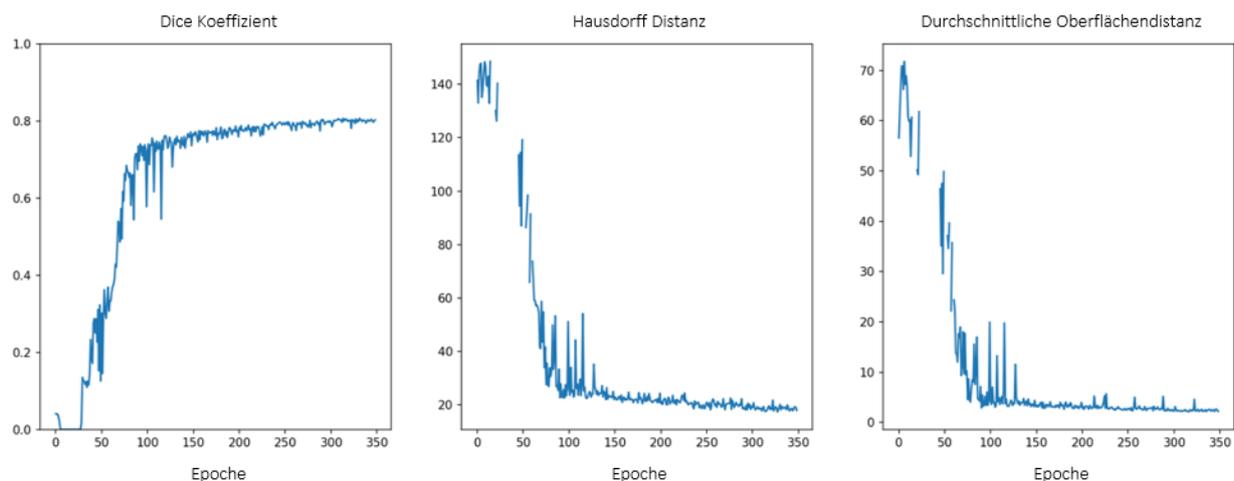


Abbildung 6. Verlauf der Metriken über die verschiedenen Epochen des Trainings. (A) zeigt den Durchschnittlichen Dice Koeffizient and welcher auch als Schlüsselmetrik für die Optimierung des Modells genutzt wurde. (B) gibt den Verlauf der Hausdorff Distanz an und (C) den Verlauf der durchschnittlichen Oberflächendistanz.

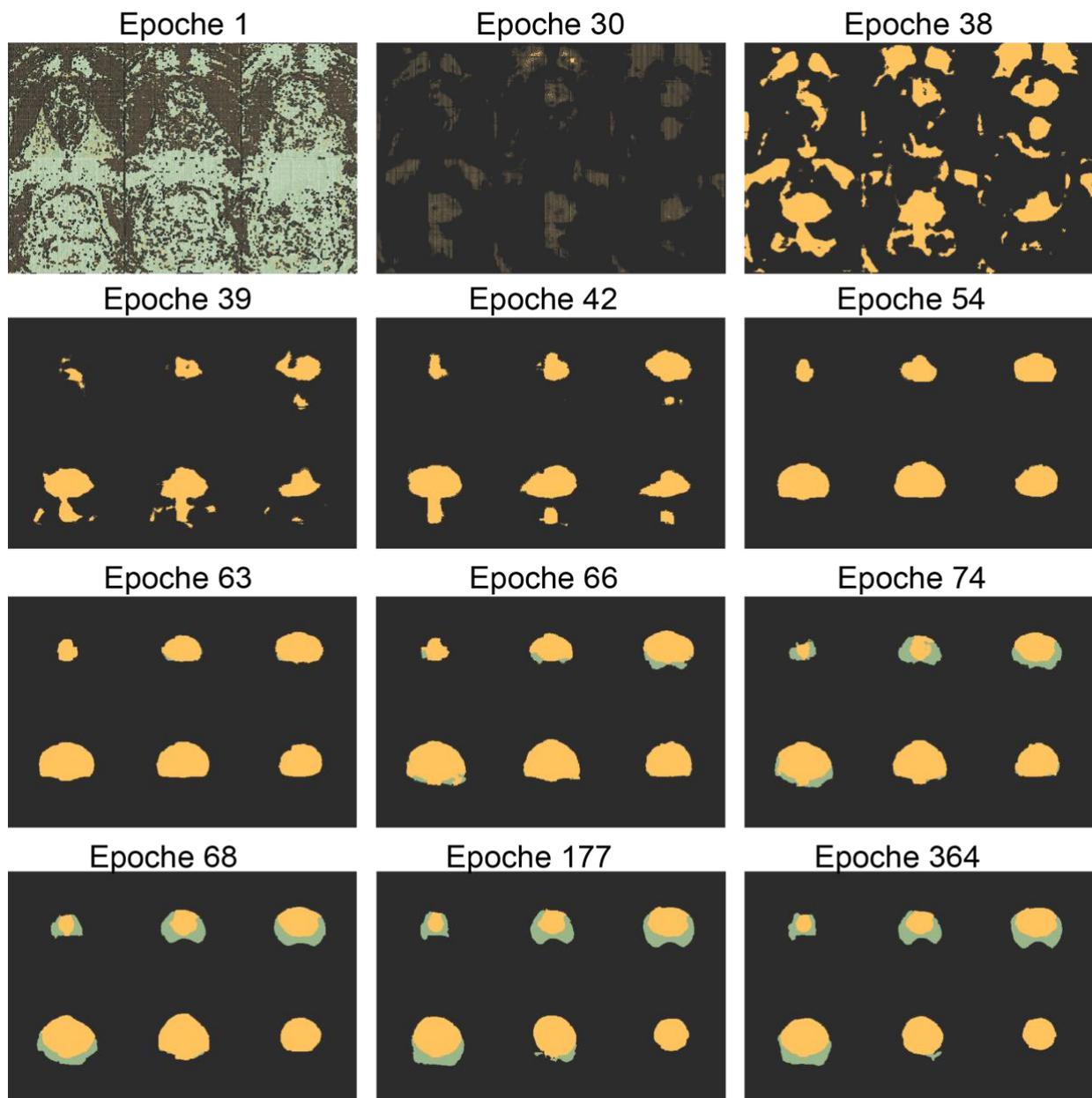


Abbildung 7. Exemplarische Segmentationen der anatomischen Zonen nach verschiedenen Trainingsepochen.

3.3.2 Vergleich zu Beurteiler 1

Die Segmentierung der zentralen Drüse zeigte eine höhere Übereinstimmung mit der Segmentierung der/des ersten Expert:in als die Segmentierung der peripheren Zone. Das Modell erzielte einen Medianen Dice Koeffizienten von 0,88 mit einem Interquartilenabstand (IQR) von 0,06 für die zentrale Drüse und einen Dice Koeffizienten von 0,75 mit einem IQR von 0,05 für die periphere Zone. Auch die Hausdorff Distanz

zeigte geringere Werte für die zentrale Drüse mit 18,2 mit einem IQR von 4,6 als die periphere Zone mit 22,8 und einem IQR von 11,3, welches auf eine höhere Übereinstimmung hindeutet. Dem entgegen gesetzt ergab die Berechnung der durchschnittlichen Oberflächendistanz geringere Werte für die periphere Zone (1,9 mit einem IQR von 0,7) als für die zentrale Drüse (2,2 mit einem IQR von 0,7). Die Übereinstimmung zwischen Modell und Beurteiler 1 ist in Abbildung 8 als Boxplots dargestellt.

3.3.3 Vergleich zu Beurteiler 2

Die Übereinstimmung zwischen den Segmentierungen des Modells und der/dem zweiten Expert:in waren geringer als zum/zur ersten Expert:in. Eine Signifikanztestung durch den Wilcoxon Test ergab hierbei den geringsten p-Wert für den Unterschied in der durchschnittlichen Oberflächendistanz der zentralen Zone mit 0,05. Bei einem korrigierten Signifikanzniveau von 0,0055 waren diese Unterschiede somit nicht signifikant.

Der Dice Koeffizient für die zentrale Drüse betrug 0,88 mit einem IQR von 0,05 und für die periphere Zone 0,73 mit einem IQR von 0,08. Die mediane Hausdorff Distanz für die Segmentation der zentralen Drüse ergab 17,5 mit einem IQR von 6,5 und 33,2 mit einem IQR von 19,5 für die periphere Zone. Die durchschnittliche Oberflächendistanz für die zentrale Drüse war ebenfalls höher als für die periphere Zone mit 2,6 und einem IQR von 0,8 bzw. 1,9 mit einem IQR von 0,3. Die Übereinstimmung zwischen Modell und Beurteiler 2 ist in Abbildung 9 als Boxplots dargestellt.

3.4 Segmentierung des Karzinoms

3.4.1 Verlauf des Trainings

Das Training des Segmentierungsmodells für die Tumore umfasste 500 Epochen mit Early Stopping, wobei das beste Ergebnis nach 287 Epochen erreicht wurde. Ähnlich wie beim Training des Modells für die Segmentierung der anatomischen Zonen erfolgte der stärkste Zuwachs an Genauigkeit innerhalb der ersten 100 Epochen von 0,009 nach Epoche 1 auf 0,29 nach Epoche 100. Im Gegensatz zum Training des Anatomiemodells

blieben die Schwankungen in der Metrik jedoch während des gesamten Trainings bestehen, trotz kontinuierlicher Absenkung der Lernrate.

Abbildung 8 zeigt den Verlauf der Schlüsselmetrik über das Training, Abbildung 9 zeigt Beispielsegmentationen nach verschiedenen Epochen.

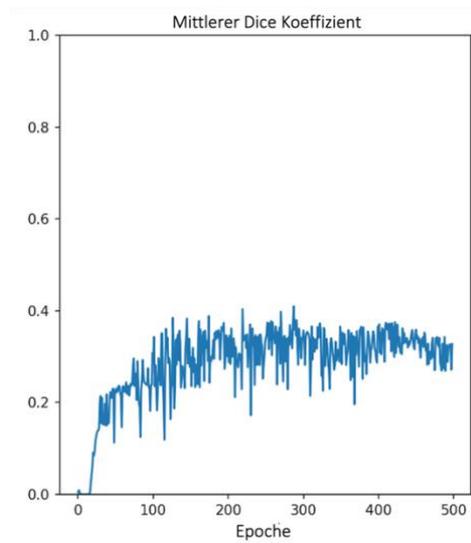


Abbildung 8 zeigt den Verlauf der Schlüsselmetrik über den Verlauf des Trainings. Da nicht alle Patienten im Validierungssatz ein Karzinom hatten, war die Berechnung der Hausdorff Distanz oder durchschnittlichen Oberflächendistanz während des Trainings nicht möglich.

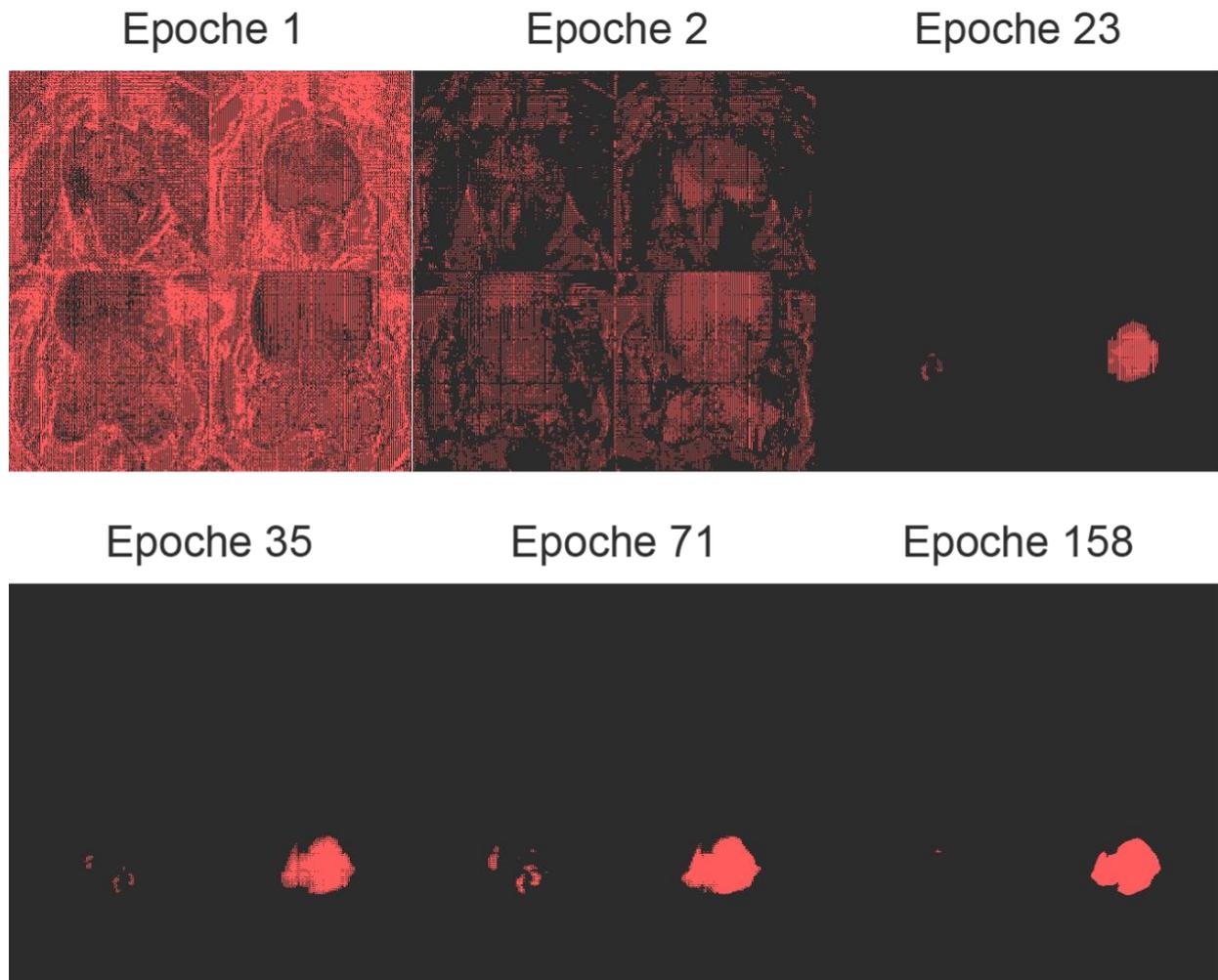


Abbildung 9. Beispielsegmentationen von Prostatakarzinomen durch das Modell nach verschiedenen Epochen des Trainings.

3.4.2 Vergleich zu den Beurteilern

Die Übereinstimmung der Segmentierung des Tumors zu den menschlichen Experten war schlechter als bei den anatomischen Zonen und die Schwankung der Werte war stärker ausgeprägt. Der mediane Dice Koeffizient zwischen Beurteiler 1 und dem Modell betrug 0,45 mit einem IQR von 0,47 und zwischen Modell und Beurteiler 2 lag er bei 0,4 mit einem IQR von 0,59. Die Hausdroff Distanz zwischen Beurteiler 1 und Modell war mit 17,4 und einem IQR von 15,4 geringer als zwischen Modell und Beurteiler 2 mit 39,5 und einem IQR von 31,7. Auch die durchschnittliche Oberflächendistanz war zwischen Modell und Beurteiler 1 war mit 17,4 und einem IQR von 15,4 geringer als zwischen Modell und Beurteiler 2 mit 19,1 und einem IQR von 20,1.

Insgesamt war die Übereinstimmung zwischen dem Modell und den Experten signifikant geringer als die Übereinstimmung der Experten untereinander, sodass nicht von einer menschenäquivalenten Performance des Modells ausgegangen werden kann ($p < 0.001$). Abbildungen 10 und 11 vergleichen die Performance des Modells zwischen anatomischen Zonen und Tumor mittels Boxplots. Tabelle 2 zeigt die Übereinstimmungen des Modells mit den menschlichen Beurteilern für die Segmentationen der Anatomie und der Tumorregionen anhand des medianen Dice Koeffizienten, der medianen Hausdorff Distanz und der medianen Oberflächendistanz.

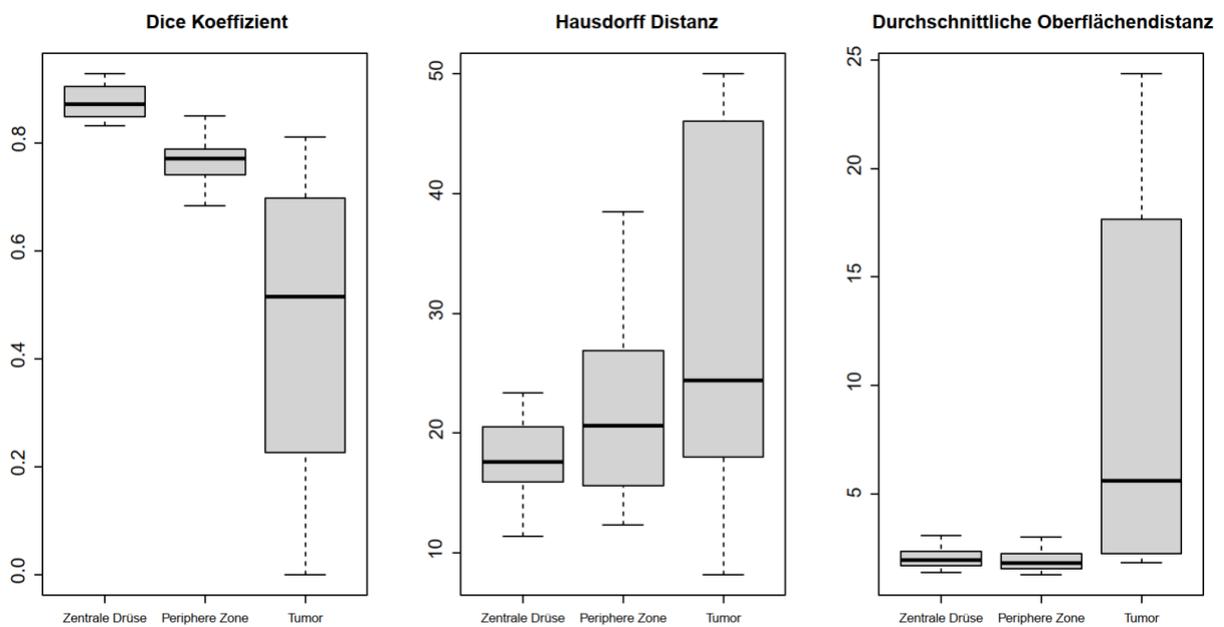


Abbildung 10. Boxplots für die Übereinstimmung zwischen der Segmentierung des Modells und Beurteiler 1.

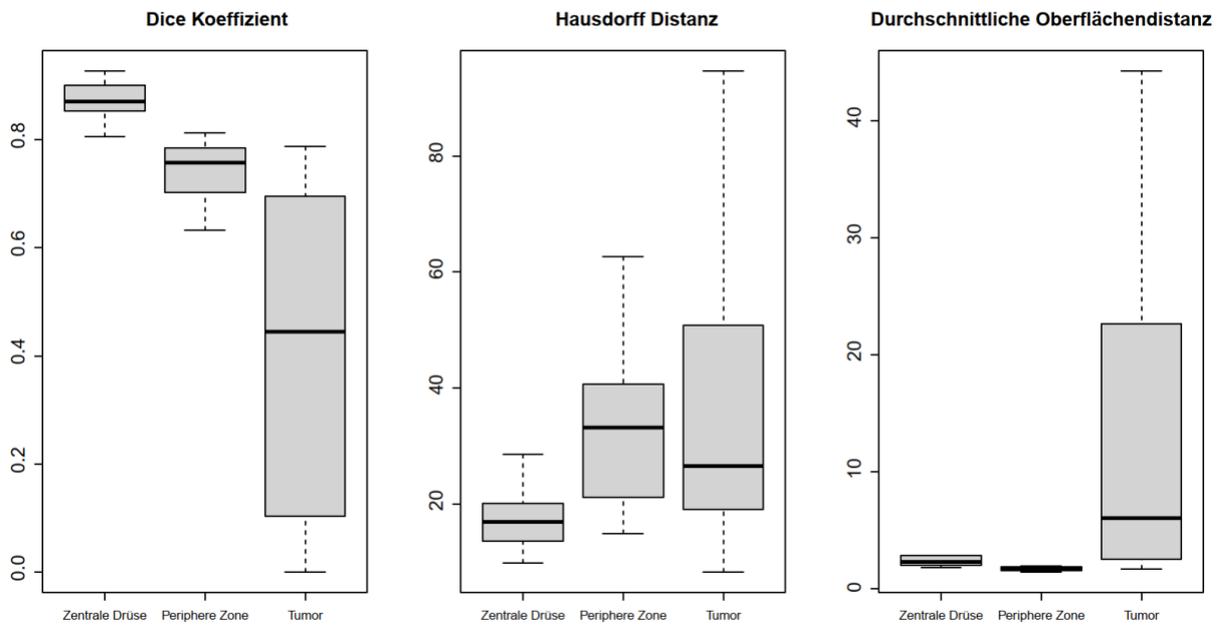


Abbildung 11. Boxplots für die Übereinstimmung zwischen der Segmentierung des Modells und Beurteiler 2.

	Metrik	Zentrale Drüse		Periphere Zone		Tumor	
		Median	IQR	Median	IQR	Median	IQR
Beurteiler 1	Dice Koeffizient	0.88	0.06	0.75	0.05	0.45	0.47
	Hausdorff Distanz	18.3	4.6	22.8	11.3	36.7	28
	Oberflächendistanz	2.2	0.7	1.9	0.7	17.4	15.4
Beurteiler 2	Dice Koeffizient	0.88	0.05	0.73	0.08	0.4	0.59
	Hausdorff Distanz	17.5	6.5	33.2	19.5	39.5	31.7
	Oberflächendistanz	2.6	0.8	1.9	0.3	19.1	20.1

Tabelle 2. Übereinstimmung zwischen der Segmentation des Modells für die zonale Anatomie und den Tumor mit der Segmentation der zwei Beurteiler angegeben als Dice Koeffizient, Hausdorff Distanz und mediane Oberflächendistanz.

4 Diskussion

4.1 Zusammenfassung der Ergebnisse

Im Rahmen dieser Arbeit konnte gezeigt werden, dass mit Hilfe eines Deep Learning Algorithmus basierend auf einer 3D-U-Net Architektur eine robuste, automatisierte Segmentierung der Prostata in biparametrischen MRT-Untersuchungen möglich ist. Hierbei konnte auch die zonale Gliederung der Prostata adäquat differenziert und getrennt segmentiert werden. Die Genauigkeit der Segmentation kam dabei einem menschlichen Expertenstandard nahe, ohne diesen jedoch erreichen zu können.

Auch die Detektion und Lokalisation histologisch gesicherter Prostatakarzinome innerhalb des Testdatensatzes, gelang mit dem Deep Learning Modell in zufriedenstellender diagnostischer Genauigkeit. Auf Grund der Architektur des Modells konnte dieses Level an Genauigkeit mit einem verhältnismäßig kleinen Trainingsdatensatz schnell erreicht werden. Zudem fand die Segmentation und Tumordetektion an einem dreidimensionalen Bilddatensatz im DICOM-Format statt, der direkt aus dem diagnostischen Bildarchiv exportiert und lediglich einmal in das NIfTI Format konvertiert und in das Netzwerk eingespeist werden konnte, was die Komplexität der menschlichen Vorverarbeitung verringert und somit die Nutzbarkeit für mögliche klinische Applikationen erhöht.

4.2 Diskussion der Methoden und Einordnung in die Literatur

Die Segmentierung der Prostata in MRT-Bildern ist eine bereits lange etablierte Technik der medizinischen Bildanalyse, welche eine wichtige Rolle für die Therapieplanung bei Prostatakrebs spielt, z.B. bei der Roboter-assistierten Chirurgie sowie bei der Bestrahlungsplanung. Zudem ist sie auch Grundlage für zukünftige Applikationen der automatisierten Detektion von Tumoren (78). Zunächst wurde die Segmentation im klinischen Alltag überwiegend händisch durch den untersuchenden Radiologen selbst durchgeführt, was allerdings zum einen sehr zeitaufwändig ist und zum anderen auch anfällig für Subjektivität und damit verbundene Ungenauigkeiten bei der Definition der Konturen mit hoher Inter- und Intraobserver Variabilität (78). Dieses Problem konnte auch im Rahmen der vorliegenden Arbeit beobachtet werden, da das Interobserver Agreement

bei der händischen Segmentation der Prostata und insbesondere der zonalen Anatomie und der Tumoreale Einschränkungen aufwies.

Im Zuge des aufkommenden Feldes der Computer-assistierten Diagnostik in der Medizin wurden diverse Ansätze der automatisierten Segmentation der Prostata entwickelt und erprobt. Dabei handelt es sich um kein einfaches Unterfangen wegen der Heterogenität des Prostatagewebes selbst sowie der komplexen Anatomie der Prostata und der umgebenden anatomischen Strukturen (78). Zudem spielen auch variable Umstände wie die Distension und Luftfüllung des Rektums und die Flüssigkeitsfüllung der Harnblase eine Rolle. Auch die nicht normierten Signalintensitätswerte in der MRT und mögliche überlagernde Artefakte komplizieren die Aufgabe zusätzlich (79).

Klein et al. (79) nutzten eine automatische Segmentierungsmethode, die auf einem Atlasvergleich basiert. Diese Methode wird atlas based segmentation genannt. Hierbei wird ein vorhandenes Wissen aus einem Referenzbild extrahiert, welches als Atlas bezeichnet wird. Der Atlas besteht in der Studie von Klein et al. aus einem Satz von MRT Aufnahmen der Prostata und entsprechenden, gelabelten Binärbildern. Es wurde eine non-rigid registration verwendet, um das Bild des Patienten mit den Atlasbildern zu fusionieren. Die am besten übereinstimmenden Atlasbilder werden ausgewählt und die finale Segmentierung entspricht dem Durchschnitt der resultierenden, deformierten label images und des Schwellenwerts nach einer majority voting rule. In der Studie wurde anhand von 50 klinischen Scans in Form von 3D Volumen untersucht, welche allesamt von 3 Experten händisch segmentiert wurden. Es konnte dabei ein DSC von 0,85 erreicht werden, was in der Studie nahe am Interobserver Agreement mit einem DSC von 0,87 lag (79). Die Vorteile in dieser Studie liegen in der Verwendung eines dreidimensionalen Bilddatensatzes, wie es auch für diese Arbeit der Fall war. Dies ermöglicht die direkte Anwendung der Segmentation auf eine exportierte Bildserie aus dem PACS ohne ein übermäßig aufwändiges post-processing und zudem sind direkt Informationen aus allen Raumebenen enthalten, was die Granularität der Segmentation verfeinert und darüber hinaus in weiteren automatisierten Diagnostikschritten wie der Detektion von Karzinomen hilfreich ist, da jeweils die angrenzenden Imageslices mit in Betracht gezogen werden können. Der Nachteil eines dreidimensionalen Ansatzes besteht in der deutlich höheren, benötigten Rechenkapazität.

Ein Nachteil der Studie von Klein et al. gegenüber der hier vorliegenden Arbeit besteht zum einen in der geringeren Fallzahl von 50 Prostata MRT-Untersuchungen sowie zum anderen darin, dass lediglich eine Segmentation des gesamten Organs vorgenommen wurde, ohne eine differenzierte Segmentation der einzelnen Zonen der Prostata. Diese Information ist jedoch für die klinische Nutzung äußerst relevant, da bekanntermaßen viele pathologische Prozesse eine deutliche zonale Prädilektion zeigen (10).

Ein weiterer Ansatz ist die Anwendung sogenannter deformable shape based models. Hierbei ist die Idee, dass eine niedrig dimensionale Darstellung die Form und den Intensitätseindruck eines Objektes wiedergeben kann. Typischerweise werden dabei die kartesischen Koordinaten bestimmter Landmarken im Randbereich des Objektes definiert und mittels principal component analysis (PCA) der Verteilung der Landmarken ein statistisches Modell der Form des Objektes erzeugt. Dieser Ansatz nennt sich active shape model (ASM) (80). Dadurch finden sich jedoch nur Informationen über die Form der Begrenzung und nicht über den Inhalt des Objektes. Im Rahmen eines active appearance model (AAM) wird zusätzlich eine PCA der Signalintensitäten des Inhaltes des Objektes durchgeführt, um so zusätzliche Informationen über das Aussehen des Objektes zu erhalten. Toth und Madabhushi (80) haben beispielsweise Texturmerkmale aus Näherungskoeffizienten der Haar-Wavelet-Transformation für die Propagierung eines shape und active appearance model (AAM/ASM) zur Segmentierung der Prostata verwendet, wobei sie das traditionelle AAM-Modell um Intensitäts- und Gradienteninformationen erweiterten. Zudem verwendeten sie ein levelset, um die Form der statistischen Modellinformationen mit einem einem multifeature landmark-free framework zu erfassen (80). Die Segmentierungen in dieser Studie wurden anhand vorverarbeiteter 2D Slices einer T2W MRT Untersuchung der Prostata durchgeführt. Insgesamt wurden 108 Untersuchungen analysiert. Als Referenz wurde die Segmentation händisch durch einen Radiologen mit Expertise auf dem Gebiet durchgeführt. Um ein Interrater Agreement zu berechnen wurden 17 Untersuchungen von einem zweiten Experten segmentiert. Der resultierende mittlere DSC in dieser Studie lag bei 0,88 und das Interrater Agreement der beiden Experten bei 0,89 (80). Die Fallzahl von 108 MRT Untersuchungen ist in dieser Studie der hier vorliegenden Arbeit ähnlich. Jedoch erscheint die angewendete Technik nicht praktikabel im klinischen Setting, da es

einer extensiven Vorverarbeitung der Bilder bedarf. Es kann zudem keine dreidimensionale Information aus den Daten gezogen werden. Darüber hinaus bedarf dieses Verfahren eines gewissen Maßes an menschlicher Intervention. Zuletzt muss auch hier angeführt werden, dass lediglich eine Segmentierung des gesamten Organs durchgeführt wurde und nicht die zonale Anatomie getrennt betrachtet wurde.

Einen Schritt weiter in Richtung einer zonalen Segmentation geht die Arbeit von Makni et al. (81), welche als erstes über die Segmentierung von PZ und zentraler Drüse (oder auch central gland = CG) der Prostata berichteten. Sie verwendeten eine modifizierte Version des bekannten Evidential-C-Means-Algorithmus zum Clustern der Voxel in ihre jeweiligen Zonen, indem sie zusätzliche multiparametrische MR-Bildgebungsinformationen nutzten, welche die Analyse der T2W, der DWI und der DCE erforderten. In einem Segmentierungskontext modelliert dieses Model Pixel = Voxel als unabhängige Muster und es sind keine Kontextinformationen mit integriert. Die Anpassung dieses Klassifikators an ein Segmentierungsschema erfordert daher die die Einbringung von räumlichen Nachbarschaftsinformationen in den Klassifizierungsprozess der Voxel. Daher wurde eine händische Vorsegmentation der gesamten Prostata für die anschließende Zuordnung der darin enthaltenen Voxel zu jeweils PZ und CG vorangestellt. Getestet wurde anhand von multiparametrischen Prostata MRT Untersuchungen von 31 Patienten. Die Untersuchungen wurden zugleich von drei menschlichen Experten händisch segmentiert und eine ground truth berechnet. Der mittlere DSC des Modells für die PZ und CG lag bei 0,76 und 0,87. Das DSC zwischen den menschlichen Experten lag bei 0,91 – 0,93 für die CG und 0,86 – 0,90 für die PZ. In der Arbeit von Makni et al. musste jedoch menschlich interveniert werden, da die zonale Segmentation nur basierend auf einer bereits vorher erfolgten händischen Segmentation des gesamten Organs funktionierte (81). Diese händische Segmentation der Prostata ist jedoch im klinischen Setting zeitaufwändig und zudem wieder der subjektiven Einschätzung und Erfahrung des Untersuchers unterworfen. Zudem kann die Notwendigkeit einer Einbeziehung der DWI und DCE Sequenzen in die Segmentation in Frage gestellt werden, da diese Sequenzen lediglich funktionale Informationen für eine nähere Gewebeklassifikation liefern und im Vergleich zur T2W in Bezug auf die glanduläre Anatomie eine Redundanz darstellen. Für eine Segmentation von Tumorarealen, wie sie im Rahmen dieser Arbeit ebenfalls getestet wurde, oder anderen

pathologischen Gewebsveränderungen erscheint dieser Ansatz hingegen sinnvoll. Zuletzt kann als Kritikpunkt an der Studie von Makni et al. im Vergleich zu der hier vorliegenden Arbeit noch angeführt werden, dass die getesteten MRT Untersuchungen von Makni et al. an einem 1,5 T Scanner mit einer Körperoberflächenspule durchgeführt wurden. Hier sind Qualitätseinbußen in Auflösung und Signal zu erwarten, was in der hier vorliegenden Arbeit durch Verwendung eines 3T Scanners der neuesten Generation vermieden wurde. Ein sehr ähnliches Herangehen basierend auf einem C-Means Clustering wurde auch von Chilali et al. (82) verfolgt. Hier wurde jedoch lediglich anhand von T2W Bildern segmentiert und mit einem vergleichsweise kleinen Datensatz von 22 MRT Untersuchungen, welche von einem einzelnen Experten zum Vergleich händisch segmentiert wurden. Daher konnte in dieser Arbeit auch kein Interrater-Agreement berechnet werden (82).

Ein Weiterer Ansatz der zonalen Segmentation wurde in der Arbeit Litjens et al. (83) verfolgt, die eine Methode der pattern recognition zur Klassifizierung der Voxel anhand anatomischer, Intensitäts- und Texturmerkmale in multiparametrischen MR-Bildern verwendeten. Ein ähnlicher Ansatz wurde zuvor auch in einigen anderen Segmentierungsansätzen an anderen Organen verwendet. Auch in dieser Arbeit wurde eine händische Segmentation des gesamten Organs zu Grunde gelegt (83). Für den pattern recognition Ansatz wird zunächst die anatomische Information eines jeden Voxels herangezogen. Diese basiert auf der Kenntnis, dass beispielsweise die PZ in der Prostata immer dorsal liegt und zum Apex hin breiter wird. Darauf aufbauend und auf der Kenntnis der Begrenzung des gesamten Organs auf Grund der vorliegenden Segmentation wurden mehrere Positions- und Abstands-Features festgelegt. Anhand der Position eines Voxels in ventrodorsaler, mediolateraler und craniocaudaler Richtung sowie seines Abstandes zur Organgrenze konnten hier entsprechende Features festgelegt werden. Über die Signalintensitäten in der T2W wurden zugleich die Muskeln um die Prostata herum identifiziert. Zuletzt werden insgesamt fünf Texturmerkmale herangezogen. Zunächst die Homogenität und die Korrelation, welche über eine co-occurrence matrix berechnet wurde. Für das Histogramm wurden 16 Graustufen in allen 2D Richtungen gemittelt. Darüber hinaus als drittes und viertes Merkmal wurden die Entropie und die Texturstärke anhand der neighborhood gray-tone difference Matrix verwendet. Auch hier wurden 16 Graustufen benutzt. Als letztes Feature wurde die lokale Binarität eines jeden

Voxels betrachtet. Hierfür wurde die Auflösung der Bilder mittels eines Gauss'schen Resamplings reduziert. Nachdem alle genannten features berechnet waren, wurde ein balancierter Trainingssatz aufgebaut. Verglichen wurde zum einen mit einer Segmentation mittels multi-atlas approach und händischer Segmentierung durch drei menschliche Experten, welche zu einer ground truth zusammengefasst wurde. Zur Validierung lagen 48 MRT Untersuchungen der Prostata vor. Das pattern recognition model erreichte dabei einen DSC von 0,89 für die CG und 0,75 für die PZ im Vergleich zu einem Interrater DSC von 0,95 – 0,97 (83). Auch in dieser Arbeit ist eine menschliche Intervention in Form einer vorherigen händischen Segmentation der gesamten Drüse sowie eines post-processings, um die initiale Abgrenzung der Zonen zu glätten. Insofern kann hier lediglich von einem semi-automatisierten Ansatz gesprochen werden, was anders als der voll automatisierte Ansatz der hier vorliegenden Arbeit, einen deutlichen zeitlichen Mehraufwand in der klinischen Routine und eine höhere Fehleranfälligkeit durch fehlende Erfahrung des Readers und Subjektivität mit sich bringt. Darüber hinaus kann auch hier der zusätzliche anatomische Informationsgehalt der verwendeten ADC Maps für die Segmentierung in Frage gestellt werden.

Ein etwas anderes Vorgehen wurde in der Arbeit von Qiu et al (84) verfolgt, in welcher aus einem 3D Bilddatensatz einer T2W MRT Untersuchung der Prostata eine Segmentation des gesamten Organs sowie der zentralen Drüse und der peripheren Zone durchgeführt wurde. Hierfür nutzten die Autoren ein spatially continuous max-flow model. Es wurden hierbei keine Informationen aus funktionalen Sequenzen mit einbezogen, sondern lediglich die T2W segmentiert, was dem Ansatz dieser Arbeit gleicht. Bei dem in der Arbeit von Qiu et al. verwendeten continuous max-flow model handelt es sich um ein global optimierungsbasiertes Segmentationsverfahren mit einer geometrischen Beschränkung und einem object appearance model (84). Die Methode verwendet hierbei eine vorbestehende Kenntnis über die Zusammensetzung der Region, nämlich beispielsweise, dass die CG und PZ innerhalb der gesamten Drüse liegen und sich nicht überlappen, und verwendet ein angepasstes Prostata-Erscheinungsbildmodell aus intensity probability density functions, um mehrere Regionen gleichzeitig zu segmentieren. (84). Für das appearance model bedarf es jedoch einer user initialization mit Referenzpunkten der Prostata und der einzelnen Zonen, was wiederum eine menschliche Interaktion mit dem Modell bedeutet und einen gewissen Aufwand erfordert.

Darüber hinaus ist dadurch auch eine gewissen Subjektivität und Abhängigkeit vom Untersucher gegeben, was in der Studie von Qiu et al. jedoch zu keiner signifikanten Inter- und Intrareader-Variabilität geführt hat. Vorteil dieser Arbeit ist jedoch die fehlende Notwendigkeit eines Preprocessings der Bilder und die direkte Verwendung des 3D Bilddatensatzes, was ähnlich zu dem Ansatz der hier vorliegenden Arbeit ist. Auch die alleinige Verwendung der T2w erscheint sinnvoll, da es die primäre Sequenz zur Extraktion anatomischer Informationen ist und Rechenkapazität und Datenmenge gespart wird, wenn auf die zusätzliche Verwendung der funktionalen Sequenzen verzichtet werden kann. Der Algorithmus von Qiu et al. wurde anhand von 18 T2W MRT Untersuchungen getestet, welche mit einer Oberflächenspule akquiriert wurden, sowie 25 MRT Untersuchungen, welche mit einer Endorektalspule akquiriert wurden. Die Auswertungen erfolgten separat, da die Untersuchungen mit Endorektalspule zwar eine bessere Bildqualität, jedoch auch höhere Artefakte beinhalteten. Alle Untersuchungen wurden an einem modernen 3T Scanner durchgeführt, was daher gut vergleichbar zu dem Testdatensatz dieser Arbeit erscheint (84). Letztlich lassen sich am besten die Ergebnisse der Untersuchungen unter Verwendung einer Oberflächenspule vergleichen. Hier kam der Algorithmus von Qiu et al. auf einen DSC von 0,89 für die gesamte Drüse sowie 0,82 für die CG und 0,69 für die PZ (84).

In den letzten Jahren traten immer mehr neuronale Netzwerke in den Vordergrund und insbesondere Tiefe CNN haben sich inzwischen als die dominanteste und vielversprechendste Technik für automatisierte Segmentierungsaufgaben durchgesetzt. CNN basieren auf der Extraktion und Analyse von Imagefeatures in hierarchischer Struktur und legen dabei insbesondere Gewicht auf die Detektion der Kanten des zu segmentierenden Objektes. Sie wurden für verschiedene Segmentierungsaufgaben in unterschiedlichen Bildgebungsmodalitäten verwendet. So nutzten beispielsweise Christ et al (85) ein fully convolutional neural network und Li et al (86) ein hybrid densely connected UNet für die Segmentierung der Leber und von Lebertumoren sowohl in MRT- als auch CT-Bildern. Vielerlei andere Organe und Strukturen des Körpers wurden in anderen Arbeiten segmentiert. Auch für MRT-Untersuchungen der Prostata gibt es zahlreiche Ansätze für die automatisierte Bildsegmentierung mit Hilfe von CNN. Zhu et al (87) setzten in einer früheren Studie von 2017 ein deeply supervised CNN für die Segmentation der gesamten Drüse anhand axialer T2W MRT Bilder ein. Das

Netzwerk leitet in früheren Phasen extrahierte Merkmale an spätere Phasen weiter, um Informationsverluste zu vermeiden. Damit die in frühen Phasen extrahierten Merkmale semantisch sinnvoll zu erhalten, wurden zudem auf jeder Stufe zusätzliche *deeply supervised layers* hinzugefügt. Dieses soll vor allem an den problematischen Arealen im Bereich des Apex und der Basis helfen, an denen die Grenzen der Drüse schwer abgrenzbar sind (87). In dieser Studie von Zhu et al. wurde die Segmentation an axialen T2W MRT-Bildern durchgeführt, wobei Schicht für Schicht zweidimensional segmentiert wurde, was dem dreidimensionalen Ansatz der hier vorliegenden Arbeit unterlegen ist, da beispielsweise in weiteren Schritten, wie der automatischen Tumordetektion die Information der angrenzenden Schichten nicht einbezogen werden kann. Zhu et al. trainierten ihr Netzwerk mit 68 MRT-Untersuchungen und testeten es lediglich an 4 zufällig ausgewählten Patienten. Während die Fallzahl für den Trainingsdatensatz zwar der hier vorliegenden Arbeit nahekommt, so ist doch der Testdatensatz sehr klein gewählt, was die Ergebnisse schwer vergleichbar macht. Zudem wurde lediglich die gesamte Drüse segmentiert und keine getrennte zonale Segmentation vorgenommen. Der DSC für die gesamte Drüse lag bei Zhu et al. bei etwa 0,89 (87). Gegenüber dem zweidimensionalen slice-by-slice Ansatz von Zhu et al. steht die Studie von Milletari et al. (88), welche ein end-to-end trainiertes, sogenanntes V-Net mit dreidimensionalen Faltungen verwendet, welches einen contracting und einen expanding path mit skip connections enthält, was der grundlegenden Architektur des U-Nets in dieser Arbeit ähnlich ist. Milletari et al. segmentierten jedoch ebenfalls nur die gesamte Drüse und nicht die einzelnen Zonen. Der mittlere DSC für die Segmentation der gesamten Drüse lag in ihrer Studie bei 0,87 (88). Der Vorteil eines U-Nets gegenüber einem klassischen CNN liegt bei Bildsegmentierungsaufgaben vor allem in der Schnelligkeit und räumlichen Konsistenz und der Anwendbarkeit auf Bildern beliebiger Größe. CNN führen bei Segmentierungsaufgaben eine Klassifikation Pixel für Pixel durch und es wird eine Region um jeden einzelnen Pixel extrahiert, welche Patch genannt wird. Die Segmentierung mittels solcher Patches durch ein CNN ist sehr zeitaufwendig, da sie sehr viele redundante Rechenoperationen erfordert. Darüber hinaus können in Abhängigkeit von der Größe der Patches lokale und globale Image Features nicht gleichzeitig festgehalten werden (89). Das klassische U-Net nach Ronneberger et al. beinhaltet einen encoder Pfad, welcher die Architektur eines klassischen CNN widerspiegelt, und einen decoder Pfad, in welchem die pooling layer des klassischen CNN ersetzt sind durch

transposed convolutions um die Dimensionalität des feature space zu erhöhen und eine höhere Lokalisierungsgenauigkeit zu erlangen (73).

Auch für die Segmentierung von Prostata Bilddatensätzen wurden bereits U-Nets eingesetzt. Rundo et al. (90) vergleichen beispielsweise verschiedene Netzwerkarchitekturen für die zonale Segmentierung der Prostata in T2W-Bildern miteinander. Das geteteste U-Net war hierbei in der Lage die Genauigkeit der beiden anderen CNN-Architekturen SegNet und pix2pix zu übertreffen (90).

Zwar existieren bislang noch recht wenige Studien, welche sich mit der automatisierten Segmentierung der einzelnen Zonen der Prostata befassen, Zabihollahy et al. (89) benutzten jedoch 2019 beispielsweise zwei parallelgeschaltete U-Nets, um anhand von T2W MRT-Bildern und ADC Maps der Prostata sowohl eine Segmentierung der gesamten Drüse, als auch der CG und der PZ vorzunehmen. Die Stärken dieser Arbeit liegen zum einen in dem verhältnismäßig großen Datensatz von 225 MRT-Untersuchungen sowie der Subgruppenanalyse an Patienten mit und ohne signifikanten Prostatakarzinomen in der Bildgebung, da die Methodik der Segmentierung sowohl für gesunde als auch erkrankte robust funktionieren sollte. Die Hinzunahme der ADC Map in der Segmentierung wird von den Autoren damit begründet, dass eine Segmentierung der ADC Map sinnvoll sei, da anhand ihrer die Detektion von Karzinomen in der bpMRT vorgenommen wird und dieses daher für zukünftige automatisierte Detektionstools wichtig sei (89). In der hier vorliegenden Studie wurde lediglich die axiale T2W-Sequenz für die zonale Annotation segmentiert und die ADC für die Karzinom-Segmentierung verwendet, wodurch eine gute Detektion der Tumore möglich. Der DSC für die Segmentierung der gesamten Drüse, der CG und der PZ lagen in der Arbeit von Zabihollahy et al. bei 0,96, 0,93 und 0,86. Als Grundwahrheit wurde eine händische Segmentierung von verschiedenen radiologischen Fachärzten mit unterschiedlichen Levels an Erfahrung genommen, die jedoch jeweils unterschiedliche Teile des Datensatzes segmentiert haben. Jedes Bild wurde hierbei nur von einem Radiologen segmentiert, was zu einer Heterogenität der Grundwahrheit führt. Dennoch sind die Ergebnisse sicherlich als sehr gut zu bezeichnen.

Rundo et al. (91) erweiterten ein U-Net um sogenannte squeeze-excitation (SE) Blöcke und verwendeten in ihrer Studie drei Datensätze aus verschiedenen Institutionen, was

die Generalisierbarkeit und Robustheit der Methodik unterstreicht. Das Prinzip basiert auf einer adaptiven Wichtung der einzelnen Channels und damit dem Hervorheben relevanter Features bei der zonalen Segmentation. Ein Ähnliches Prinzip liegt auch dem Netzwerk der hier vorliegenden Arbeit zu Grunde. Die Gesamtzahl der Patienten für die Studie von Rundo et al. betrug dabei 80, was deutlich weniger als der Datensatz der hier vorliegenden Arbeit ist. Rundo et al. verglichen ihr Netzwerk zu anderen gängigen Architekturen, welche in diesem Abschnitt ebenfalls bereits diskutiert wurden und konnten mit ihrem Enc-Dec-USE-Net bessere Ergebnisse erzielen. Der DSC lag für die drei Datensätze jeweils für die CG zwischen 0,87 und 0,93 sowie für die PZ zwischen 0,8 und 0,92 (91).

Meyer et al. (92) verwendeten ebenfalls ein 3D-U-Net ähnlich dem in dieser Arbeit verwendeten Netzwerk. Für die Segmentierungen wurden in ihrer Studie jedoch auch die anderen orthogonalen Ebenen in T2W verwendet, um das Problem der Anisotropie der Voxel in rein axialen Bildstapeln zu beheben. Auf diese Weise konnten Meyer et al. hochaufgelöste isotrope Segmentierungen der Prostata vornehmen. Dabei arbeiteten sie mit einem öffentlichen Datensatz mit insgesamt 40 Untersuchungen. Eine Stärke dieser Studie ist die zusätzliche Berechnung der Hausdorff Distanz und Oberflächendistanz als zusätzliche Parameter der Performance Bewertung des Netzwerkes. Leider wurde lediglich die Segmentation der gesamten Drüse vorgenommen und nicht die der einzelnen Zonen, was keinen hinreichenden Vergleich zu der hier vorliegenden Arbeit zulässt.

Die automatisierte Segmentation der Prostata bietet eine der Grundlagen für die Detektion und Segmentation von Prostatakarzinomen. Wang et al. (93) verglichen in ihrer Studie ein CNN mit einem nicht Deep Learning basierten Ansatz für die automatisierte Detektion von Prostatakarzinomen. Hierbei konnten sie feststellen, dass der Deep Learning basierte Ansatz des CNN überlegen ist. Die area under the receiver operating characteristics curve (AUC) betrug 0,84 für das CNN und 0,70 für den nicht Deep Learning basierten Ansatz.

In der Studie von Hosseinzadeh et al. (94) wurde ebenfalls ein Deep Learning Algorithmus für die Erkennung von PI-RADS ≥ 4 verwendet, wobei eine AUC von 0,88 und eine Sensitivität von 87% erreicht werden konnte. Zugleich konnte festgestellt werden, dass die vorherige Kenntnis der zonalen Anatomie in Form einer Segmentation signifikant zur Genauigkeit der Detektion beiträgt. Machireddy et al. (95) nutzten ein U-

Net mit Implementation spezieller Aufmerksamkeiten für die Segmentation der Karzinomregion anhand der T2W und ADC Map. Sie konnten hierbei einen DSC von 0,59 erzielen. Es lässt sich somit auch für die Karzinomdetektion und -Segmentation erkennen, dass neuronale Netzwerke die vielversprechendste Methodik darstellen. Zugleich wird die Relevanz der zonalen Segmentation als Grundlage unterstrichen.

4.3 Diskussion der Ergebnisse

In dieser Studie wird ein trainiertes Deep Learning Modell für die automatisierte Segmentierung der zonalen Anatomie und karzinomsuspekter Areale in der biparametrischen MRT der Prostata vorgestellt. Die beste Segmentierungsleistung wurde hierbei für die CG erzielt, gefolgt von der PZ und Bereichen mit Verdacht auf Prostatakrebs. Diese Verteilung der Genauigkeit spiegelt sich auch in den vorliegenden Vergleichsstudien in der Literatur wider (77 – 91) und korrespondiert mit der Komplexität der jeweiligen geometrischen Formen. Demnach ist die zumeist einfach sphärisch geformte CG simpler zu erfassen als die etwas komplexer und von Schicht zu Schicht heterogener geformten PZ. Demgegenüber am schwierigsten begreifbar sind die Tumorareale, welche teilweise bizarr amorph gestaltet sind.

Die in dieser Arbeit erzielten Ergebnisse für die Segmentierung der Tumorareale ähneln denen, die in der neueren Literatur veröffentlicht wurden, und stellen einen realistischen Maßstab dar. De Vente et al. (96) berichteten über einen DSC von 0,37 +/- 0,05 für die Segmentierung von klinisch signifikanten Karzinomen nach fünffacher Kreuzvalidierung, was geringer ist als die in dieser Studie erzielten Ergebnisse. Die Herausforderung der Segmentation karzinomsuspekter Areale liegt hauptsächlich in ihrer Heterogenität sowohl bezogen auf ihre Form und Begrenzung als auch auf die Signalintensitäten in den verschiedenen Sequenzen. Auch für menschliche Experten stellt die genaue Erfassung der Begrenzung karzinomsuspekter Areale eine Hürde dar, was sich auch in dieser Arbeit anhand der Interrater Übereinstimmung erkennen lässt, welche für die Segmentierung der Tumorareale deutlich geringer ausfiel als für die CG und PZ (siehe Tabelle 1). Während in einigen Arbeiten sehr hohe Sensitivitäten in der Detektion von Karzinomen erzielt werden konnten (92 – 94), so stellt die Segmentierung doch eine größere Schwierigkeit dar. Dennoch scheint eine Segmentierung auch hier für den klinischen Alltag sinnvoll, beispielsweise für die Volumetrie zur genaueren Verlaufsbeurteilung von

Karzinomen unter active surveillance oder auch für die Durchführung von MRT gestützten Fusionsbiopsien.

Für die Segmentierung der Anatomie der Prostata und ihrer einzelnen Zonen lag die Interrater Übereinstimmung in einem guten Bereich, welcher in etwa mit den berichteten Werten aus der Literatur übereinstimmt (79, 80). Es wurde zwar in manchen Studien von Makni et al. und Litjens et al. (81, 83) auch von Interrater Übereinstimmungen mit einem DSC von über 0,9 berichtet, jedoch wurde hier teilweise in öffentlichen Datensätzen segmentiert, welche zuvor im Hinblick auf besonders hohe Qualität der Bildgebung ausgewählt und zusammengestellt wurden. Daher sind manche Einschränkungen wie Artefakte oder Signalinhomogenitäten gegebenenfalls weniger vertreten, was die Qualität der Segmentationen beeinflussen kann.

Für die Segmentierung der Prostata durch das Deep Learning Modell sind die in der vorliegenden Studie berichteten Ergebnisse im Kontext früherer Studien in Bezug auf den DSC im Mittelfeld angesiedelt. Zwar konnten bessere Ergebnisse erzielt werden als in den Arbeiten von Klein et al. (79), Toth et al. (80), Makni et al. (81) und Qiu et al. (84), jedoch auch etwas niedrigere als in einigen anderen der früheren Studien. So erreichten Sanford et al. einen DSC von 0,92 für die gesamte Prostata und 0,90 für die TZ (97), während Bardis et al. einen mittleren DSC von 0,94 für die gesamte Prostata im Vergleich zu 0,91 für die TZ und 0,77 für die TZ berichteten (98). Tabelle 3 zeigt die erzielten DSC Werte der diskutierten früheren Studien und der hier vorliegenden Arbeit im Vergleich. Zwar lagen die DSC Werte dieser Arbeit teilweise etwas unter denen aus einigen der Vergleichsstudien, jedoch zeigte diese Arbeit auch, dass der DSC Koeffizient allein keine genaue Bewertung der Modelleistung ermöglicht. Während für die zentrale Drüsensegmentierung ein DSC Koeffizient von 0,88 erreicht werden konnte, der um 0,01 höher lag als die Interrater Übereinstimmung, was auf eine dem Menschen vergleichbare Leistung hindeutet, zeigten die Hausdorff Distanz und die mittlere Oberflächendistanz immer noch an, dass das Modell der menschlichen Leistung etwas unterlegen war. Daher erscheint es generell ratsam, alle drei vorgeschlagenen Metriken einzubeziehen, um eine ausgewogene Bewertung der Segmentierungsleistung zu erhalten.

Studie	DSC ganze Drüse	DSC CG	DSC TZ
Qiu et al.	-	0,82	0,69
Klein et al.	0,85	-	-
Makni et al.	-	0,87	0,76
Toth et al.	0,88	-	-
<i>Adams et al.</i>	-	<i>0,88</i>	<i>0,76</i>
Litjens et al.	-	0,89	0,75
Sanford et al.	0,9	-	-
Bardis et al.	0,94	0,91	0,77
Zabihollahy et al.	-	0,93	0,86

Table 3. Vergleich der Ergebnisse vorheriger Studien mit der hier vorliegenden Arbeit.

4.4 Übersicht der Stärken und Limitationen

Als eine Stärke des verwendeten Deep Learning Netzwerkes ist ganz eindeutig die parallele Verwendung biparametrischer Bildinformationen im Training zu erwähnen, welche der realen Befundung durch einen menschlichen Befunder gleicht. Hierbei werden durch das Netzwerk für die Karzinomsegmentation gleichzeitig Informationen aus der T2W-Sequenz sowie den DWI-Sequenzen und der ADC Map genutzt. Dem gegenüber wird in vielen anderen Studien lediglich eine Sequenz verwendet oder es werden separat in verschiedenen Trainingsdurchläufen die unterschiedlichen Sequenztypen nacheinander in das Netzwerk eingebracht.

Insbesondere in der Detektion von Prostatakarzinomen kann der parallele biparametrische Trainingsansatz als vorteilhaft bewertet werden. In diesem Zusammenhang könnte man gleichzeitig allerdings als mögliche Schwäche anführen, dass in dieser Arbeit keine kontrastmittelverstärkten DCE-Sequenzen einbezogen wurden. Die DCE bietet in der klinischen Routine zwar durchaus wertvolle Informationen über die pathologische Neovaskularisation der Tumorregion, spielt jedoch zunehmend eine untergeordnete Rolle. Dieses lässt sich auch an der aktuellen PI-RADS 2.1 Richtlinie erkennen, in welcher die DCE lediglich zur Differenzierung einer unklaren Läsion mit einem PI-RADS 3 Score in der DWI-Bildgebung beiträgt. In diesem Falle führt ein nachweisbares Frühenhancement des Kontrastmittels zu einem Upgrading auf eine PI-RADS 4 Läsion. Jedoch besagt auch die PI-RADS 2.1 Richtlinie, dass bei guter

Bildqualität der T2W- und DWI-Sequenzen in der Regel auf eine DCE verzichtet werden kann (35). Dies ist insbesondere auf den finanziellen Aspekt, die Untersuchungsdauer und Sicherheitsbedenken in Bezug auf Gadolinium haltige Kontrastmittel zurückzuführen. Aus diesem Grunde sehen wir den biparametrischen Ansatz dieser Arbeit als durchaus passend für die Realität im klinischen Alltag an.

Eine weitere Stärke liegt sicherlich in der Qualität der zu Grunde liegenden menschlichen Annotation der Prostata MRT Untersuchungen für das Training des Netzwerkes. Im Rahmen dieser Studie wurden die Bilddaten von zwei Fachärzten für Radiologie mit mehrjähriger Erfahrung in der Befundung multiparametrischer MRT der Prostata annotiert. Beide Beurteiler haben jeweils eine Subspezialisierung auf Q1-Niveau nach Qualitätsstandard der AG für Uroradiologie der DRG. Aus diesem Grunde kann sowohl von einer guten Qualität der Annotationen des Trainingsdatensatzes sowie von einer soliden Grundwahrheit des Testdatensatzes ausgegangen werden.

Die einfache Anwendbarkeit des Netzwerkes in einem potenziellen klinischen Setting sollte ebenfalls hervorgehoben werden. Für das Training und die Nutzung des Netzwerkes zur Segmentation oder Karzinomdetektion kann ein regulärer DICOM Bildstapel aus einem klinischen Bildarchiv exportiert und nach schneller Konvertierung als kompletter Bildstapel in das Programm eingeladen werden. Im Weiteren findet eine voll automatisierte Vorverarbeitung der Bilder statt, welche keiner menschlichen Intervention bedarf. Die Bilder werden automatisch zugeschnitten und etwa 20% unnütze Bildinformation vom Rand verworfen. Dadurch wird Rechenkapazität und Zeit bei der weiteren Verarbeitung gespart. Zudem werden die Bilder resampled und einzelne Bildpunkte auf gemeinsame Koordinaten gesetzt, um die parallele Betrachtung der DWI und der T2W durch das Netzwerk zu ermöglichen. Außerdem wird das Training des Algorithmus durch eine automatisierte Augmentation der Bilder unterstützt. Hierbei werden Alterationen der Bilder wie ein willkürliches, künstliches Rauschen, eine Verzerrung der Bilder, künstliche Bildartefakte und Schwankungen der Signalintensitäten vorgespielt, um das Netzwerk robuster zu machen. Ziel ist es hierbei, dass das Netzwerk lernt, gewisse, für die Segmentation irrelevante Störfaktoren in den Bildern zu ignorieren. Die Anwendung von Bildaugmentationen im Training ist ein Vorteil für eine zukünftige Anwendung des Netzwerkes auf klinische Bilddaten, bei denen Artefakte oder etwaige

Einschränkungen der Bildqualität eine Rolle spielen. Auch hierin liegt eine Stärke unserer Studie, denn es wurden Fälle aus der klinischen Routine verwendet, welche entsprechenden Qualitätseinschränkungen wie Artefakten durch Bewegung oder Fremdmaterial und Luft unterliegen und teils komplexe Anatomien auf Grund von pathologischen Prozessen wie BPH, Entzündung, Einblutung oder Tumoren aufweisen. Darin unterscheiden sich klinische Daten oft von öffentlichen Datensätzen, welche meist ausschließlich vorselektierte Bilder in hoher Qualität enthalten. In dem Sinne liegt der hier verwendete Datensatz näher an der Alltagsrealität der Radiologie. So kann beispielsweise eine stark ausgeprägte BPH mit massiver Hypertrophie der TZ zu einer nahezu vollständigen Verdrängung der PZ mit starker Ausdünnung derselben führen. Dieses macht eine Erkennung und verlässliche Segmentation der PZ und auch die Detektion dort vorliegender Tumore schwierig. Zudem kann die zonale Anatomie durch zonenübergreifende, große Tumorprozesse mit diffuser Infiltration gestört sein, was die Abgrenzung der einzelnen Zonen voneinander und auch die Begrenzung der Organkapsel für eine Segmentation komplex oder gar unmöglich macht. Hierin liegt jedoch auch eine Schwäche der Arbeit, denn diese Inhomogenität des Datensatzes mit entsprechenden Komplikationen erschwert das Training des Netzwerkes und verringert die Genauigkeit der Segmentation im Vergleich zu anderen Studien in der Literatur, welche sich auf öffentliche Datensätze stützen. Dieser Aspekt wurde bereits in der Diskussion der Ergebnisse angesprochen.

Zuletzt muss erwähnt werden, dass es sich bei unserem Datensatz um eine eher kleine Stichprobe handelt und das Training des Netzwerkes auch hierdurch etwas ausgebremst wird. Eine mögliche Perspektive wäre hier in Zukunft das Hinzuziehen öffentlicher Prostata Bilddatensätze. Dadurch sollte sich die Performance der Segmentation und insbesondere der Karzinomdetektion weiter steigern lassen. Dennoch sollte erwähnt werden, dass der hier verwendete und annotierte Datensatz im Vergleich zur vorliegenden Literatur durchaus als groß angesehen werden kann. Insbesondere die Anzahl der histologisch gesicherten Prostatakarzinome ist hierbei hervorzuheben.

4.5 Schlussfolgerungen und Ausblick

Zusammenfassend kann anhand der vorliegenden Arbeit und auch der zuvor erwähnten und zitierten anderen Studien gezeigt werden, dass mit Hilfe künstlicher Intelligenz eine

voll automatisierte Segmentation der Prostata und ihrer zonalen Anatomie möglich ist und durchaus auch die Qualität und Sicherheit menschlicher Expertenstandards erreichen kann. Insbesondere unsere Arbeit zeigt hierbei auch den Weg in eine klinische Alltagstauglichkeit auf, da der menschliche Interventionsbedarf, die Vorverarbeitung der Bilddaten durch den Befunder und auch die Filterung der Bilddaten nach Qualität bzw. der Ausschluss niederqualitativer Untersuchungen nicht notwendig sind. Damit ergibt sich eine realistische Perspektive für die Unterstützung des Radiologen in seinem klinischen Alltag durch den Algorithmus in sonst sehr zeitintensiven Tätigkeiten. Derzeit werden manuelle Segmentierungen auf Grund des Zeitaufwandes meistens in der Befundungsroutine nicht durchgeführt. Eine Segmentierung ist jedoch durchaus von Nutzen zur Erstellung dreidimensionaler Modelle beispielsweise zu Navigationszwecken in der Planung und Durchführung von diagnostischen Punktionen oder interventionellen Therapien sowie ggf. auch für die Roboter assistierte Chirurgie, welche gerade für die Prostata zunehmend an Bedeutung gewinnt. Die vollautomatisierte Segmentation kann hier also einen entscheidenden Beitrag leisten.

Die automatisierte Detektion und Segmentation von Tumoren kann zukünftig sicherlich anhand größerer Datensätze in ihrer Genauigkeit noch signifikant gesteigert werden und dann ebenfalls ein hilfreiches Werkzeug für den Befundersteller darstellen, indem beispielsweise auf suspekte Areale hingewiesen wird, die dann durch den menschlichen Reader näher bewertet werden können. Zudem spielt die aktive Überwachung bei niedriggradigen Karzinomen eine wichtige Rolle und kann ebenfalls entscheidend durch die künstliche Intelligenz unterstützt werden. So kann durch die Segmentation des Tumors eine Volumenberechnung erfolgen, die im Weiteren eine genauere Verlaufsbeurteilung hinsichtlich eines möglichen Progresses ermöglicht im Vergleich zu einer konventionellen Durchmesser messung. Darüber hinaus unterliegt das Netzwerk nicht den subjektiven Einflüssen durch zwei potenziell unterschiedliche menschliche Befundersteller für die Untersuchungen, welche möglicherweise in ihrer Art zu messen differieren und somit zu verschiedenen Ergebnissen kommen, was die Verlaufsbeurteilung in ihrer Zuverlässigkeit deutlich einschränkt.

Letztlich bietet sich als weiterer Ausblick noch der Nutzen der automatisierten Segmentation des Organs und eines eventuell vorliegenden Tumors für weitere darauf

aufbauende Analysen. So könnten beispielsweise Analysen bestimmter Bildfeatures ergänzt werden, um eine Vorhersage über das Grading des Tumors oder eine Risikostratifizierung für kapselüberschreitendes Wachstum vorzunehmen. Hierfür könnten anatomische Informationen aus der T2W und auch Signalintensitäten herangezogen werden. Darüber hinaus könnten auch quantitative Bildparameter wie die numerischen Werte der ADC Map verwendet werden. Auch die Verwendung von Mapping Sequenzen für die quantitative Bestimmung der T1- und T2-Relaxationszeiten könnten hinzugenommen werden, um diese in die Analyse durch das Netzwerk einzubeziehen. Hierzu liegen bereits Daten bezüglich eines Nutzens dieser Parameter für das non-invasive Tumorigradung vor (99).

Literaturverzeichnis

1. Coakley FV, Hricak H. Radiologic anatomy of the prostate gland: a clinical approach. *Radiologic Clinics of North America*. 2000;38(1):15-30.
2. Lee CH, Akin-Olugbade O, Kirschenbaum A. Overview of prostate anatomy, histology, and pathology. *Endocrinology and metabolism clinics of North America*. 2011;40(3):565-75, viii.
3. Hammerich KH, Ayala GE, Wheeler TM. *Anatomy of the prostate gland and surgical pathology of prostate cancer*. Cambridge University, Cambridge. 2009:1-10.
4. Chen H, Zhang CA, Gill HS. Anatomical study of the prostatic urethra using vinyl polysiloxane casts. *The Prostate*. 2020;80(3):241-6.
5. Aaron L, Franco OE, Hayward SW. Review of prostate anatomy and embryology and the etiology of benign prostatic hyperplasia. *Urologic Clinics*. 2016;43(3):279-88.
6. Ittmann M. *Anatomy and histology of the human and murine prostate*. Cold Spring Harbor perspectives in medicine. 2018;8(5):a030346.
7. Ayala AG, Ro JY, Babaian R, Troncoso P, Grignon DJ. The prostatic capsule: does it exist? Its importance in the staging and treatment of prostatic carcinoma. *The American journal of surgical pathology*. 1989;13(1):21-7.
8. Viliers A, Steg A, Boccon-Gibod L. *Anatomy of the prostate: review of the different models*. *European urology*. 1991;20:261-8.
9. Lowsley OS. The development of the human prostate gland with reference to the development of other structures at the neck of the urinary bladder. *American journal of Anatomy*. 1912;13(3):299-349.
10. McNeal JE. Regional morphology and pathology of the prostate. *American journal of clinical pathology*. 1968;49(3):347-57.
11. Mcneal JE. The prostate and prostatic urethra: a morphologic synthesis. *The Journal of urology*. 1972;107(6):1008-16.

12. McNeal JE. Origin and evolution of benign prostatic enlargement. *Investigative urology*. 1978;15(4):340-5.
13. McNeal JE. The zonal anatomy of the prostate. *The prostate*. 1981;2(1):35-49.
14. Selman SH. The McNeal prostate: a review. *Urology*. 2011;78(6):1224-8.
15. Roehrborn CG. Benign prostatic hyperplasia: an overview. *Reviews in urology*. 2005;7(Suppl 9):S3.
16. Hricak H, Doms GC, McNeal JE, Mark AS, Marotti M, Avallone A, Pelzer M, Proctor EC, Tanagho EA. MR imaging of the prostate gland: normal anatomy. *American journal of roentgenology*. 1987;148(1):51-8.
17. Lokeshwar SD, Harper BT, Webb E, Jordan A, Dykes TA, Neal Jr DE, Terris MK, Klaassen Z. Epidemiology and treatment modalities for the management of benign prostatic hyperplasia. *Translational andrology and urology*. 2019;8(5):529.
18. Isaacs JT, Coffey DS. Etiology and disease process of benign prostatic hyperplasia. *The Prostate*. 1989;15(S2):33-50.
19. Vuichoud C, Loughlin KR. Benign prostatic hyperplasia: epidemiology, economics and evaluation. *Can J Urol*. 2015;22(Suppl 1):1-6.
20. Taub DA, Wei JT. The economics of benign prostatic hyperplasia and lower urinary tract symptoms in the United States. *Current Prostate Reports*. 2006;4(2):81-90.
21. BLACKLOCK NJ. Anatomical factors in prostatitis. *British journal of urology*. 1974;46(1):47-54.
22. Nickel JC. Prostatitis: myths and realities. *Urology*. 1998;51(3):362-6.
23. Meares Jr EM. Prostatitis. *The Medical clinics of North America*. 1991;75(2):405-24.
24. Krieger JN, Nyberg Jr L, Nickel JC. NIH consensus definition and classification of prostatitis. *Jama*. 1999;282(3):236-7.

25. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA: a cancer journal for clinicians*. 2015;65(1):5-29.
26. Barnes B, Kraywinkel K, Nowossadeck E, Schönfeld I, Starker A, Wienecke A, Wolf U. Bericht zum Krebsgeschehen in Deutschland 2016. Robert Koch-Institut; 2016.
27. Grönberg H. Prostate cancer epidemiology. *The Lancet*. 2003;361(9360):859-64.
28. Crawford ED. Epidemiology of prostate cancer. *Urology*. 2003;62(6):3-12.
29. Gerharz EW, Heidenreich A. Aktualisierte S3-Leitlinie Prostatakarzinom. Im Fokus Onkologie. 2022;25(1):22-5.
30. Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. *Cochrane Database of Systematic Reviews*. 2013(1).
31. Hayes JH, Barry MJ. Screening for prostate cancer with the prostate-specific antigen test: a review of current evidence. *Jama*. 2014;311(11):1143-9.
32. Smith RA, Andrews KS, Brooks D, Fedewa SA, Manassaram-Baptiste D, Saslow D, Wender RC. Cancer screening in the United States, 2019: a review of current American Cancer Society guidelines and current issues in cancer screening. *CA: a cancer journal for clinicians*. 2019;69(3):184-210.
33. Mottet N, Bellmunt J, Briers E, Van den Bergh R, Bolla M, Van Casteren N, Cornford P, Culine S, Joniau S, Lam T. Guidelines on prostate cancer. *Eur Urol*. 2014;65(1):124-37.
34. Humphrey PA. Histological variants of prostatic carcinoma and their significance. *Histopathology*. 2012;60(1):59-74.
35. Chatfield M, Hicks L, Kaur D, Vivian-Davis C. ACR PI-RADS: prostate imaging and reporting and data system. Version 2.1. Reston, VA, American College of Radiology; 2019.
36. Litwin MS, Tan H-J. The diagnosis and treatment of prostate cancer: a review. *Jama*. 2017;317(24):2532-42.

37. Mohler J, Bahnson RR, Boston B, Busby JE, D'Amico A, Eastham JA, Enke CA, George D, Horwitz EM, Huben RP. Prostate cancer. *Journal of the National Comprehensive Cancer Network*. 2010;8(2):162-200.
38. Hamm B, Asbach P. Magnetic resonance imaging of the prostate in the PI-RADS Era. *Diseases of the Abdomen and Pelvis 2018-2021*. 2018:99-115.
39. Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, Rouviere O, Logager V, Fütterer JJ. ESUR prostate MR guidelines 2012. *European radiology*. 2012;22(4):746-57.
40. Woo S, Suh CH, Kim SY, Cho JY, Kim SH, Moon MH. Head-to-head comparison between biparametric and multiparametric MRI for the diagnosis of prostate cancer: a systematic review and meta-analysis. *American Journal of Roentgenology*. 2018:W226-W41.
41. Stabile A, Giganti F, Rosenkrantz AB, Taneja SS, Villeirs G, Gill IS, Allen C, Emberton M, Moore CM, Kasivisvanathan V. Multiparametric MRI for prostate cancer diagnosis: current status and future directions. *Nature Reviews Urology*. 2020;17(1):41-61.
42. Akin O, Sala E, Moskowitz CS, Kuroiwa K, Ishill NM, Pucar D, Scardino PT, Hricak H. Transition zone prostate cancers: features, detection, localization, and staging at endorectal MR imaging. *Radiology*. 2006;239(3):784-92.
43. Mehralivand S, Shih JH, Harmon S, Smith C, Bloom J, Czarniecki M, Gold S, Hale G, Rayn K, Merino MJ. A grading system for the assessment of risk of extraprostatic extension of prostate cancer at multiparametric MRI. *Radiology*. 2019;290(3):709-19.
44. Johnson LM, Turkbey B, Figg WD, Choyke PL. Multiparametric MRI in prostate cancer management. *Nature reviews Clinical oncology*. 2014;11(6):346-53.
45. Hambrock T, Somford DM, Huisman HJ, van Oort IM, Witjes JA, Hulsbergen-van de Kaa CA, Scheenen T, Barentsz JO. Relationship between apparent diffusion coefficients at 3.0-T MR imaging and Gleason grade in peripheral zone prostate cancer. *Radiology*. 2011;259(2):453-61.

46. Pagani E, Bizzi A, Di Salle F, De Stefano N, Filippi M. Basic concepts of advanced MRI techniques. *Neurological Sciences*. 2008;29(3):290-5.
47. Zelhof B, Pickles M, Liney G, Gibbs P, Rodrigues G, Kraus S, Turnbull L. Correlation of diffusion-weighted magnetic resonance data with cellularity in prostate cancer. *BJU international*. 2009;103(7):883-8.
48. Wu L-M, Xu J-R, Ye Y-Q, Lu Q, Hu J-N. The clinical value of diffusion-weighted imaging in combination with T2-weighted imaging in diagnosing prostate carcinoma: a systematic review and meta-analysis. *American Journal of Roentgenology*. 2012;199(1):103-10.
49. Haider MA, Van Der Kwast TH, Tanguay J, Evans AJ, Hashmi A-T, Lockwood G, Trachtenberg J. Combined T2-weighted and diffusion-weighted MRI for localization of prostate cancer. *American journal of roentgenology*. 2007;189(2):323-8.
50. Engelbrecht MR, Huisman HJ, Laheij RJ, Jager GJ, van Leenders GJ, Hulsbergen-Van De Kaa CA, de la Rosette JJ, Blickman JG, Barentsz JO. Discrimination of prostate cancer from normal peripheral zone and central gland tissue by using dynamic contrast-enhanced MR imaging. *Radiology*. 2003;229(1):248-54.
51. Franiel T, Hamm B, Hricak H. Dynamic contrast-enhanced magnetic resonance imaging and pharmacokinetic models in prostate cancer. *European radiology*. 2011;21(3):616-26.
52. Puech P, Sufana-Iancu A, Renard B, Lemaitre L. Prostate MRI: can we do without DCE sequences in 2013? *Diagnostic and interventional imaging*. 2013;94(12):1299-311.
53. Berman RM, Brown AM, Chang SD, Sankineni S, Kadakia M, Wood BJ, Pinto PA, Choyke PL, Turkbey B. DCE MRI of prostate cancer. *Abdominal Radiology*. 2016;41(5):844-53.
54. Oto A, Yang C, Kayhan A, Tretiakova M, Antic T, Schmid-Tannwald C, Eggener S, Karczmar GS, Stadler WM. Diffusion-weighted and dynamic contrast-enhanced MRI of prostate cancer: correlation of quantitative MR parameters with Gleason score and tumor angiogenesis. *American Journal of Roentgenology*. 2011;197(6):1382-90.

55. Wang L, Mazaheri Y, Zhang J, Ishill NM, Kuroiwa K, Hricak H. Assessment of biologic aggressiveness of prostate cancer: correlation of MR signal intensity with Gleason grade after radical prostatectomy. *Radiology*. 2008;246(1):168-76.
56. Taghipour M, Ziaei A, Alessandrino F, Hassanzadeh E, Harisinghani M, Vangel M, Tempany CM, Fennessy FM. Investigating the role of DCE-MRI, over T2 and DWI, in accurate PI-RADS v2 assessment of clinically significant peripheral zone prostate lesions as defined at radical prostatectomy. *Abdominal radiology*. 2019;44(4):1520-7.
57. Ullrich T, Schimmöller L. Perspective: a critical assessment of PI-RADS 2.1. *Abdominal Radiology*. 2020;45(12):3961-8.
58. Bass E, Pantovic A, Connor M, Gabe R, Padhani A, Rockall A, Sokhi H, Tam H, Winkler M, Ahmed H. A systematic review and meta-analysis of the diagnostic accuracy of biparametric prostate MRI for prostate cancer in men at risk. *Prostate Cancer and Prostatic Diseases*. 2020:1-16.
59. Greer MD, Shih JH, Lay N, Barrett T, Kayat Bittencourt L, Borofsky S, Kabakus IM, Law YM, Marko J, Shebel H. Validation of the dominant sequence paradigm and role of dynamic contrast-enhanced imaging in PI-RADS version 2. *Radiology*. 2017;285(3):859-69.
60. Ullrich T, Quentin M, Arsov C, Laqua N, Abrar D, Hiester A, Albers P, Antoch G, Schimmöller L, editors. Value of dynamic contrast-enhanced (DCE) MR imaging in peripheral lesions in PI-RADS-4 patients. *RöFo-Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*; 2020: © Georg Thieme Verlag KG.
61. Wells III WM. *Medical image analysis—past, present, and future*. Elsevier; 2016.
62. Gonzalez RC, Woods RE. *Digital image processing*. Prentice hall Upper Saddle River, NJ; 2002.
63. Lehmann T, Oberschelp W, Pelikan E, Reppes R. *Bildverarbeitung für die Medizin: Grundlagen, Modelle, Methoden, Anwendungen*: Springer-Verlag; 2013.

64. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. *Medical image analysis*. 2017;42:60-88.
65. Pohle R. Computerunterstützte bildanalyse zur auswertung medizinischer bilddaten. 2004.
66. Kooi T, Litjens G, Van Ginneken B, Gubern-Mérida A, Sánchez CI, Mann R, den Heeten A, Karssemeijer N. Large scale deep learning for computer aided detection of mammographic lesions. *Medical image analysis*. 2017;35:303-12.
67. Minaee S, Boykov YY, Porikli F, Plaza AJ, Kehtarnavaz N, Terzopoulos D. Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2021.
68. Ian Goodfellow YB, Aaron Courville. *Deep Learning*: MIT Press; 2016.
69. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-44.
70. Albawi S, Mohammed TA, Al-Zawi S, editors. Understanding of a convolutional neural network. 2017 International Conference on Engineering and Technology (ICET); 2017 21-23 Aug. 2017.
71. Fawzi A, Achuthan A, Belaton B. Brain Image Segmentation in Recent Years: A Narrative Review. *Brain Sci*. 2021;11(8).
72. Long J, Shelhamer E, Darrell T, editors. Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2015.
73. Ronneberger O, Fischer P, Brox T, editors. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*; 2015: Springer.
74. Nyúl LG, Udupa JK. On standardizing the MR image intensity scale. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*. 1999;42(6):1072-81.

75. Vovk U, Pernus F, Likar B. A review of methods for correction of intensity inhomogeneity in MRI. *IEEE transactions on medical imaging*. 2007;26(3):405-21.
76. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, Gee JC. N4ITK: improved N3 bias correction. *IEEE transactions on medical imaging*. 2010;29(6):1310-20.
77. He K, Zhang X, Ren S, Sun J, editors. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Proceedings of the IEEE international conference on computer vision*; 2015.
78. Cheng R, Roth HR, Lay NS, Lu L, Turkbey B, Gandler W, McCreedy ES, Pohida TJ, Pinto PA, Choyke PL. Automatic magnetic resonance prostate segmentation by deep learning with holistically nested networks. *Journal of medical imaging*. 2017;4(4):041302.
79. Klein S, Van Der Heide UA, Lips IM, Van Vulpen M, Staring M, Pluim JP. Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Medical physics*. 2008;35(4):1407-17.
80. Toth R, Madabhushi A. Multifeature landmark-free active appearance models: application to prostate MRI segmentation. *IEEE Transactions on Medical Imaging*. 2012;31(8):1638-50.
81. Makni N, Iancu A, Colot O, Puech P, Mordon S, Betrouni N. Zonal segmentation of prostate using multispectral magnetic resonance images. *Medical physics*. 2011;38(11):6093-105.
82. Chilali O, Puech P, Lakroum S, Diaf M, Mordon S, Betrouni N. Gland and zonal segmentation of prostate on T2W MR images. *Journal of digital imaging*. 2016;29(6):730-6.
83. Litjens G, Debats O, van de Ven W, Karssemeijer N, Huisman H, editors. A pattern recognition approach to zonal segmentation of the prostate on MRI. *International Conference on Medical Image Computing and Computer-Assisted Intervention*; 2012: Springer.

84. Qiu W, Yuan J, Ukwatta E, Sun Y, Rajchl M, Fenster A. Dual optimization based prostate zonal segmentation in 3D MR images. *Medical image analysis*. 2014;18(4):660-73.
85. Christ PF, Ettliger F, Grün F, Elshaera MEA, Lipkova J, Schlecht S, Ahmaddy F, Tatavarty S, Bickel M, Bilic P. Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks. *arXiv preprint arXiv:170205970*. 2017.
86. Li X, Chen H, Qi X, Dou Q, Fu C-W, Heng P-A. H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. *IEEE transactions on medical imaging*. 2018;37(12):2663-74.
87. Zhu Q, Du B, Turkbey B, Choyke PL, Yan P, editors. Deeply-supervised CNN for prostate segmentation. 2017 international joint conference on neural networks (IJCNN); 2017: IEEE.
88. Milletari F, Navab N, Ahmadi S-A, editors. V-net: Fully convolutional neural networks for volumetric medical image segmentation. 2016 fourth international conference on 3D vision (3DV); 2016: IEEE.
89. Zabihollahy F, Schieda N, Krishna Jeyaraj S, Ukwatta E. Automated segmentation of prostate zonal anatomy on T2-weighted (T2W) and apparent diffusion coefficient (ADC) map MR images using U-Nets. *Medical physics*. 2019;46(7):3078-90.
90. Rundo L, Han C, Zhang J, Hataya R, Nagano Y, Militello C, Ferretti C, Nobile MS, Tangherloni A, Gilardi MC. CNN-based prostate zonal segmentation on T2-weighted MR images: a cross-dataset study. *Neural Approaches to Dynamics of Signal Exchanges*: Springer; 2020. p. 269-80.
91. Rundo L, Han C, Nagano Y, Zhang J, Hataya R, Militello C, Tangherloni A, Nobile MS, Ferretti C, Besozzi D. USE-Net: Incorporating Squeeze-and-Excitation blocks into U-Net for prostate zonal segmentation of multi-institutional MRI datasets. *Neurocomputing*. 2019;365:31-43.

92. Meyer A, Mehrtash A, Rak M, Schindele D, Schostak M, Tempny C, Kapur T, Abolmaesumi P, Fedorov A, Hansen C, editors. Automatic high resolution segmentation of the prostate from multi-planar MRI. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018); 2018: IEEE.
93. Wang X, Yang W, Weinreb J, Han J, Li Q, Kong X, Yan Y, Ke Z, Luo B, Liu T. Searching for prostate cancer by fully automated magnetic resonance imaging classification: deep learning versus non-deep learning. *Scientific reports*. 2017;7(1):1-8.
94. Hosseinzadeh M, Saha A, Brand P, Slootweg I, de Rooij M, Huisman H. Deep learning–assisted prostate cancer detection on bi-parametric MRI: minimum training data size requirements and effect of prior knowledge. *European Radiology*. 2021:1-11.
95. Machireddy A, Meermeier N, Coakley F, Song X, editors. Malignancy detection in prostate multi-parametric MR images using U-net with attention. 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); 2020: IEEE.
96. De Vente C, Vos P, Hosseinzadeh M, Pluim J, Veta M. Deep learning regression for prostate cancer detection and grading in bi-parametric MRI. *IEEE Transactions on Biomedical Engineering*. 2020;68(2):374-83.
97. Sanford TH, Zhang L, Harmon SA, Sackett J, Yang D, Roth H, Xu Z, Kesani D, Mehralivand S, Baroni RH. Data augmentation and transfer learning to improve generalizability of an automated prostate segmentation model. *American Journal of Roentgenology*. 2020;215(6):1403-10.
98. Bardis M, Houshyar R, Chantaduly C, Tran-Harding K, Ushinsky A, Chahine C, Rupasinghe M, Chow D, Chang P. Segmentation of the prostate transition zone and peripheral zone on MR images with deep learning. *Radiology: Imaging Cancer*. 2021;3(3):e200024.
99. Makowski MR, Bressemer KK, Franz L, Kader A, Niehues SM, Keller S, Rueckert D, Adams LC. De Novo Radiomics Approach Using Image Augmentation and Features From T1 Mapping to Predict Gleason Scores in Prostate Cancer. *Investigative Radiology*. 2021;56(10):661-8.

Eidesstattliche Versicherung

„Ich, Günther Engel, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: KI-gestützte Segmentierung der Prostata, AI-based segmentation of the prostate, selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren/innen beruhen, sind als solche in korrekter Zitierung kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) werden von mir verantwortet.

Ich versichere ferner, dass ich die in Zusammenarbeit mit anderen Personen generierten Daten, Datenauswertungen und Schlussfolgerungen korrekt gekennzeichnet und meinen eigenen Beitrag sowie die Beiträge anderer Personen korrekt kenntlich gemacht habe (siehe Anteilserklärung). Texte oder Textteile, die gemeinsam mit anderen erstellt oder verwendet wurden, habe ich korrekt kenntlich gemacht.

Meine Anteile an etwaigen Publikationen zu dieser Dissertation entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Erstbetreuer/in, angegeben sind. Für sämtliche im Rahmen der Dissertation entstandenen Publikationen wurden die Richtlinien des ICMJE (International Committee of Medical Journal Editors; www.icmje.org) zur Autorenschaft eingehalten. Ich erkläre ferner, dass ich mich zur Einhaltung der Satzung der Charité – Universitätsmedizin Berlin zur Sicherung Guter Wissenschaftlicher Praxis verpflichte.

Weiterhin versichere ich, dass ich diese Dissertation weder in gleicher noch in ähnlicher Form bereits an einer anderen Fakultät eingereicht habe.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§§156, 161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

Datum

Unterschrift

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Publikationsliste

1. Assessment of the extracellular volume fraction for the grading of clear cell renal cell carcinoma: first results and histopathological findings.
Adams LC, Jurmeister P, Ralla B, Bressemer KK, Fahlenkamp UL, Engel G, Siepmann S, Wagner M, Hamm B, Busch J, Makowski MR.
Eur Radiol. 2019 Nov;29(11):5832-5843. doi: 10.1007/s00330-019-06087-x. Epub 2019 Mar 18.
2. Feasibility of gadoxetate disodium enhanced 3D T1 MR cholangiography (MRC) with a specific inversion recovery prepulse for the assessment of the hepatobiliary system.
Fahlenkamp UL, Adams LC, Böker SM, Engel G, Huynh Anh M, Wagner M, Hamm B, Makowski MR.
PLoS One. 2018 Sep 5;13(9):e0203476. doi: 10.1371/journal.pone.0203476. eCollection 2018.
3. Improved visualisation of hepatic metastases in gadoxetate disodium-enhanced MRI: Potential of contrast-optimised (phase-sensitive) inversion recovery imaging.
Fahlenkamp UL, Engel G, Adams LC, Böker SM, Huynh Anh M, Wagner M, Hamm B, Makowski MR.
PLoS One. 2019 Mar 6;14(3):e0213408. doi: 10.1371/journal.pone.0213408. eCollection 2019.
4. Use of quantitative T2 mapping for the assessment of renal cell carcinomas: first results.
Adams LC, Bressemer KK, Jurmeister P, Fahlenkamp UL, Ralla B, Engel G, Hamm B, Busch J, Makowski MR.
Cancer Imaging. 2019 Jun 7;19(1):35. doi: 10.1186/s40644-019-0222-8.
5. Assessing venous thrombus in renal cell carcinoma: preliminary results for unenhanced 3D-SSFP MRI.
Adams LC, Ralla B, Engel G, Diederichs G, Hamm B, Busch J, Fuller F, Makowski MR.

Clin Radiol. 2018 Aug;73(8):757.e9-757.e19. doi: 10.1016/j.crad.2018.04.003. Epub 2018 May 17.

6. Evaluation of osseous cervical foraminal stenosis in spinal radiculopathy using susceptibility-weighted magnetic resonance imaging.
Engel G, Bender YY, Adams LC, Boker SM, Fahlenkamp UL, Wagner M, Diederichs G, Hamm B, Makowski MR.
Eur Radiol. 2019 Apr;29(4):1855-1862. doi: 10.1007/s00330-018-5769-4. Epub 2018 Oct 15.
7. Native T1 mapping for assessment of the perilesional zone in metastases and benign lesions of the liver.
Fahlenkamp UL, Ziegeler K, Adams LC, Böker SM, Engel G, Makowski MR.
Sci Rep. 2020 Jul 30;10(1):12889. doi: 10.1038/s41598-020-69819-w.
8. Intracellular accumulation capacity of gadoxetate: initial results for a novel biomarker of liver function.
Fahlenkamp UL, Ziegeler K, Adams LC, Böker SM, Engel G, Makowski MR.
Sci Rep. 2020 Oct 22;10(1):18104. doi: 10.1038/s41598-020-75145-y.
9. Correlation of Native Liver Parenchyma T1 and T2 Relaxation Times and Liver Synthetic Function Tests: A Pilot Study.
Fahlenkamp UL, Kunkel J, Ziegeler K, Neumann K, Adams LC, Engel G, Böker SM, Makowski MR.
Diagnostics (Basel). 2021 Jun 20;11(6):1125. doi: 10.3390/diagnostics11061125.

Danksagung

Mein herzlicher Dank gilt folgenden Personen, die mich in der Durchführung meines Promotionsvorhabens maßgeblich unterstützt haben:

Zunächst meiner Erstbetreuerin Frau PD Dr. med. Lisa Adams für die Vermittlung des Promotionsthemas sowie die stetige Erreichbarkeit und Hilfsbereitschaft bei jeglichen Fragen und Problemen. Nach vielen Jahren der gemeinsamen klinischen Tätigkeit war es eine schöne Erfahrung, auch wissenschaftlich zusammen zu arbeiten und eine derart unkomplizierte und freundschaftliche Betreuung zu erfahren.

Des Weiteren möchte ich auch meinem Zweitbetreuer Herrn PD Dr. med. Keno Bressem danken für seine ausdauernde Unterstützung in technischen Fragen sowie für seine äußerst kompetenten und geduldigen Erklärungen einer für mich sehr komplexen und neuen Materie. Darüber hinaus bin ich auch großen Dank für seine konstruktiven Anmerkungen und schnellen Rückmeldungen schuldig.

Schlussendlich wäre die Anfertigung dieser Arbeit ohne die Unterstützung meiner Eltern nicht möglich gewesen. Sie haben meinen Weg bis hier her durch ihre Motivation, ihren Zuspruch und Rückhalt in jeglicher Lebenslage bereitet. Worte können meiner Dankbarkeit keinen Ausdruck verleihen.

Bestätigung eines akkreditierten Statistikers



CharitéCentrum für Human- und Gesundheitswissenschaften

Charité | Campus Charité Mitte | 10117 Berlin

Institut für Biometrie und klinische Epidemiologie (iBikE)

Direktor: Prof. Dr. Frank Konietschke

Name, Vorname: Engel, Günther
Emailadresse: guenther.engel@charite.de
Matrikelnummer: 4000545
PromotionsbetreuerIn: PD Dr. Keno Bressem,
PD Dr. Lisa Adams
Promotionsinstitution / Klinik: Klinik für Radiologie

Postanschrift:
Charitéplatz 1 | 10117 Berlin
Besucheranschrift:
Reinhardtstr. 58 | 10117 Berlin
Tel. +49 (0)30 450 562171
frank.konietschke@charite.de
<https://biometrie.charite.de/>



Bescheinigung

Hiermit bescheinige ich, dass Herr Günther Engel innerhalb der Service Unit Biometrie des Instituts für Biometrie und klinische Epidemiologie (iBikE) bei mir eine statistische Beratung zu einem Promotionsvorhaben wahrgenommen hat. Folgende Beratungstermine wurden wahrgenommen:

- Termin 1: 05.10.2021
- Termin 2: 14.04.2022

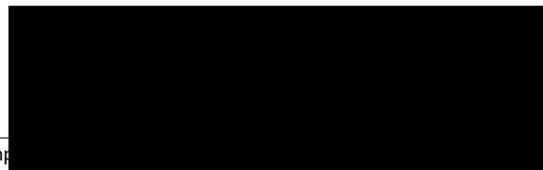
Folgende wesentliche Ratschläge hinsichtlich einer sinnvollen Auswertung und Interpretation der Daten wurden während der Beratung erteilt:

- deskriptive Auswertung der Daten
- da retrospektive Analyse, explorative Interpretation der Ergebnisse

Diese Bescheinigung garantiert nicht die richtige Umsetzung der in der Beratung gemachten Vorschläge, die korrekte Durchführung der empfohlenen statistischen Verfahren und die richtige Darstellung und Interpretation der Ergebnisse. Die Verantwortung hierfür obliegt allein dem Promovierenden. Das Institut für Biometrie und klinische Epidemiologie übernimmt hierfür keine Haftung.

Datum: 21.04.2022

Unterschrift BeraterIn, Institutsstempel



Campus Charité Mitte
Charitéplatz 1 | D-10117 Berlin
Sitz: Reinhardtstr. 58