

Aus der
Radiologischen Universitätsklinik Tübingen
Abteilung Diagnostische und Interventionelle Radiologie

**Künstliche Intelligenz in der Rekonstruktion von
Computertomographien – Einfluss von Deep Learning
Algorithmen auf die Bildqualität von Niedrigdosis-CT-
Untersuchungen bei Kindern und Jugendlichen**

**Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin**

**der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen**

**vorgelegt von
Schmid, Ulrich Gereon**

2025

Dekan: Professor Dr. B. Pichler

1. Berichterstatter: Privatdozent Dr. I. Tsiflikas

2. Berichterstatter: Professorin Dr. V. Ellerkamp

Tag der Disputation: 01.07.2025

Widmung

Für meine Eltern.

Inhalt

Abbildungs- und Tabellenverzeichnis	III
Abkürzungsverzeichnis	V
1 Einleitung	1
1.1 Computertomographie (CT) bei pädiatrischen Patienten	1
1.2 Besonderheiten in der Kinderradiologie	2
1.2.1 Besonderheiten und Strahlenschutz bei Kindern	2
1.2.2 Empfehlungen in der pädiatrischen Bildgebung	4
1.2.3 Indikationen für eine CT-Untersuchung in der Kinderradiologie	5
1.3 Technik der Computertomographie	6
1.3.1 Entwicklung CT	7
1.3.2 Aufbau CT und Aufnahmeprozess	9
1.3.3 Bildrekonstruktion	11
1.4 Protokollparameter und Bildqualität bei CT-Scans	15
1.4.1 Akquisitionsparameter	15
1.4.2 Rekonstruktionsparameter	17
1.4.3 Bildqualität	17
1.5 Dosimetrie	19
1.6 Dosisreduktion	22
1.6.1 Strahlenschutz Grundsätze	22
1.6.2 ALARA-Prinzip	22
1.6.3 Möglichkeiten von Strahlendosisreduktion	22
1.7 Künstliche Intelligenz bei Rekonstruktion von CT-Aufnahmen	25
1.7.1 Machine Learning, Artificial Intelligence, Deep Learning	25
1.7.2 Convolutional Neural Network	26
1.7.3 Anwendung künstlicher Intelligenz in der CT	27

1.8	Fragestellung der Arbeit.....	30
2	Material und Methoden	31
2.1	Patientenstudiendesign, Population und Strahlendosis	31
2.2	Bildaufnahme und Rekonstruktionsparameter	32
2.3	Objektive Bildqualität.....	33
2.4	Subjektive Bildqualität	33
2.5	Diagnostische Genauigkeit	34
2.6	Time to diagnosis	35
2.7	Phantomversuch	35
2.8	Statistische Auswertung.....	36
3	Ergebnisse.....	38
3.1	Patientenkohorte	38
3.2	Objektive Bildqualität.....	40
3.3	Subjektive Bildqualität	42
3.4	Diagnostische Genauigkeit	45
3.5	Time to diagnosis	60
3.6	Phantomversuch	62
4	Diskussion	64
4.1	Objektive Bildqualität.....	64
4.2	Subjektive Bildqualität	65
4.3	Diagnostische Genauigkeit	66
4.4	Time to diagnosis	67
4.5	Phantomversuch	68
4.6	Limitationen.....	69
4.7	Schlussfolgerung.....	69
5	Zusammenfassung	70
6	Literaturverzeichnis.....	72
7	Erklärung zum Eigenanteil	75
8	Veröffentlichungen.....	76
9	Danksagung.....	77

Abbildungs- und Tabellenverzeichnis

Abbildung 1: Darstellung eines Dual Source Computertomographen nach Siemens Healthineers: Die Geschichte der CT bei Siemens 14	10
Abbildung 2: Aufgliederung der Patientenkohorte 27	39
Abbildung 3: Bildrauschen der CT-Aufnahmen aus der Patientenstudie, separiert nach der Rekonstruktionsmethode 27	41
Abbildung 4: Verteilung der subjektiven gemittelten Bildqualität aufgeteilt in die einzelnen Rekonstruktionsmethoden 27	43
Abbildung 5: Visualisierung der diagnostischen Genauigkeit in Bezug auf peribronchiales Cuffing.27	48
Abbildung 6: Visualisierung der diagnostischen Genauigkeit in Bezug auf Mucus Plugging.27	50
Abbildung 7: Visualisierung der diagnostischen Genauigkeit in Bezug auf Milchglasinfiltrate.27	52
Abbildung 8: Visualisierung der diagnostischen Genauigkeit in Bezug auf Air Trapping.27	54
Abbildung 9: Visualisierung der diagnostischen Genauigkeit in Bezug auf Kavitäten.27	56
Abbildung 10: Visualisierung der diagnostischen Genauigkeit in Bezug auf Konsolidierungen.27	58
Abbildung 11: Time to diagnosis aufgetragen zur Berufserfahrung mit Aufteilung in die verschiedenen Rekonstruktionsmethoden 27	61

Abbildung 12: Darstellung des mittleren Bildrauschens der Phantomaufnahmen bei unterschiedlicher Strahlendosis und Rekonstruktionsmethode im Vergleich zu den 100 % KI-Aufnahmen	64
Tabelle 1: HU wichtiger menschlicher Organe abgeändert aus „Wie funktioniert CT?“ (S. 7) ¹³	14
Tabelle 2: Aufteilung der Studienkohorte nach Untersuchungsgrund bei Pneumonie	40
Tabelle 3: Interraterreliabilität bei Bewertung der CT-Aufnahmen unterschiedlicher Rekonstruktionsmethoden ²⁷	44
Tabelle 4: Interraterreliabilität bei Ermittlung der diagnostischen Genauigkeit anhand eines Severity Scores ²⁷	46
Tabelle 5: p-Werte des Severity Scores der einzelnen pathologischen Korrelate im Vergleich zu den KI-Aufnahmen	59
Tabelle 6: Ergebnisse der Mixed-Effekt Analyse bei Vergleich des Bildrauschens der verschiedenen Phantomschans mit unterschiedlicher Strahlendosis im Vergleich mit den 100 % KI-Aufnahmen	63

Abkürzungsverzeichnis

Abb.	Abbildung
AAPM	American Association of Physicists in Medicine
ADMIRE 2	Advanced Modeled Iterative Reconstruction strength 2
AEC	Automatic Exposure Control
ALARA	As Low As Reasonably Achievable
BfS	Bundesamt für Strahlenschutz
cm	Zentimeter
CNN	Convolutional Neural Network
CNR	Contrast-to-Noise ratio
CT	Computertomographie
CTA	CT-Angiographie
CTDI _{vol}	Computed Tomography Dose Index
DECT	Dual Energy Computertomographie
DL	Deep Learning
DLP	Dosislängenprodukt
DNS	Desoxyribonukleinsäure
DSCT	Dual Source Computertomographie
ED	Effektivdosis
et al.	et alii/aliae/alia
FBP	filtered back projection
GE	General Electric
GCS	Glasgow Coma Score
GNN	Generative Neural Network

Gy	Gray
HU	Hounsfield Unit
Hrsg.	Herausgeber
IQR	Interquartile Range
IR 2	iterative Rekonstruktion Stärke 2
Kernel	Faltungskern
KF	Knochenfenster
kg	Kilogramm
KI	künstliche Intelligenz
KM	Kontrastmittel
kV	Kilovolt
kW	Kilowatt
LF	Lungenfenster
mA	Milliampere
mAs	Milliamperesekunde
MBIR	model-based iterativ reconstruction
MDCT	Multidetektor-Computertomographie
mGy	Milligray
MIP	maximum intensity projection
mm	Millimeter
ML	Machine Learning
MPR	Multiplanare Rekonstruktion
MRT	Magnetresonanztomographie

ms	Millisekunde
mSv	Millisievert
PACS	Picture Archiving and Communication System
pCAP	pediatric community acquired pneumonia
ReLu	rectified linear units
RIS	Radiologisches Informationssystem
ROI	Region of interest
s	Sekunde
SD	standard deviation
SHT	Schädelhirntrauma
SNR	Signal-to-Noise ratio
SOP	Standard Operating Procedures
SSD	shaded surface display
SSDE	size specific dose estimate
SSK	Strahlenschutzkommission
Tab.	Tabelle
ULD-CT	Ultra Low Dose – Computertomographie
Vgl.	Vergleich
WF	Weichteilfenster
wFBP	weighted filtered back projection
ZNS	Zentrales Nervensystem
%	Prozent

Hinweis

Die in der Dissertation gewählte männliche Form bezieht sich immer zugleich auf weibliche, männliche und diverse Personen. Ausschließlich zum Zweck der besseren Lesbarkeit wird auf die geschlechtsspezifische Schreibweise sowie auf eine Mehrfachbezeichnung verzichtet. Alle personenbezogenen Bezeichnungen sind somit geschlechtsneutral zu verstehen.

1 Einleitung

1.1 Computertomographie (CT) bei pädiatrischen Patienten

Bildgebende Verfahren sind fester Bestandteil bei Bearbeitung klinischer Fragestellungen bei erwachsenen wie pädiatrischen Patienten. Es stehen zahlreiche unterschiedliche Bildgebungsverfahren zur Auswahl, die fallspezifisch angewendet werden können. Zu den klinisch am häufigsten angewendeten Verfahren gehören Röntgenaufnahmen, Sonographie, Computertomographie (CT) und Magnetresonanztomographie (MRT). Zur Anfertigung von CT-Aufnahmen und Röntgenbildern wird ionisierende Röntgenstrahlung benötigt. Sonographie- und MRT-Aufnahmen können ohne ionisierende Strahlung erstellt werden.

Schnittbildgebende Verfahren haben sich im klinisch-stationären als auch im ambulanten Sektor zunehmend etabliert und werden immer häufiger angewendet. Der Verlauf der Anzahl an durchgeführten CT-Untersuchungen in Deutschland in den letzten Jahren ist steigend. So hat die Anzahl an durchgeführten CT-Untersuchungen zwischen 2007 und 2016 um 45 % zugenommen.¹ Hohe Bildqualität, starke diagnostische Genauigkeit und gute Verfügbarkeit machen schnittbildgebende Untersuchungstechniken attraktiv. Der Bedarf an schnittbildgebenden Untersuchungstechniken besteht bei der Versorgung von erwachsenen Patienten sowie auch im pädiatrischen Bereich. In der Kinderradiologie wird sehr sorgfältig auf den Einsatz von Röntgenstrahlung geachtet, da Kinder als besonders strahlensensibel gelten. Genaueres folgt in 1.2 Besonderheiten in der Kinderradiologie. So gelten in der Kinderradiologie die konventionelle Radiographie und die Sonographie als Basisuntersuchungen. Trotzdem haben sich CT-Untersuchungen für bestimmte pädiatrische Fragestellungen etabliert, da die konventionelle Radiographie und die Sonographie bei bestimmten Fragestellungen an ihre Grenzen stoßen. Dies ist z.B. bei der Suche nach Lungenmetastasen oder Infiltraten der Lunge bei immunsupprimierten pädiatrischen Patienten der Fall. Weitere Beispiele zur

Indikation von CT-Untersuchungen im Kindesalter folgen in Abschnitt 1.2.3 Indikationen für eine CT-Untersuchung in der Kinderradiologie.

1.2 Besonderheiten in der Kinderradiologie

1.2.1 Besonderheiten und Strahlenschutz bei Kindern

Im Allgemeinen gilt es, die Strahlenbelastung für den Patienten zu minimieren. Jeder Einsatz von bildgebenden Verfahren, die mit Röntgenstrahlung arbeiten, benötigt eine Indikation, welche die Durchführung rechtfertigt. Dies gilt besonders bei pädiatrischen Patienten. Die Gründe zur besonderen Vorsicht sind vielfältig.

Kinder gelten als besonders strahlensensibel. Dies liegt u.a. daran, dass die Mitoserate in kindlichem Gewebe entwicklungsbedingt höher als die Mitoserate in Gewebe erwachsener Personen ist. Dies wird dadurch bedingt, dass die Mitosephase des Zellzyklus besonders anfällig für Strahlenschäden ist, da während dieser Phase die Teilung des Zellkerns und der Zelle stattfindet.² Säuglinge und Kinder, die sich in der Wachstumsphase befinden, besitzen mehr Zellen, die eine hohe Zellteilungsrate besitzen und somit vulnerabel sind, als dies bei erwachsenen Personen der Fall ist.³ Des Weiteren enthalten Gewebe von Kindern häufig einen höheren Wassergehalt. Dadurch wird Röntgenstrahlung in höherem Ausmaß absorbiert und gestreut. Folge ist, dass im Vergleich zu einer erwachsenen Person eine höhere Strahlendosis bei Kindern zur Durchdringung einer gleichdicken Gewebsschicht benötigt wird.³

Aufgrund der sich nach einer Untersuchung mit Röntgenstrahlung anschließenden voraussichtlich längeren Lebenszeit bei Kindern ist zudem die Wahrscheinlichkeit höher, dass sich Spätfolgen der Röntgenstrahlung manifestieren.⁴ Die durchschnittliche Latenz zwischen Bestrahlung und Auftreten eines Malignoms der Schilddrüse beträgt 10 - 40 Jahre.⁵ Eine 80 jährige Person ist folglich einem geringeren Risiko für Spätschäden ausgesetzt als ein Säugling.

Weiterhin können Untersuchungen mit ionisierender Strahlung das Erbgut von Kindern schädigen. Da Kinder als potenzielle Eltern angesehen werden können, können Schäden potenziell an Nachkommen weitergegeben werden. Die Höhe des Risikos ist jedoch wenig bekannt und schwer abzuschätzen, da es sich um

einen multifaktoriellen Vorgang handelt, bei dem viele andere teratogene Noxen ebenfalls Einfluss haben können.⁵

Kinder besitzen eine andere Anatomie als Erwachsene. Abweichend von der Anatomie eines Erwachsenen sind unter anderem die Körperproportionen, Größe und Verteilung von blutbildendem rotem Knochenmark im Körper. Genaueres folgt im nächsten Abschnitt.

Der kindliche Körper besitzt andere Proportionen als die eines Erwachsenen. Organe und andere anatomische Strukturen liegen näher nebeneinander. So kann es bei Körperstammuntersuchungen (Thorax, Abdomen) bei Säuglingen und Kleinkindern zur Strahlenbelastung von Organen kommen, die nicht Teil der Untersuchung sind. Dies basiert auf Effekten der Streustrahlung oder der Nicht-Ausblendbarkeit aus dem Strahlenfeld aufgrund anatomischer Gegebenheiten des Kindes. Als Beispiel ist folgender Zusammenhang zu nennen: Bei einer Röntgen-Thoraxaufnahme (z.B. Aufgrund einer Pneumonie) von einem pädiatrischen Patienten können strahlenempfindliche Organe, wie die Schilddrüse und die Ovarien, im strahlenexponierten Feld liegen. In der Regel bezieht sich die Fragestellung bei einer Röntgen-Thoraxaufnahme jedoch nicht auf die Schilddrüse oder Ovarien. Es werden in diesem Fall Schilddrüse und ggf. Ovarien mit Strahlung belastet, die nicht zur Beantwortung der Fragestellung abgebildet hätten werden müssen.⁵

Des Weiteren besitzen Kinder eine andere Verteilung des blutbildenden roten Knochenmarks als Erwachsene. Kinder besitzen in den Knochen der Extremitäten und des Schädels anteilmäßig mehr rotes blutbildendes Knochenmark als Erwachsene. Dies führt, je nach Körperregion unvermeidbar, bei Röntgenaufnahmen zu einer höheren mittleren Knochenmarks-Dosis bei Kindern als bei Erwachsenen.⁴ Besonders tritt ein Unterschied bei pädiatrischen Aufnahmen des Schädels und der Extremitäten auf.⁴

Strahlenschäden sind grundsätzlich in stochastische und deterministische Strahlenschäden aufteilbar. Bei deterministischen Strahlenschäden existiert abhängig vom betroffenen Gewebe eine Schwellendosis, ab welcher es zur Organschädigung kommt. Knochenmark besitzt bei einzeitiger und kurzzeitiger

Bestrahlung eine Schwellendosis von 0,5 Sv. Im Vergleich werden bei einer CT-Thorax eines Erwachsenen durchschnittlich 5-7 mSv appliziert.² Deterministische Strahlenschäden sind im Rahmen der radiologischen Diagnostik somit in der Regel nicht zu erreichen. Von Interesse sind hauptsächlich stochastische Strahlenschäden.

Bei stochastischen Strahlenschäden handelt es sich um Strahlenschäden, die im Gegensatz zu deterministischen Strahlenschäden keine Schwellendosis besitzen. Stochastische Strahlenschäden treten zufällig nach einer Exposition mit ionisierender Strahlung auf. Die Eintrittswahrscheinlichkeit eines stochastischen Strahlenschadens hängt von der Dosis ab, die schwere des stochastischen Strahlenschadens ist jedoch dosisunabhängig.² Des Weiteren besteht ein Zusammenhang zwischen Zeitpunkt der Strahlenexposition und Manifestation von Spätfolgen. Kinder, die im Alter von zehn Jahren ionisierender Strahlung ausgesetzt waren, haben im Vergleich zu Patienten, die zwischen dem 30. und 40. Lebensjahr exponiert waren, ein drei- bis vierfach erhöhtes Lebenszeitrisiko stochastische Strahlenschäden zu erleiden.⁴

1.2.2 Empfehlungen in der pädiatrischen Bildgebung

Bei pädiatrischen Patienten sollte die Indikation für die Anwendung ionisierender Strahlung streng gestellt werden. Sofern möglich sollten bildgebende Verfahren ohne ionisierende Strahlung, wie Sonographie und MRT, zum Einsatz kommen. Die Basisuntersuchung im Kindesalter ist bei vielen Fragestellungen die Sonographie. Genannt seien hier die Diagnostik von ZNS, der Abdominalorgane, Nieren und ableitenden Harnwege.⁴ Seltener erfolgen auch Aufnahmen des Thorax.⁴ Sollte sich eine Fragestellung mittels Sonographie nicht beantworten lassen, so ist auch der Einsatz anderer Bildgebungsmethoden in Betracht zu ziehen. Dabei sollten als Alternative, sofern möglich, Verfahren den Vorzug erhalten, welche keine ionisierende Strahlung verwenden oder in geringem Ausmaß. Viele pädiatrische Fragestellungen lassen sich mit einer CT- wie auch MRT-Aufnahme beurteilen. In der ZNS-, Weichteil- und Skelettdiagnostik ist die MRT nahezu unersetzlich.⁴ Weitere Anwendungsbereiche, die oftmals mit CT-

Diagnostik beantwortet wurden, inzwischen aber auch durch MRT-Aufnahmen beantwortet werden können, ist z.B. die Quantifizierung der Lungenbelüftung und Lungenperfusion im Rahmen einer Erkrankung an zystischer Fibrose.⁶

Sofern medizinisch möglich sollte die MRT in der Pädiatrie bei gleicher Aussagekraft der CT vorgezogen werden. Auch wirtschaftliche Belange sind bei Kindern gegenüber Strahlenexposition von untergeordneter Rolle, wenn dadurch auf Strahlenanwendung verzichtet werden kann.⁴

1.2.3 Indikationen für eine CT-Untersuchung in der Kinderradiologie

Trotz zahlreicher Weiterentwicklungen bildgebender Verfahren gibt es Fragestellungen und Situationen, in denen die MRT oder Sonographie ungeeignet ist, um eine gegebene Fragestellung zu beantworten. In solchen Fällen muss auf andere Verfahren mit ionisierender Strahlung zurückgegriffen werden. Dazu zählen CT und Röntgenaufnahmen. Diese können gewisse Fragestellungen, z.B. Ausschluss von Frakturen, beantworten, die durch Sonographie oder MRT nicht oder nicht eindeutig beantwortet werden können. Stößt eine Bildgebungsmethode ohne ionisierende Strahlung an ihre Grenzen, so muss abgewogen werden, ob für die Verwendung von ionisierender Strahlung zur Bildgebung eine Indikation besteht. Je nachdem welcher Schluss gezogen wird, muss entweder auf bildgebende Diagnostik verzichtet werden oder die Strahlenexposition in Kauf genommen werden.

Die Strahlenschutzkommission (SSK) hat 2019 eine überarbeitete Fassung ihrer Orientierungshilfe für bildgebende Verfahren herausgebracht. Je nach klinischer Fragestellung werden unterschiedliche Bildgebungsverfahren empfohlen. Dabei wird zwischen primärer und weiterführender Untersuchung unterschieden.

Es gibt diverse Indikationen für CT-Untersuchungen im Kindesalter. Hauptsächlich handelt es sich bei den Indikationen um weiterführende Untersuchungen oder Notfälle, in der die CT zum Einsatz kommt. Ein paar Fälle sollen hier genannt bzw. besprochen werden.

Ein akut vorliegendes Schädelhirntrauma (SHT) mit neurologischer Symptomatik stellt eine relative bzw. absolute Indikation für eine CT-Untersuchung dar. Ob es

sich um eine relative oder absolute Indikation handelt, ist in diesem Fall vom pädiatrischen Glasgow Coma Score (GCS) abhängig.³ Befindet sich der Patient in einem stabilem Zustand bei akutem SHT und besitzt neurologische Defizite, ist bei Verfügbarkeit über eine MRT nachzudenken.³

Bei einer pädiatrisch ambulant erworbenen Pneumonie (pCAP) wird die Diagnose primär klinisch gestellt. Bildgebende Verfahren, wie eine Röntgen-Thorax-Aufnahme, werden individuell und erst bei persistierender, unklarer Symptomatik angewandt. Schnittbildverfahren wie die CT können hinzugezogen werden, um parenchymatöse, hiläre und mediastinale Prozesse in hoher Auflösung darzustellen und zu differenzieren.⁷

Festzuhalten ist, dass die Indikation für eine CT-Untersuchung im Kindesalter sehr streng gestellt wird. Eine CT-Untersuchung ist oftmals nicht Mittel der ersten Wahl. Vielmehr dient die CT-Untersuchung im Kindesalter dazu, in speziellen unklaren Fällen die Diagnostik zu unterstützen. Klinische Diagnosefindung ohne ionisierende Strahlung hat häufig Vorrang.

Trotz alledem besitzt die CT je nach Fragestellung gegenüber anderen Bildgebungsmethoden entscheidende Vorteile. Dazu zählen u.a. die hohe räumliche Auflösung, die gute Verfügbarkeit, eine hohe diagnostische Genauigkeit sowie die kostengünstige und zeitlich schnelle Durchführbarkeit. In manchen Fragestellungen und Situationen überwiegen die Vorteile der CT, sodass die Risiken der Strahlung in Kauf genommen werden. Aufgrund dessen gilt es, die Dosis bei CT-Untersuchungen zu reduzieren. Genauer folgt in Punkt 1.6 Dosisreduktion.

1.3 Technik der Computertomographie

Im Folgenden soll die Entwicklung der Computertomographie und der grobe Aufbau eines Computertomographen beschrieben werden.

1.3.1 Entwicklung CT

Im Jahr 1972 wurde der erste kommerzielle medizinische CT-Scanner vom englischen Ingenieur Hounsfield gebaut. Das Gerät besaß verglichen mit dem heutigen technischen Stand eine niedrige Bildauflösung (80 x 80 Pixel) und lange Scanzeiten (300 s pro Schicht). Geeignet war dieser CT-Scanner ausschließlich für Aufnahmen des Kopfes.⁸

Bis Ende der 1980er Jahre gab es keine dramatischen Veränderungen der CT-Scanner. Lediglich die erhebbare Schichtdicke sank von 13 mm auf 1 mm. Ebenso wurde die Scandauer gesenkt sowie die Auflösung erhöht.⁸

Anfang der 1990er Jahre wurde die Entwicklung von Spiral-CT-Scannern u.a. durch Willi Kalender vorangetrieben. Grundlage der Entwicklung war eine kontinuierlich rotierende Röntgenröhre und Detektoreinheit. Der dadurch ermöglichte kontinuierlich erfolgende Tischvorschub bringt mehrere Vorteile mit sich.

Die Vorteile der Spiral-CT sind sehr vielfältig und so haben sich die klinischen Möglichkeiten der CT in vielerlei Hinsicht stark erweitert. Die Spiral-CT ist Grundlage zahlreicher inzwischen etablierter klinischer Untersuchungstechniken. Dazu gehören u.a. die CT-Angiographie (CTA), maximum intensity projection (MIP) und surface shaded displays (SSD). Ebenso zu erwähnen ist, dass durch die spiralförmige Aufnahmetechnik eine Überlappung der einzelnen aufgenommenen Schichten entsteht. Dadurch lassen sich Partialvolumeneffekte minimieren. Des Weiteren wird dadurch eine beliebige Schnittführung für Betrachtungsebenen ermöglicht. Aufnahmen können dadurch nicht nur in axialer Schnittebene begutachtet werden, sondern in beliebigen Varianten. Man spricht von multiplanarer Reformatierung (MPR).⁸ MPR ermöglicht beispielsweise bei einer CTA des Herzens das Erstellen einer Betrachtungsebene, die den Vierkammerblick ermöglicht.

1998 erfolgte die Markteinführung von Mehrzeilen Detektor Computertomographen (MDCT) durch die größeren CT-Gerätehersteller. Neuartig war bei diesen CT-Scannern, dass der Detektor nicht mehr nur aus einer Zeile besteht, sondern mehrere Detektorreihen besitzt. Dies ermöglicht die

Abdeckung von mehr Scanvolumen pro Umdrehung der Röntgenröhren-Detektoreinheit (Gantryeinheit). Infolgedessen wurden weitere Untersuchungsmethoden ermöglicht, da sich unter anderem die Scanzeit verkürzte. So folgte Anfang der 2000er Jahre die Erstellung der ersten EKG getriggerten Herz-CT Untersuchung.⁸

In den darauffolgenden Jahren wurden weitere MDCT-Scanner entwickelt, welche mehr Detektorzeilen besitzen. Erst wurden CT-Scanner mit 64 Zeilen, später mit 256 oder 320 Detektorzeilen entwickelt. MDCT-Scanner sind, auch aufgrund ihrer hohen Leistungsfähigkeit, bis zum heutigen Tag zahlreich im klinischen Einsatz. Ein konkretes Beispiel der Leistungsfähigkeit eines MDCT: Bei einer Kollimation von 256 x 0,625 mm deckt der Detektor 16 cm ab. So ist es möglich mit nur einer Umdrehung der Gantryeinheit das gesamte Herz aufzunehmen. Artefakte wie Stufenbildung können so bei Herz-CT Aufnahmen weitgehend minimiert werden.⁸ Ebenso ermöglichen MDCT-Scanner und Ultra Low Dose CT Protokolle (ULD-CT) im pädiatrischen Anwendungsbereich eine Thoraxuntersuchung in unter einer Sekunde.⁹

Mitte der 2000er Jahre erfolgte die Einführung von Dual Source CT (DSCT) Geräten. Besonderheit dieser CT-Scanner ist die Verwendung von zwei, oft im 90 Grad Winkel, versetzt angeordneten Röntgenröhren. Beide Röntgenröhren können simultan arbeiten. Dabei kann je nach Fragestellung und Aufnahmeprotokoll die gleiche oder unterschiedliche Beschleunigungsspannungen für die Röntgenröhren gewählt werden. So kann entweder bei Verwendung beider Röntgenröhren unter gleicher Beschleunigungsspannung die Scanzeit noch weiter gesenkt werden. Durch die verkürzte Aufnahmezeit entsteht die Möglichkeit von CT-Untersuchungen des Thorax bei pädiatrischen Patienten ohne Atemstopp bei erhaltener Bildqualität.¹⁰ Alternativ können die Röntgenröhren mit unterschiedlicher Beschleunigungsspannung, auch Dual Energy CT (DECT) genannt, betrieben werden. Man erhält bei DECT-Aufnahmen zwei unterschiedliche Datensätze von einem Patienten in einer Untersuchung. Dies ermöglicht es, mehr Informationen von dem untersuchten Patienten zu gewinnen. Beispielsweise besteht die Möglichkeit, bei einer CTA-Untersuchung mit iodhaltigem Kontrastmittel eine

Röntgenröhre im Niederspannungsbereich bei ca. 60-70 kV, was dem Absorptionsmaximum von Iod entspricht, zu betreiben und die zweite Röntgenröhre mit höherer Spannung anzusteuern (z.B. 140 kV). Je nach Beschleunigungsspannung wird die abgegebene Röntgenstrahlung von unterschiedlichen Geweben bzw. Materialien unterschiedlich stark abgeschwächt. Es handelt sich dabei um den Effekt der materialabhängigen charakteristischen Schwächung von Röntgenstrahlung. Bei Verwendung unterschiedlicher Röhrensparnungen (Dual Energy CT-Aufnahme) kann so die spezifische Absorption einzelner Materialien im Bild dargestellt werden. Der Kontrast von iodhaltigem Kontrastmittel zu umliegenden Gewebe, wie z.B. einer Gefäßwand, kann deutlicher hervorgehoben werden.¹¹ DECT-Aufnahmen werden gerne zur Charakterisierung von Gewebe oder zur Calcium-Quantifizierung genutzt.⁸

Die letzte größere Änderung bezüglich der Bauart von CT-Scannern erfolgte im Jahr 2021 und betrifft vor allem die Detektoreinheit. Diese Photon Counting CT-Scanner befinden sich inzwischen in klinischer Nutzung. Die Besonderheit ist, dass die verwendeten Detektoren direkt einzelne Photonen erfassen und in ein elektrisches Signal umwandeln können. Bisher arbeiteten die meisten CT-Scanner mit solid-state scintillation Detektoren. Diese wandeln Röntgenstrahlung in einem zweistufigen Prozess erst in sichtbares Licht um, welches anschließend in ein elektronisches Signal umgewandelt und ausgewertet wird.¹² Welche Auswirkungen der Einsatz von Photon Counting CT-Scannern im klinischen Alltag mit sich bringt, ist nicht abschließend geklärt und Bestandteil aktueller Forschung.

1.3.2 Aufbau CT und Aufnahmeprozess

Wie in 1.3.1 Entwicklung CT beschrieben gibt es unterschiedliche Bauformen von CT-Scannern. Hier soll der grobe Aufbau eines DSCT-Scanners nach Bauart der dritten Generation beschrieben werden.

Moderne Dual Source CT-Scanner nach Bauart der dritten Generation bestehen aus einem beweglichen Patiententisch sowie aus einer rotierenden Gantry. Die

Gantryeinheit enthält bei einem DSCT je zwei Röntgenröhren mit zugehörigen Detektoreinheiten, einen Generator und röhren- sowie detektorseitige Blenden. Der Generator dient der Stromversorgung der beiden Röntgenstrahler. Moderne CT-Scanner entfalten Leistungsspitzen von bis zu 100 kW, die abgedeckt werden müssen. Die röhrenseitige Blende dient zur Kollimierung der Schichtdicke des Röntgenfächers. So kann die Ausbreitung des Röntgenfächers in Patientenlängsrichtung begrenzt werden. Röntgenröhren und Detektoreinheiten rotieren mit konstanter Geschwindigkeit um den Patienten herum, während der Patiententisch mit kontinuierlichem Vorschub (Prinzip Spiral-CT) durch die Gantry hindurchbewegt wird. Dadurch wird eine kontinuierliche Aufnahme von Messwerten ermöglicht.¹³

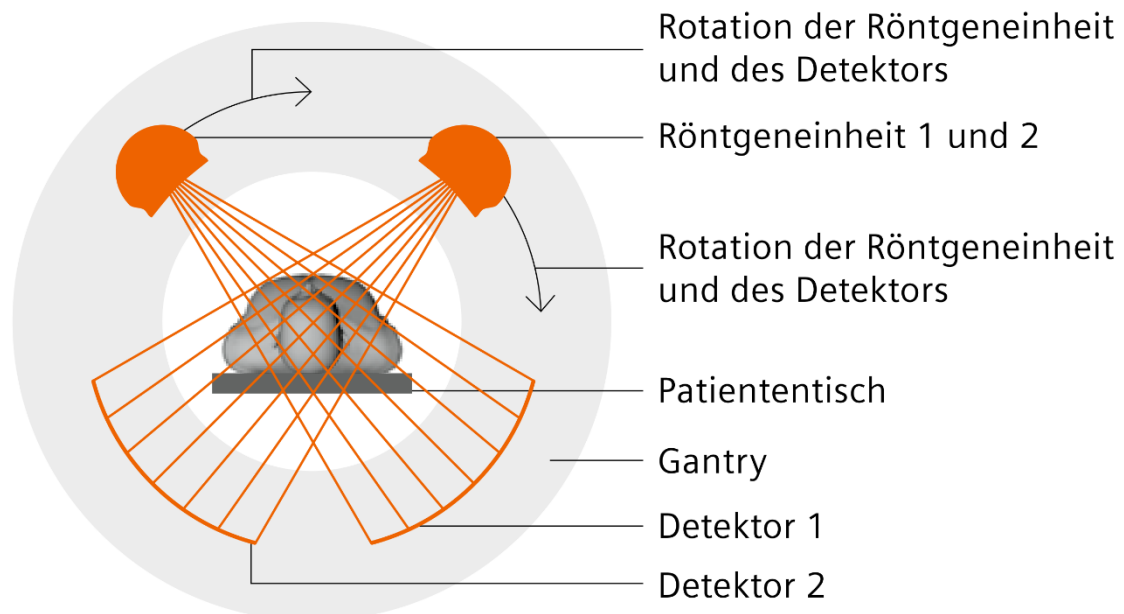


Abbildung 1: Darstellung eines Dual Source Computertomographen nach Siemens Healthineers: Die Geschichte der CT bei Siemens¹⁴

Dies ermöglicht kurze Untersuchungszeiten und daraus folgend eine hohe zeitliche Auflösung, wodurch Aufnahmen von beweglichen Organen wie bspw. des Herzens möglich werden.⁸

Die abgegebene Röntgenstrahlung wird durch den Patientenkörper abgeschwächt. Je nach Organ und somit Gewebe- bzw. Materialtyp wird die

Röntgenstrahlung unterschiedlich stark durch den Patientenkörper abgeschwächt. Die Detektoreinheiten messen die Intensität der Röntgenstrahlung, die nach Durchdringen des Patienten noch vorhanden ist. Die gemessene Intensität wird in ein elektronisches Signal umgewandelt. Dies geschieht heutzutage bei den meisten CT-Scannern durch einen Festkörper-Szintillationsdetektor. Der Szintillator wandelt die den Patienten durchdrungene Röntgenstrahlung in sichtbares Licht um, welches anschließend von einer Silizium-Photodiode erfasst und in ein elektronisches Signal umgewandelt wird.^{13,8}

Die Detektorelemente bestehen aus einem strahlensensiblen Material, häufig kommen Cadmium-Wolframat oder Gadolinium-Oxid zum Einsatz.¹³ Der Wirkungsgrad der CT-Detektoren entspricht ca. 80 %.¹³

Das elektronische Signal kann anschließend in mehreren Schritten in ein Schnittbild umgewandelt werden, welches zur Befundung geeignet ist. Dazu müssen mehrere Schritte durchlaufen werden. Es lässt sich aus der Differenz zwischen abgegebener Strahlenintensität und die durch den Patienten abgeschwächte gemessenen Strahlenintensität ein Schwächungswert errechnen, der im weiteren Verlauf in ein Bild umgewandelt wird.

Dazu genaueres in 1.3.3 Bildrekonstruktion.

1.3.3 Bildrekonstruktion

Die Bildrekonstruktion beschäftigt sich mit der Erstellung eines Bildes aus den erhobenen Schwächungsprofilen, die aus unterschiedlich gemessenen Projektionen erstellt wurden. Ziel ist es, den bestmöglichen Dreiklang aus hochauflösendem Bild, niedriger Strahlendosis und schneller Bildrekonstruktion zu erreichen.¹⁵

Ein Verfahren zur CT-Bildrekonstruktion ist die gefilterte Rückprojektion der einzelnen Schwächungsprofile. Die Messwerte der Schwächungsprofile werden entlang des ursprünglichen Strahlengangs ins Bild übertragen. Jeder Messstrahl, der durch den Patienten verläuft, entspricht einer fiktiven Linie im späteren CT-

Bild. Jede erdenkbare Linie mit unterschiedlicher räumlicher Ausrichtung kreuzt Pixel des späteren Bildraums. Alle Pixel, die durch einen Strahlengang bzw. die gedachte Linie gekreuzt werden, werden von dem Messwert dieses Strahlengangs beeinflusst. Die CT-Werte der Pixel, die von einem Strahlengang gekreuzt werden, erhalten den Schwächungswert dieses bestimmten Strahlengangs. Dieser Vorgang wird mit allen erhobenen Projektionen durchgeführt. So entsteht Schritt für Schritt eine Werte-Matrix, die sich als Bild darstellen lässt. Wird dieses Verfahren für jede aufgenommene Projektion bzw. Schicht durchgeführt, erhält man ein Abbild des Messkörpers.¹³

Das Abbild des Messobjekts, welches durch diese Verfahren entsteht, ist allerdings unscharf und verwaschen. Behoben werden kann dieser Effekt durch Verwendung eines Hochpassfilters vor Durchführung der Rückprojektion. Mathematisch lassen sich solche Verfahren als Faltungskern beschreiben. Das Ergebnis nach Faltung ist dann ein abbildungstreu und scharf dargestelltes Bild.¹³

Es gibt unterschiedliche Faltungskerne (Kernel). Man kann vereinfacht zwischen scharfen und weichen Faltungskernen unterscheiden.

Scharfe Faltungskerne (Kernel) bieten sich zur Darstellung von Knochen oder anderen Organen, welche hohe Kontrastunterschiede besitzen, an. Dazu zählt z.B. auch die Lunge. Scharfe Faltungskerne produzieren tendenziell Bilder mit scharfen Kanten und Abgrenzungen, produzieren jedoch ein höheres Bildrauschen.¹³

Die Verwendung von weichen Faltungskernen resultiert in CT-Bildern mit geringerer Bildschärfe, die zwangsläufig zu einer schlechteren Abgrenzbarkeit einzelner Strukturen führt. Es können kleine Details im Bild verloren gehen, jedoch ist das Niveau des Bildrauschens erniedrigt. Somit sind weiche Faltungskerne besser dazu geeignet, großflächige Strukturen mit kleinen Dichteunterschieden gut darzustellen. Anwendung finden die weichen Faltungskerne z.B. bei der Befundung von Leberläsionen.¹³

Weitere Bildrekonstruktionsmethoden existieren, sind allerdings vom Aufbau komplexer.

Iterative Rekonstruktionsalgorithmen haben einige Besonderheiten im Vergleich zur gefilterten Rückprojektion. Die CT-Bilder werden nicht analytisch, wie bei der gefilterten Rückprojektion, errechnet, sondern es werden einige Korrekturschritte in der Bildrekonstruktion eingefügt. Bei der Bildrekonstruktion werden nicht vorhandene Projektionsdaten vom Algorithmus künstlich erzeugt. Die künstlich erzeugten Daten werden anschließend wiederholend (= iterativ) mit den tatsächlich gemessenen Rohdaten verglichen und schrittweise an das potenziell der Wirklichkeit genauer entsprechende Bild angepasst. So entsteht ein Korrekturbild, welches in weiteren Durchgängen des Algorithmus weiterbearbeitet werden kann und zur Anpassung des ursprünglichen Bildes dient. Dazu lassen iterative Rekonstruktionsalgorithmen zahlreiche weitere Annahmen und Information in die Bilderrechnung mit einfließen. Bspw. werden Effekte wie Poissonsches Schrotrauschen mitbeachtet oder das Fehlen von Bilddaten aus bestimmten Scanbereichen toleriert bzw. ausgeglichen. So können Bildrauschen und die örtliche Auflösung moduliert werden. Man beachte, dass bei CT-Aufnahmen, welche iterativ rekonstruiert wurden, Bildrauschen und örtliche Auflösung nicht mehr zwangsläufig zusammenhängen. Bildrauschen kann durch iterative Rekonstruktionsverfahren gesenkt werden. Ebenso lassen sich Bildartefakte bis zu einem bestimmten Grad minimieren. Das Verfahren ist jedoch sehr rechenintensiv, wodurch längere Rekonstruktionszeiten als bei der gefilterten Rückprojektion entstehen.^{13,16}

Neuerdings werden auch Rekonstruktionsmethoden für CT-Aufnahmen entwickelt, welche auf künstlicher Intelligenz basieren. Dazu mehr in 1.7 Künstliche Intelligenz bei Rekonstruktion von CT-Aufnahmen.

In der CT-Bildgebung besteht ein Zusammenhang zwischen Bildschärfe und Bildrauschen. Bei gleichbleibender Strahlendosis und gleichbleibendem Bildrekonstruktionsalgorithmus ist eine Erhöhung der Bildschärfe in der Regel nur mit einem erhöhten Bildrauschen möglich.¹³

Als nächster Schritt im Weg von Intensitätsmessung zum fertigen CT-Bild gilt es die ermittelten und bearbeiteten Röntgen-Absorptions-Koeffizienten in CT-Werte umzurechnen.

Dazu gilt folgender Zusammenhang:

$$\text{CT-Werte} = 1000 * \frac{\mu - \mu_w}{\mu_w}$$

μ_w entspricht dabei dem Röntgen-Absorptionskoeffizienten von Wasser und beträgt 0,192/cm.

CT-Werte werden in der Einheit Hounsfield Unit (HU) angegeben.

Wasser erhält somit immer einen CT-Wert von 0 HU. Folgend in einer Tabelle die HU-Werte wichtiger Organe.

Gewebe	HU
Luft	-1000
Lunge	- 700 bis -950
Fett	- 50
Wasser	0
Muskulatur	50
Blut	50 bis 65
Knochen	bis zu 2000

Tabelle 1: HU wichtiger menschlicher Organe abgeändert aus „Wie funktioniert CT?“ (S. 7)¹³

Die ermittelten CT-Werte dienen zur anschließenden Darstellung als Grauwerte. Bei einem CT-Bild erstreckt sich die CT-Werteskala von -1024 bis 3071. Es ergeben sich insgesamt 4096 Abstufungen. Potenziell könnten 4096 unterschiedliche Graustufen abgebildet werden. Das menschliche Auge kann jedoch nur ca. 60 Graustufen unterscheiden. Werden CT-Bilder betrachtet, kommt in der Regel eine Fensterung zum Einsatz, die für jede Fragestellung bzw. Gewebe frei gewählt werden kann. Jedes gewählte Fenster besitzt eine Spanne von CT-Werten, innerhalb der die Graustufenzuteilung erfolgt. Dadurch entsteht folgender Effekt, dass alle CT-Werte unterhalb der Fensterung schwarz, alle CT-

Werte oberhalb der Fensterung weiß dargestellt werden. Innerhalb der Fensterung herrscht ein gleichmäßig verteilter Verlauf an Graustufen von schwarz zu weiß.¹³

1.4 Protokollparameter und Bildqualität bei CT-Scans

Grundsätzlich können verschiedene Parameter bei der Erstellung von CT-Aufnahmen variiert werden. Zum einen können Parameter der Bildakquisition beeinflusst werden, zum anderen kann im Anschluss an das Aufnahmeverfahren die Rekonstruktion sowie Nachbearbeitung der gemessenen Daten unterschiedlich erfolgen. Beides soll hier besprochen werden.

1.4.1 Akquisitionsparameter

Folgende Akquisitionsparameter können mehr oder weniger frei gewählt werden: Kollimation, Pitch, Volumenpitch, Rotationszeit, Röhrenstrom und Röhrenspannung.

Kollimation

CT-Scanner verfügen über Kollimatoren. Ein Kollimator ist eine Blende, mit der ein Strahlenfächer aus der vom Röntgenstrahler abgegebenen Strahlung geformt werden kann. Variiert wird die Kollimation. Unter Kollimation versteht man die Kollimatoröffnung. Bei handelsüblichen CT-Scannern kommen zwei Kollimatoren zum Einsatz. Ein Kollimator befindet sich zwischen Röntgenstrahler und Patient, der andere Kollimator ist zwischen Patient und Detektor lokalisiert. Die Aufgabe des röntgenstrahlernahen Kollimators ist die Kontrolle und Formung des Strahlenfächers. Die Längsausdehnung des Strahlenfächers kann so angepasst und begrenzt werden. Ziel der Verwendung von Kollimatoren ist, den Patienten vor nicht bildgebender Strahlung zu schützen. Der detektornahe Kollimator dient zur Abschirmung des Detektors vor Streustrahlung. Je nach Wahl der Kollimation wird die örtliche Auflösung der Aufnahme beeinflusst.¹³

Pitch

Der Pitch-Faktor p ist dimensionslos und wird bei der Spiral-CT verwendet. Beschrieben wird der kontinuierliche Tischvorschub während einer vollständigen Gantryumdrehung. Mit beeinflusst wird der Pitch-Faktor durch die Detektorbreite bei Mehrzeilen-CT. Dabei zählen jedoch nur aktivierte Detektorzeilen.¹³

Volumenpitch

Alternativ kann der Volumenpitch angegeben werden. Im Unterschied zum Pitch-Faktor zählen nur die gewählte Kollimation und der Tischvorschub. Die Anzahl aktivierter Detektorzeilen wird nicht mit eingerechnet.¹³ Je nach CT-Scanner-Hersteller wird der Pitch- oder Volumenpitchwert angegeben.

Rotationszeit

Die Rotationszeit ist die Zeit, die für eine Rotation der Gantry benötigt wird. Beeinflusst u.a. die Aufnahmedauer einer CT-Untersuchung. Wird in Millisekunden (ms) angegeben.

Röhrenstrom

Als Röhrenstrom wird der Strom bezeichnet, der zum Heizen durch die Kathode der Röntgenröhre fließt. Angegeben wird der Röhrenstrom in mA. Häufig erfolgt die Angabe des Röhrenstrom-Zeit-Produkts. Dabei wird die Rotationszeit mit dem Röhrenstrom multipliziert. Die Einheit ist mAs. Das Röhrenstrom-Zeit-Produkt dient als Maß für die Belichtung. Zur besseren Vergleichbarkeit wird häufig das effektive Röhrenstrom-Zeit-Produkt angegeben. Es normiert das Röhrenstrom-Zeit-Produkt und die Belichtungszeit für eine einzelne Schicht.¹³

Röhrenspannung

Die Röhrenspannung ist als die in der Röntgenröhre verwendete Spannung zwischen Kathode und Anode definiert. Je nach Variation kann das Spektrum der Röntgenstrahlung beeinflusst werden.¹³

1.4.2 Rekonstruktionsparameter

Nachdem eine CT-Untersuchung durchgeführt wurde, sind zahlreiche Messwerte erhoben worden. Die erhobenen Messwerte werden im Prozess der Bildrekonstruktion in ein Bild umgewandelt. Hierzu gibt es Verfahren und Möglichkeiten, den Prozess durch Wahl unterschiedlicher Parameter zu beeinflussen.

Variiert werden kann u.a. das **Rekonstruktionsinkrement**. Darunter wird der Abstand von zwei Schichten verstanden. Bei einer Spiral-CT ist das Rekonstruktionsinkrement frei wählbar und hat Einfluss auf den Überlappungsgrad von Schichten.¹³

Ebenso kann die **Schichtdicke** der einzelnen Bilder gewählt werden. Jedoch können nur Schichtdicken gewählt werden, die gleich groß oder größer als die verwendete Detektorkollimation sind.¹³

Rekonstruktionsalgorithmen

Die Wahl des Rekonstruktionsalgorithmus hat großen Einfluss auf die spätere Bildqualität von CT-Aufnahmen. Häufig kommt eine gefilterte Rückprojektion oder ein iterativer Rekonstruktionsalgorithmus zur Anwendung. Die gefilterte Rückprojektion und der grobe Aufbau einer iterativen Rekonstruktion wurde bereits in 1.3.3 Bildrekonstruktion besprochen.

1.4.3 Bildqualität

Die Bildqualität von CT-Aufnahmen kann unterschiedlich beurteilt werden. Es kann zwischen subjektiven und objektiven Bewertungskriterien unterschieden werden.

Bildrauschen einzelner Pixel kann als ein gutes Maß für die objektive Bildqualität herangezogen werden. Pixelrauschen in CT-Aufnahmen setzen sich zum einen aus Quantenrauschen, zum anderen aus elektronischem Rauschen zusammen. Quantenrauschen entsteht durch ein gewisses Bildrauschen der eigentlichen Messwerte. Elektronisches Rauschen hat seine Ursache im Detektorelement des

CT-Scanners und der anschließenden digitalen Weiterverarbeitung des elektronischen Signals. Dabei gilt anzunehmen, dass das Pixelrauschen einer Poisson-Verteilung folgt. Ausdrücken lässt sich das Bildrauschen als Standardabweichung der CT-Werte im CT-Bild. Pixelrauschen und Dosis verhalten sich proportional zueinander. Auf CT-Aufnahmen lässt sich der Zusammenhang wie folgt darstellen:¹³

$$\sigma = \frac{1}{\sqrt{ST \cdot N}}$$

Wobei σ dem Rauschen, ST der Schichtdicke und N der Anzahl der Photonen entspricht.

Daraus lässt sich ableiten, dass das Pixelrauschen reziprok proportional zur Quadratwurzel der Schichtdicke und Anzahl der Photonen zusammenhängt. Bildrauschen lässt sich demnach auf verschiedene Art und Weise mindern. Einige seien hier genannt:

- Steigerung des Röhrenstrom-Zeit-Produkts
- Wahl größerer Schichtdicken
- Variation der Bildrekonstruktion

Alle genannten Varianten können Bildrauschen reduzieren. Jedoch wird sich verringertes Bildrauschen in allen genannten Fällen durch andere, mehr oder weniger tolerierbare Nachteile erkauft. Diese sind u.a. erhöhte Strahlendosis, verringerte Ortsauflösung oder komplexere Bildrekonstruktionsmethoden.¹³

Der Kontrast in Bildern ist ein weiteres Qualitätsmerkmal. Er beschreibt den Helligkeitsverlauf zwischen zwei nebeneinander liegenden Bildpunkten. Besitzt ein Bild keinen Kontrast, so kann keine nennenswerte Bildinformationen entnommen werden. Auch hier besteht ein Zusammenhang zur applizierten Dosis. Kontrast und Dosis verhalten sich proportional zueinander. Ebenso hat auch das untersuchte Gewebe Einfluss auf den Kontrast eines Bildes. Ausgedrückt wird Kontrast als ein Verhältnis zum Pixelrauschen.¹³

Es gilt also, dass durch erhöhte Röhrenstrom-Zeit-Produkte und somit erhöhter Strahlendosis Bildrauschen minimiert und geringe Kontraste besser dargestellt werden können. Für klinische CT-Untersuchungen hat dies zur Folge, dass zur Darstellung von Gewebe mit hohem Kontrast (Lunge, Knochen, Kontrastmittel-Untersuchungen) geringere Dosen für eine solide Bildgebung notwendig sind als Gewebe mit niedrigem Kontrast (Abdomen, Gehirn).¹³

1.5 Dosimetrie

Durchdringt Röntgenstrahlung eine Materie so wird Energie an die bestrahlte Materie übertragen. Die übertragene Energie bezeichnet man auch als Strahlendosis. Zur Ermittlung und Angabe der Strahlendosis stehen in der Radiologie unterschiedliche Parameter zur Verfügung. Jeder Parameter gibt unterschiedliche Informationen über die Strahlenexposition einer Untersuchung.

Energiedosis

Die Energiedosis ist ein Maß der mittleren Energie, die bei einer Untersuchung von einem Patienten durch Absorption aufgenommen wird. Dabei ist entscheidend, wieviel Energie, angegeben in Joule, pro kg Materie absorbiert wird. Somit wird die Energiedosis in der Einheit Gray (Joule / kg) angegeben und es handelt sich um eine physikalische Messgröße.¹⁷

Äquivalentdosis

Ergänzend zur Energiedosis gibt es das Maß der Äquivalentdosis, sie bezieht die Art der verwendeten Strahlung mit ein und versucht so die unterschiedliche Wirkung verschiedener Strahlenarten mit ihren unterschiedlichen Energieabgabeverhalten in biologischem Gewebe bei Bestrahlung abzubilden. Bei Verwendung von Röntgenstrahlung entspricht die Äquivalenzdosis numerisch der Energiedosis. Die Einheit der Äquivalentdosis ist jedoch Sievert. Man kann festhalten, dass die Einheit Gray somit zur Beschreibung physikalisch gemessenen Strahlendosen genutzt wird und die Einheit Sievert zur Bewertung einer Strahlendosis auf ein biologisches System verwendet wird.¹⁷

Effektivdosis

Bei der Effektivdosis handelt es sich um eine Äquivalentdosis. Sie gilt als Maß zum Abschätzen des Strahlenrisikos des Patienten. Die Effektivdosis bezieht die unterschiedliche Vulnerabilität einzelner Organe in die Berechnung mit ein. Ebenso wird in die Ermittlung der effektiven Dosis die verwendete Strahlenart miteinbezogen. Die Ermittlung der effektiven Dosis wird anhand eines mathematischen Modells geschätzt. Angegeben wird die Effektivdosis in der Einheit Millisievert (mSv).^{13,8}

Zum Vergleich werden in der Literatur für eine CT-Thorax eines Erwachsenen ca. 5-7 mSv angegeben.¹³ Die natürliche jährliche effektive Dosis eines in Deutschland lebenden Menschen wird auf 2 bis 3 mSv geschätzt.¹ Ein Hin- und Rückflug von Frankfurt nach New York liegt bei ca. 0,14 mSv.¹⁸

Zur Beschreibung der Strahlendosis bei CT-Untersuchungen gibt es weitere Parameter, die die Besonderheiten der CT-Technik mit einbeziehen. Zu den wichtigsten zählen der Computed Tomography Dose Index, die Size-specific dose estimate und das Dosislängenprodukt

Computed Tomography Dose Index

Bei CT-Untersuchungen werden Patienten schichtweise von Röntgenstrahlung durchdrungen. Dies führt zu Bestrahlung, die sich nicht exakt auf eine definierbare Schicht verteilt. Effekte wie Streustrahlung haben ebenfalls einen nennenswerten Anteil an der abgegebenen Strahlendosis pro Schicht. Der $CTDI_{vol}$ beschreibt die von einer definierten Schichtdicke erhaltene Strahlendosis in der CT mit Einbezug der Streustrahlung, die auf Nachbarschichten gewirkt hat. Mit einbezogen werden häufig 50 mm auf- und abwärts der Schicht, die von Interesse ist. Es handelt sich bei Wahl einer 50 mm Pufferzone um den $CTDI^{100}$. Der $CTDI_{vol}$ ist eine standardisierte Berechnung der absorbierten Strahlung durch das Scanobjekt und ist nicht zwingend repräsentativ für die vom Patienten absorbierte Strahlendosis.

Es ist zu beachten, dass der $CTDI_{vol}$ je nach Körperbau des Patienten Ungenauigkeiten mit sich bringen kann. Ursache ist, dass die Grundlage zur Berechnung des $CTDI_{vol}$ auf Eichmessungen an einem Phantommesskörper basieren. Diese besitzen oftmals 16 cm bzw. 32 cm Durchmesser und werden als Vergleich für die spätere Berechnung des $CTDI_{vol}$ herangezogen. Der Phantommesskörper entspricht nur einer Näherung eines Standardpatienten. Weicht der Körperbau in Form oder Größe des Patienten vom Phantommesskörper nennenswert ab, so entspricht die tatsächlich erhaltene Strahlendosis des Patienten nicht zwangsläufig mit der Dosis überein, welche über den $CTDI_{vol}$ berechnet wurde.⁸

Der $CTDI_{vol}$ beschreibt eine Dosis, die ein bestimmtes Scanvolumen absorbiert, wenn das Scanvolumen einem homogenen Phantommesskörper entsprechen würde. Angegeben wird der $CTDI_{vol}$ in mGy. In der Spiral-CT hängen $CTDI_{vol}$ und Pitch reziprok proportional zusammen.¹³

Size-specific dose estimate

Size-specific dose estimate (SSDE) ist ein Parameter, der von der American Association of Physicists in Medicine (AAPM) eingeführt wurde. Er dient der Anpassung des $CTDI_{vol}$, welches auf einem 32 cm bzw. 16 cm Phantomkörper basiert, an die echte Patientengröße. Dazu gibt es Konversionsfaktoren, welche je nach lateraler bzw. anteroposteriorer Ausdehnung des Patientenkörper gewählt werden können und mit dem $CTDI_{vol}$ multipliziert werden. Der so gebildete SSDE-Parameter kann nun genutzt werden, um Dosisvergleiche zwischen unterschiedlich großen Patienten mit gleicher Untersuchung zu ziehen.⁸

Dosislängenprodukt

Das Dosislängenprodukt (DLP) berechnet sich aus dem $CTDI_{vol}$. Dazu wird das Integral entlang der Z-Achse (Längsachse des Patienten) gebildet. Die Einheit des DLP ist somit $mGy \cdot cm$. Das DLP beschreibt die summierte Energie einer gesamten CT-Untersuchung.

1.6 Dosisreduktion

1.6.1 Strahlenschutz Grundsätze

Bildgebende Diagnostik mit Röntgenstrahlung hat in der Radiologie einen bedeutenden Stellenwert und muss mit Bedacht sowie unter Beachtung korrekter Indikation verwendet werden. Der Nutzen einer Bildgebung mit Röntgenstrahlung sollte stets das potenzielle Krankheitsrisiko des Patienten überwiegen. Dies wurde bereits in Punkt 1.1 Computertomographie (CT) bei pädiatrischen Patienten erörtert.

Strahlenschutz besteht aus genauem Abwägen, ob eine Untersuchung überhaupt durchgeführt werden sollte und der Minimierung der Strahlenbelastung für den Patienten, wenn eine Röntgenuntersuchung durchgeführt wird.

1.6.2 ALARA-Prinzip

Das Prinzip As Low As Reasonably Achievable (ALARA) wurde in den 1980er von Hendee et al entwickelt.¹⁹ Es fordert einen umsichtigen Einsatz von Röntgenstrahlung. Dabei soll die Strahlendosis so weit gesenkt werden wie möglich, aber unter Erhalt einer Bildqualität, die zur Diagnosefindung noch ausreichend ist. Das ALARA-Prinzip hat sich mittlerweile zu einem europäischen Sicherheitsstandard etabliert und findet sich in nationalen Gesetzgebungen.^{13,19}

1.6.3 Möglichkeiten von Strahlendosisreduktion

Es gibt verschiedene Möglichkeiten einer Strahlendosisreduktion für Patienten. Die einfachste ist auf eine Bildgebung mit Röntgenstrahlung verzichten zu

können. Dies ist jedoch nicht immer möglich. Fällt die Wahl der Bildgebung auf eine CT-Untersuchung, so besitzen CT-Scanner verschiedene Möglichkeiten, Strahlendosis bei CT-Aufnahmen einzusparen und dabei eine ähnlich hohe Bildqualität zu produzieren. Wie sich bestimmte Akquisitionsparameter auf die Strahlendosis bei CT-Scans auswirken, wurde bereits in Punkt 1.4.1 Akquisitionsparameter angesprochen.

Ein paar gerätespezifische Möglichkeiten zur Strahlendosisreduktion sollen nun ergänzt werden.

Modulation des Röhrenstroms

Eine effektive Möglichkeit Dosis einzusparen, besteht in der Modulation des Röhrenstroms. Der Röhrenstrom bestimmt die Anzahl der entsendeten Photonen pro Zeit. Röhrenstrom und Dosis sind proportional miteinander verknüpft. Dazu siehe auch 1.4.1 Akquisitionsparameter.

Der Röhrenstrom kann manuell vor einer Untersuchung oder in Echtzeit während einer Untersuchung automatisch angepasst werden. Bei der automatischen Röhrenstrommodulation spricht man auch von einer automatischen Expositionskontrolle (Automatic Exposure Control = AEC).

Grundlage der AEC ist die Auswertung des vor einer CT-Untersuchung angefertigten Topogramms. Anhand des Topogramms werden für den Patienten spezifisch in X-, Y- und Z-Richtung ein schwächungsbasierter Röhrenstrom berechnet, der in der Untersuchung dann ortsspezifisch angewandt wird. Im Idealfall entsteht eine CT-Aufnahme mit gleichbleibendem Bildrauschen trotz unterschiedlichen Röhrenstroms. Je nach Hersteller und Produkt variieren die Ansätze der Röhrenstrommodulation leicht. Strahlendosisreduktionen zwischen 20 - 68 % wurden von Mulkens et al. 2005 beschrieben.²⁰ Schlanke Patienten profitieren mehr von der Röhrenstrommodulation als adipöse Patienten, bei denen weniger Dosis eingespart werden kann.^{13,8}

Wahl der Röhrenspannung

Durch geschickte Wahl der Röhrenspannung kann eine Dosisreduktion erreicht werden. Besonders bei CT-Untersuchungen von kleinen Volumina, also normal gewichtigen Patienten, kann von einer niedrigeren Röhrenspannung bei ähnlichem Bildrauschen profitiert werden. Bei schlanken Patienten kann durch eine Absenkung der Röhrenspannung bei gleichzeitiger Erhöhung des Röhrenstroms die Contrast-to-Noise ratio (CNR) gesteigert werden. Ebenso kann bei der Verwendung von iodhaltigem Kontrastmittel die erhöhte Absorptionsfähigkeit von Iod bei niedrigeren Röhrenspannungen ausgenutzt werden. Dies eignet sich besonders bei CT-Angiographien. Wie der Röhrenstrom kann auch die Röhrenspannung automatisch gewählt werden. Dazu wird das vor der CT-Untersuchung erstellte Topogramm herangezogen.⁸

Eine Strahlendosisreduktion von 14,7 % über alle CT-Scans kann durch Verwendung einer automatischen Röhrenspannungswahl erreicht werden.⁸

Spectral Shaping

Bei CT-Aufnahmen ohne Verwendung von Kontrastmittel tragen niederenergetische Photonen weniger zur Bildgebung bei als hochenergetische Photonen. Ziel des Spectral Shaping ist es, niederenergetische, der Bildgebung wenig zuträgliche Röntgenstrahlung mittels Zinn-Filter zu entfernen. Der Patient erhält eine geringere Strahlendosis.⁸

Genauer folgt in Punkt 4.1 Objektive Bildqualität.

Ebenso können die Parameter, die in 1.4.1 Akquisitionsparameter besprochen wurden, variiert werden und zu einer Senkung der Strahlendosis führen. Dies betrifft vor allem die Wahl des Rekonstruktionsalgorithmus und des Pitch-Faktors.

1.7 Künstliche Intelligenz bei Rekonstruktion von CT-Aufnahmen

Es gibt zahlreiche Methoden, aus den in einer CT-Untersuchung erhobenen Schwächungswerten ein CT-Bild zu erstellen. Herkömmliche Bildrekonstruktionsverfahren sind die gefilterte Rückprojektion (FBP) und iterative Rekonstruktionsalgorithmen, die bereits in 1.3.3 Bildrekonstruktion beschrieben wurden. Seit Einführung der CT in den klinischen Alltag haben CT-Scanner und Rekonstruktionsalgorithmen eine gewaltige Entwicklung durchlaufen. Der erste auf dem Markt erhältliche CT-Scanner von EMI verwendete noch einen iterativen Ansatz der Bildrekonstruktion. Aufgrund der rasant steigenden Auflösung der CT-Scanner und der damals begrenzt zur Verfügung stehenden Rechenleistung in den 1970-80er Jahren, war es schnell nicht mehr möglich, auf iterative Rekonstruktionsalgorithmen zurückzugreifen. FBP-Algorithmen setzten sich durch, bis in den 2000er Jahren iterative Rekonstruktionsalgorithmen in die CT-Bildrekonstruktion zurückkehrten.²¹

1.7.1 Machine Learning, Artificial Intelligence, Deep Learning

Zunächst soll ein kurzer Einschub zur Thematik Künstlicher Intelligenz (KI) und neuronalen Netzen folgen. Häufig werden im Zusammenhang mit künstlicher Intelligenz die Begriffe Machine Learning (ML) und Deep Learning (DL) verwendet. Es handelt sich dabei nicht um Synonyme, sondern um leicht unterschiedliche Verfahren bzw. Untergruppen von künstlicher Intelligenz.

Unter künstlicher Intelligenz werden Computeralgorithmen verstanden, welche im Stande sind Aufgaben zu lösen für die normalerweise menschliche Intelligenz notwendig wäre.²²

Machine Learning ist eine Untergruppe von künstlicher Intelligenz. Es handelt sich dabei um Algorithmen, die ohne vorgegebene Regeln von zur Verfügung gestellten Daten lernen können, um ein Problem zu lösen.²²

Als Deep Learning wird eine weitere Untergruppe von Machine Learning verstanden. Besonderheit ist hier, dass die Methode mit neuronalen Netzen mit unterschiedlichen Verschaltungsebenen arbeitet. In der Bildverarbeitung

kommen hier gerne Convolutional Neural Networks (CNN) oder Generative Neural Networks (GNN) zum Einsatz.²²

1.7.2 Convolutional Neural Network

Convolutional Neural Networks (CNN) werden häufig zur Bildverarbeitung eingesetzt. CNN kombinieren durch ihren Aufbau eine gewisse Verschiebungs-, Skalierungs- und Verzerrungsinvarianz, die es dem neuronalen Netz ermöglicht sich an die zu lösende Aufgabe anzupassen.²³

Ein CNN besteht aus mehreren Ebenen. Jede Ebene hat unterschiedliche Aufgaben. Es müssen Aufgaben der Datenaufnahme, Aktivierung bzw. Filterung, Datenverarbeitung und Output des Netzwerks abgedeckt werden.²³

Erste Ebene der Datenverarbeitung ist ein sogenannter Convolutional-Layer (Faltungskern), welche rezeptive Felder und feste Gewichtung besitzt. Die zu verarbeitenden Daten werden hier in das CNN eingebracht. Der Output aus der Eingangsebene passiert anschließend eine nicht lineare Aktivierungsfunktion.²³ Aktivierungsfunktionen dienen als eine Art Sortierfunktion des Eingangssignals. Sie sollen unwichtige von wichtigen Signalen am Eingang unterscheiden. Gerne werden hier rectified linear units (ReLU) oder Sigmoidfunktionen verwendet. Auf diese Art und Weise wird eine sogenannte Activation Map erstellt. Es folgt im Anschluss meist ein Pooling zur Minimierung der Datenmenge.²³ Dazu stehen verschiedene Verfahren zur Auswahl. Ergebnis der Durchführung von Pooling ist ein Datensatz mit weniger Auflösung als das Ursprungsbild. Dadurch ist der gewonnene Datensatz nach durchgeführtem Pooling weniger anfällig bzw. sensibel auf Veränderungen im Input-Bild. Der Output aus der Pooling Layer wird durch weitere Ebenen oder Pooling Layers des CNN geschleust, bis ein Ausgabewert am Ende der Datenverarbeitung steht.

Neurone einer Ebene eines CNN sind mit Neuronen der nächsttieferen Ebene des Netzwerkes verbunden. Vor Beginn der Trainingsphase ist jedes Neuron einer Ebene mit jedem Neuron der nächsten Ebene verbunden. Erhält ein Neuron ein Signal, kann es den Eingangswert durch seine Gewichtung und Bias

verändern und ein Signal an die nächsten Neuronen weitergeben. Durch Training des Netzwerkes und der damit verbundenen Veränderung des Bias und Gewichtung passt sich das Netzwerk der Problemstellung an. Verbindungen zwischen Neuronen können funktionell durch eine Gewichtung und ein Bias von null aufgelöst werden. Zum Training eines CNN gibt es verschiedene Methoden, meist werden mehrere verschiedene Trainings-, Validierungs- und weitere Testdatensätze verwendet. Auch hier gibt es verschiedene Methoden zur Auswahl von Größe und Inhalt von Trainingsdatensätze, worauf hier aber nicht näher eingegangen werden soll.^{24,25,23,22}

Neuronale Netze können durch unterschiedliche Methoden trainiert werden. Supervised Learning mit der dazugehörigen Backpropagation sollen hier als Stichworte beispielhaft genannt werden.

1.7.3 Anwendung künstlicher Intelligenz in der CT

Bisherige Rekonstruktionsalgorithmen basieren auf Methoden und Modellen, bei denen ein Mensch Intelligenz zur Erstellung der Methode aufbringen musste. Solche Bildrekonstruktionsmethoden stoßen ab einem bestimmten Punkt an ihre Grenzen und lassen sich nur noch immer schwieriger verbessern. Durch die Verwendung von KI-basierten Methoden gibt es neue Entwicklungsmöglichkeiten zur Verarbeitung von Bildmaterial.

Ziel der Weiterentwicklung der Rekonstruktionsalgorithmen ist es, die Bildqualität weiter zu steigern, Dosis einzusparen und kürzere Rekonstruktionszeiten der Bilder zu erreichen. Mit bisherigen Rekonstruktionsmethoden konnte meist nur einer der drei Parameter verbessert werden, jedoch durch Inkaufnahme einer Verschlechterung eines anderen Parameters. Als Beispiel sei hier die FBP-Rekonstruktionsmethode genannt. Bei der Verwendung von FBP-Algorithmen besitzen die rekonstruierten Bilder ein hohes Bildrauschen und Tendenz zur Entwicklung von Streifenartefakten, da diese wenig bis gar nicht korrigiert werden. Die Rekonstruktionsdauer bei der Verwendung von FBP-Rekonstruktionsalgorithmen ist jedoch sehr niedrig. Es wird also eine mäßige Bildqualität für eine kurze Rekonstruktionsdauer in Kauf genommen.^{21,15}

Die Verwendung von komplett modellbasierten iterativen Rekonstruktionsalgorithmen (MBIR) hilft Bilder zu produzieren, welche eine bessere Signal-to-Noise Ratio (SNR) besitzen und eine niedrigere Strahlendosis zur Erstellung der CT-Aufnahmen benötigen. Jedoch werden statistische und optische Modelle als Grundlage zur sehr rechenintensiven Bildrekonstruktion benötigt. Die Rekonstruktionszeit ist bei gleicher Rechenleistung und Auflösung höher als bei FBP-Algorithmen.¹⁵

In den letzten Jahren sind Ansätze zur Bildrekonstruktion entstanden, welche auf künstlicher Intelligenz basieren. Sie versprechen hohe Bildqualität bei niedriger Strahlendosis und schneller Bildrekonstruktion.¹⁵

Im Folgenden soll ein kurzer Überblick über den aktuellen Stand von KI-basierten Rekonstruktionsalgorithmen und weiteren KI-Ansätzen in der Bildverarbeitung von medizinischen Bildern gegeben werden.

Die aktuelle Forschung und Entwicklungsarbeit von KI-basierter Software zur Bildrekonstruktion und Bildverarbeitung von medizinischen Bildern versucht zum einen althergebrachte Fragestellungen besser zu beantworten als bisherige Lösungsansätze und zum anderen neue, bisher unlösbare Fragestellungen zu beantworten.

Einige bereits umgesetzte Beispiele, in denen KI-basierte Software zum Einsatz kommt, werden im folgenden Abschnitt genannt.

Rausch- und Artefaktreduzierung

Reduzierung von Bildrauschen und Artefakten ist ein Problem, mit dem sich die CT-Bildgebung seit ihrer Erfindung beschäftigt. Wird Dosis bei CT-Scans eingespart und mit analytischen Bildrekonstruktionen, wie der wFBP, gearbeitet, steigt das Bildrauschen und die Aufnahmen können zahlreiche Artefakte (z.B. Streakartefakte) besitzen. KI-Rekonstruktionen versuchen tatsächliche Bildinformation und Rauschen bzw. Artefakte voneinander zu unterscheiden. So kann als Resultat die eigentliche Bildinformation besser dargestellt werden.²⁵

Rekonstruktion aus nicht vollständigen Datensätzen

Bei Anwendungen, die möglichst kurze Scandauer erfordern, wie interventionellen Eingriffen oder Herz-CT-Scans, kann es zu einer nicht vollständigen Abdeckung einer Schicht im Radonraum kommen. Solche Lücken in der Datenerhebung beeinflussen bei der Bildrekonstruktion die spätere Bildqualität der Aufnahmen. KI-Bildrekonstruktionen können Zusammenhänge erkennen, ob eine tatsächliche Bildinformation vorliegt, es sich um ein Artefakt handelt oder eine eigentlich nötige Projektion fehlt. Datenlücken können von KI-Bildrekonstruktionen gefüllt werden. Dadurch werden keine neuen Informationen gewonnen, da die errechneten Daten nicht auf tatsächlich gemessenen Daten beruhen. Es werden lediglich die gemessenen Daten bereinigt, an Artefakten reduziert und mit niedrigerem Rauschlevel als Bild dargestellt. Dadurch soll die Interpretierbarkeit der Aufnahmen gesteigert werden, wodurch die Diagnosefindung erleichtert werden soll.²⁵

Rekonstruktion mit verbesserter Auflösung

Ein weiterer Ansatz beschäftigt sich mit dem Versuch aus CT-Aufnahmen mit niedriger Ortsauflösung, Aufnahmen mit höherer Ortsauflösung zu generieren. Dies geschieht auf Grundlage von CT-Aufnahmen mit hoher Auflösung, aus denen ein Algorithmus gelernt hat. Erhält der Algorithmus Bilddaten, werden zusätzlich neue Bilddaten aus den Erfahrungswerten des Algorithmus berechnet. Eine tatsächliche Bildinformation liegt jedoch nicht vor. Die Ergebnisse, die von Umehara et al. 2018 beschrieben wurden, wirken sehr natürlich und hochauflösend.^{25,26} Ob ein diagnostischer Gewinn in diesem Ansatz liegt, muss noch geklärt werden.²⁵

Erstellung von Pseudo-CT-Bildern aus MRT-Datensätzen

Strahlendosisreduktion kann durch Unterlassen einer Behandlung mit Röntgenstrahlung erreicht werden. Ein weiterer KI-basierter Ansatz beschäftigt sich mit der Generierung von Pseudo-CT-Bildern aus MRT-Datensätzen.

Grundlage sind echte Patientendaten aus MRT-Aufnahmen, welche dann zu errechneten CT-Datensätzen umgerechnet werden.²⁵

Limitierungen

Es gibt zahlreiche weitere Ansätze KI-basierter Software zur Bilderstellung oder Rekonstruktion von medizinischem Bildmaterial. Alle Ansätze haben Limitationen. So muss stets bewusst sein, dass KI-Software nie neue Informationen über einen Patienten generiert, sondern ausschließlich die Darstellung und Verarbeitung von real durch Messung erhobenen Daten beeinflussen bzw. verbessern kann. KI-Algorithmen können Informationen auswerten und Zusammenhänge erkennen und daraus bisher nicht erkennbare Informationen darstellen. Aus den erhobenen Daten werden mehr bzw. andere bisher nicht betrachtete Informationen entnommen.²⁵

Auf ausgiebige Testverfahren und eine korrekte Konzeption bei Entwicklung der Software der Algorithmen sollte nicht verzichtet werden. Bilder, die mit KI-basierten Algorithmen erstellt werden, zeichnen sich oft durch einen herausstechenden Bildeindruck aus.²⁵ Dies birgt die Gefahr, dass solche Bilder ohne Hinterfragen als korrekt angenommen werden, obwohl ggf. Informationen vom Algorithmus nicht korrekt verändert wurden.²⁵

1.8 Fragestellung der Arbeit

In der Einleitung wurden verschiedene Themenkomplexe der CT angesprochen und beschrieben. Eine gute Bildqualität und niedrige Strahlendosis sind Kernpunkte für die klinische Alltagstauglichkeit von CT-Aufnahmen. Bei herkömmlichen Verfahren zur Bildrekonstruktion in der CT wird das Verbesserungspotential immer kleiner, da die Modelle der Bildrekonstruktionsalgorithmen bereits sehr weit entwickelt sind und nur noch in kleinen Schritten Verbesserungen erzielt werden können. KI-basierte Rekonstruktionsalgorithmen versprechen nochmals die Bildqualität zu verbessern und die Strahlendosis zu senken.

Nachfolgender Teil der Arbeit ist in einen klinischen Teil und einen Phantomversuch aufgeteilt. Der klinische Teil der Arbeit beschäftigt sich mit dem Vergleich von ULD-CT-Aufnahmen, die mit einem FBP-, iterativen und KI-basierten Bildrekonstruktionsalgorithmus nachbearbeitet wurden. Es wird unter anderem die Auswirkung der Bildrekonstruktionsmethode auf Bildqualität, diagnostische Genauigkeit und Time to diagnosis untersucht.

Der Phantomversuch soll der Frage nachgehen wie und ob Strahlendosis bei CT-Aufnahmen durch die Verwendung von KI-basierten Bildrekonstruktionsalgorithmen im Vergleich zu iterativen und FBP-Algorithmen eingespart werden kann.

2 Material und Methoden

2.1 Patientenstudiendesign, Population und Strahlendosis

Eine a-priori-Power-Analyse wurde mit der Softwarelösung G*Power (Version 3.1.9.7, Franz Faul, Universität Kiel, Deutschland) durchgeführt. Daraus wurde die notwendige Stichprobengröße ($f = 1,18$, $\alpha = 0,05$, $1-\beta = 0,95$) auf 100 Patienten ermittelt. Die Ethik-Kommission der Medizinischen Fakultät der Universität Tübingen genehmigte für dieses Projekt die retrospektive Analyse und Datenerhebung von Patienten der Radiologischen Universitätsklinik (#167/2022BO2). Zunächst erfolgte eine Datenbanksuche, die mit folgenden Stichwörtern durchgeführt wurde: Pneumonie UND Pädiatrie UND Computertomographie. Wenn ein Patient im angegebenen Zeitraum mehrere Thorax-ULD-CT-Untersuchungen erhalten hatte, wurde nur die jüngste Untersuchung eingeschlossen und die anderen verworfen. Weitere Ausschlusskriterien waren eine andere klinische Indikation als der Verdacht auf eine Lungenentzündung, eine kontrastverstärkte Bildgebung und andere Bildaufnahmeprotokolle als die ULD-CT, welche in der Kinderradiologie Tübingen im klinischen Einsatz standardisiert mittels eines spektral gefilterten 100 kV Protokolls mit Zinn-Filter (Sn100) durchgeführt wird. So wurden die neusten ULD-CT-Aufnahmen von 100 Patienten gesammelt. Dazu wurde das Alter und

Geschlecht der Patienten zum Zeitpunkt der Untersuchung aus den klinischen Berichten erfasst. Für den Vergleich der Strahlendosis wurde die SSDE (Size Specific Dose Estimated in mGy) des Patienten mit Hilfe der Dosis-Management-Software DoseM[®] (Infinit Europe GmbH, Frankfurt am Main, Deutschland) errechnet.

2.2 Bildaufnahme und Rekonstruktionsparameter

Alle CT-Untersuchungen der Patientenstudie wurden ohne Kontrastmittel und auf demselben Computertomographen erstellt (CT der 3. Generation, SOMATOM Force; Siemens Healthineers, Erlangen, Deutschland). Für die Untersuchung wurde eine schwächungsbasierte Röhrenstrommodulation (CARE Dose4D, Referenz 190 mAs) verwendet. Die Röhrenspannung war auf Single-Source 100 kV mit Zinnfilter (Sn100) eingestellt. Die Kollimation wurde auf 0,6 x 192/128 mm, der Pitch-Faktor auf 2,8 ("Flash-Modus") und die Gantry-Rotationszeit auf 0,25 s eingestellt. Zur Rekonstruktion aller Bilder wurde ein mittelweicher Kernel (Br40d) verwendet. Die CT-Datensätze wurden in axialem Schnitt mit einer Schichtdicke und einem Inkrement von 1 mm unter Verwendung mit einer gewichteten gefilterten Rückwärtsprojektion (Weighted Filtered Back Projection, wFBP) und einem iterativen Rekonstruktionsalgorithmus (Advanced Modeled Iterative Reconstruction strength 2 (ADMIRE[®]), Siemens Healthineers, Erlangen, Deutschland) rekonstruiert, welches im weiteren Verlauf iterative Rekonstruktion Stärke 2 (IR2) genannt wird. IR2 entspricht dem aktuell verwendeten Nachbearbeitungsprozess in der Diagnostischen und Interventionellen Radiologie der Universitätsklinik Tübingen. Zusätzlich wurde eine neuartige KI-basierte Nachbearbeitungssoftware (PixelShine[®], AlgoMedica, Sunnyvale, CA, USA) verwendet, um das Bildrauschen der wFBP-Bilder zu senken. Die KI rekonstruierte Datenreihe wird im nachfolgenden KI genannt. Insgesamt entstanden so drei Datensätzen pro Patient.

2.3 Objektive Bildqualität

Alle drei Bildserien der Patientenstudie (wFBP, IR2, KI) wurden in die Open-Source-ImageJ-Distribution FIJI (Version 1.53k, Wayne Rasband, National Institutes of Health-NIH, Maryland, USA) geladen. Die wFBP-Serien wurden verwendet, um in 30 verschiedenen Regions of Interest (ROI) auszumessen. Die verwendeten ROI wurden so gewählt, dass sie einen Durchmesser von ≥ 1 cm besessen haben und wurden bilateral in homogener paraspinaler Muskulatur platziert. Auf jeder Seite der Dornfortsätze wurden drei ROIs in 5 aufeinanderfolgenden Schichten platziert. Die verwendeten Positionen der ROIs wurden bei jedem Patienten manuell in möglichst homogene Bereiche der paraspinalen Muskulatur gelegt. Die Positionen der ROIs wurden für alle drei Bildserien je Patient an derselben Stelle belassen. Dies erfolgte automatisiert, um eine hohe Ortskonsistenz der Messungen durch die verschiedenen Bildserien hinweg zu erhalten. In jeder ROI wurde der Mittelwert und die Standardabweichung der CT-Zahlen in Hounsfield-Einheiten (HU) gemessen bzw. ermittelt. Die Standardabweichung der HU wurde als Maß für Bildrauschen definiert.²⁷

Analog zur Patientenstudie wurde in der Phantomstudie ebenso über 30 ortskonsistente ROIs in 5 aufeinanderfolgenden Schichten das Bildrauschen bestimmt.

2.4 Subjektive Bildqualität

Die Bilddatensätze wurden von einem Mitarbeiter der Diagnostischen und Interventionellen Radiologie der Universitätsklinik Tübingen, der mit diesem Projekt nicht in Verbindung stand, anonymisiert und randomisiert.

Insgesamt 8 verblindete Teilnehmer mit unterschiedlich viel Erfahrung in der Befundung von pädiatrischen ULD-CT-Thoraxaufnahmen bewerteten unabhängig voneinander die subjektive Bildqualität. Die Staffelung des Erfahrungsgrades sah wie folgt aus: Rater 1 = Medizinstudent am Ende eines dreimonatigen Praktikums, Rater 2-6 = Ärzte in Weiterbildung in der Radiologie

mit 1, 2, 3, 4 und 5 Jahren Erfahrung, Rater 7-8 = Fachärzte für Radiologie mit 6 und 7 Jahren Erfahrung).²⁷

Die subjektive Bildqualität wurde nach den diagnostischen Anforderungen gemäß dem Kapitel "Chest, General" der European Guidelines on Image Quality in Computed Tomography bewertet.²⁷

Dazu zählen unter anderem die visuell scharf abgrenzbare Darstellung von Strukturen wie der thorakalen Aorta, der Trachea inklusive Carina, des Oesophagus, der pleuramediastinale Übergang, große und mittlere Blutgefäße der Lunge, segmentale Bronchen und die Abgrenzung von Pleura und Brustkorbwand.

Jeder Rater vergab eine Gesamtnote für die Kriterien anhand einer 5-stufigen Likert-Skala (1 = schlecht, 2 = suboptimal, 3 = befriedigend, 4 = gut, 5 = ausgezeichnet).

2.5 Diagnostische Genauigkeit

Jeder Rater füllte für die subjektive Bildqualitätsanalyse eine eigens angelegte semiquantitative Severity-Score Tabelle aus. Die Tabelle diente anschließend zum Bilden eines Summenscores, mit dem die Aufnahmen nach subjektiven Auffälligkeiten in drei Gruppen aufgeteilt wurden (0 = keine Beeinträchtigung, 1 = 0-50 % Beeinträchtigung, 2 = 51-100 % Beeinträchtigung). Zur Bestimmung der diagnostischen Genauigkeit wurden Kriterien aufgestellt. Dazu gehörten sechs häufige pneumoniebezogene Pathologien. Aufgenommen in die Wertung wurden peribronchiale Verdickungen („cuffing“), Mucus Plugging, Milchglastrübungen, Kavitäten, Konsolidierungen und Air Trapping. Die Bewertung der Aufnahmen erfolgte in sechs unterschiedlichen Lungenregionen. Dazu zählten der rechte Oberlappen, rechte Mittellappen, rechte Unterlappen, linke Oberlappen, das Lingulasegment und der linke Unterlappen.²⁷

2.6 Time to diagnosis

Die Time to diagnosis, hier definiert als die Zeit, die benötigt wurde, um die korrekte Diagnose zu stellen und die subjektive Bewertung der Bildqualität ausfüllen zu können. Als korrekte Diagnose wurde die im klinischen Befundbericht dokumentierte Diagnose der jeweiligen Aufnahme angenommen. Die Time to diagnosis wurde für jeden Rater gemessen, um potenzielle Unterschiede in der Time to diagnosis in Bezug auf die verschiedenen Rekonstruktionsalgorithmen zu untersuchen. Zusätzlich wurden eine posthoc-Subgruppenanalysen nach der Berufserfahrung durchgeführt, um mögliche Abhängigkeiten zwischen Berufserfahrung, Rekonstruktionsalgorithmus und Time to diagnosis zu untersuchen.²⁷

2.7 Phantomversuch

Das Institut für Diagnostische und Interventionelle Radiologie des Universitätsklinikums Tübingen verwendet bereits die niedrigstmöglichen Doseinstellungen für die ULD-CT des pädiatrischen Thorax. Wie eingangs erläutert, ist die Strahlendosis in der Computertomographie jedoch eng mit dem zu erwartenden Bildrauschen verknüpft. Neben dem uns vorliegenden KI Algorithmus zur Entrauschung können ähnliche Effekte auch durch Aufnahmen mit höherer Strahlendosis erreicht werden. Es wurde ein 16 cm Wasserphantom (Siemens slice thickness phantom, Model-Nr. 10742380) unter sequenzieller Dosiserhöhung gescannt, um die Möglichkeiten des KI-basierten Algorithmus zur Reduzierung der Strahlendosis zu untersuchen.

In der Phantomstudie wurden dieselben Scanparameter aus der klinischen Routine verwendet, mit der Ausnahme, dass CARE Dose4D (automatische Belichtungskontrolle) ausgeschaltet wurde. Stattdessen wurde das mittlere $CTDI_{vol}$ unserer Studienpopulation für einen vergleichbaren 100%-Scan verwendet. Innerhalb der Versuchsreihe wurde das Röhrenstrom-Zeit-Produkt schrittweise um 25 % der ursprünglichen Strahlendosis gesteigert. So wurden CT-Aufnahmen von dem Phantom mit einer Strahlendosis im Bereich von 100 % bis 600 % im Vergleich zur Studienpopulation erstellt. Die CT-Aufnahmen wurden

anschließend auf ihr Bildrauschen untersucht und verglichen. Besonderer Fokus wurde auf den Vergleich gelegt, bei welcher Strahlendosis die drei Bildrekonstruktionsalgorithmen das gleiche Bildrauschen besitzen. Dazu wurde ein 100 % Scan vom Phantom durchgeführt. Die Aufnahme des 100 % Scans und Rekonstruktion mittels wFBP + PixelShine, im weiteren Verlauf nur noch KI genannt, diente als Vergleichsaufnahme für alle anderen Scans und Rekonstruktionsvarianten. Verglichen wurde die 100 % KI-Aufnahme mit Scans höherer Dosis, welche mit wFBP oder IR2 rekonstruiert wurden. Anschließend wurde ermittelt, bei welcher Dosis die herkömmlichen Rekonstruktionsalgorithmen eine äquivalente Bildqualität im Vergleich zu den 100 % KI-Aufnahmen besitzen.

Vorherige Studien haben bereits eine Input-unabhängige Performance des Algorithmus bestätigt.²⁸ Deshalb wurden in diesem Teil der Studie nur die 100 % Dosis wFBP-Bilder mit dem KI-basierten Rekonstruktionsalgorithmus entrauscht und mit den anderen Bildrekonstruktionsalgorithmen und Dosisstufen verglichen. Es wurden 23 Datensätze (1 x 100% KI, 11 x wFBP 100%-600%, 11 x IR2 100%-600%) erstellt.

Die statistische Auswertung erfolgte mittels einer Mixed-Effekt Analyse.

2.8 Statistische Auswertung

Die statistische Analyse und graphische Aufbereitung wurde mit GraphPad Prism Version 9.3.1 für Windows (GraphPad Software, San Diego, Kalifornien, USA) durchgeführt. Die Datenverteilung wurde mit dem Shapiro-Wilk-Test geprüft. Normal verteilte Variablen wurden als Mittelwert \pm SD und nicht normal verteilte Variablen als Median und Interquartilsbereich (IQR) angegeben.

Die Datenauswertung erfolgte mit Mixed-Effect-Modell Analysen. Eine Greenhouse-Geisser-Korrektur wurde im Falle der Verletzung der Sphärizität angewandt. Eine zweistufige Step-Up-Korrektur nach Benjamini, Krieger und Yekutieli wurde verwendet, um Fehler erster Ordnung bei posthoc Mehrfachvergleichen zu minimieren.

Ein korrigierter p-Wert von $p \leq 0,05$ wurde als hinweisend auf statistische Signifikanz festgelegt.

Der Spearman-Korrelationskoeffizient, angegeben als r , wurde zur Überprüfung der Übereinstimmung zwischen den Bildbetrachtern bei der subjektiven Bildqualitätsanalyse und bei der Erstellung der Severity-Score Tabellen im Teilversuch Diagnostische Genauigkeit verwendet. Korrelationskoeffizienten zwischen 0 – 0,3 wurden als vernachlässigbar, 0,31 – 0,5 als gering, 0,51-0,70 als mäßig, 0,71 – 0,9 als gut und 0,91 – 1,00 als hervorragend gewertet.

3 Ergebnisse

3.1 Patientenkohorte

Insgesamt wurden 317 pädiatrische Ultra Low-Dose Computer Tomographieaufnahmen (ULD-CT) aus den Jahren 2020 und 2021 auf Eignung zum Einschluss in die Patientenkohorte überprüft. Davon wurden insgesamt 217 ausgeschlossen: 103 der 217 ausgeschlossenen Patienten wurden mit Aufnahmetechniken ohne ULD-CT-Protokoll untersucht. Weitere 95 der 217 ausgeschlossenen Patienten erhielten eine CT-Untersuchung, aber nicht aufgrund der Fragestellung einer Pneumonie. Des Weiteren wurden 19 Aufnahmen ausgeschlossen, da es sich um Doppelt-Untersuchungen von Patienten handelte. Es erfüllten 100 Untersuchungen die Einschlusskriterien. Alle 100 Untersuchungen wurden mit demselben CT-Scanner durchgeführt. Es kam ein Dual-Source Scanner der dritten Generation zum Einsatz (Siemens SOMATOM Force).²⁷

Die Studienkohorte umfasste 45 weibliche und 55 männliche Patienten, wobei das Durchschnittsalter der gesamten Kohorte 10 ± 7 Jahre betrug. Das Durchschnittsalter nach Aufteilung anhand von Geschlecht ergab folgende Verteilung weiblich: 10 ± 8 Jahre; männlich: 10 ± 6 Jahre.

Die durchschnittliche größenspezifische Strahlendosis (SSDE) der Patienten betrug $0,23 \pm 0,09$ mGy. Auf die Geschlechter aufgeteilt, entfiel die SSDE wie folgt weiblich: $0,25 \pm 0,11$ mGy; männlich: $0,22 \pm 0,06$ mGy.

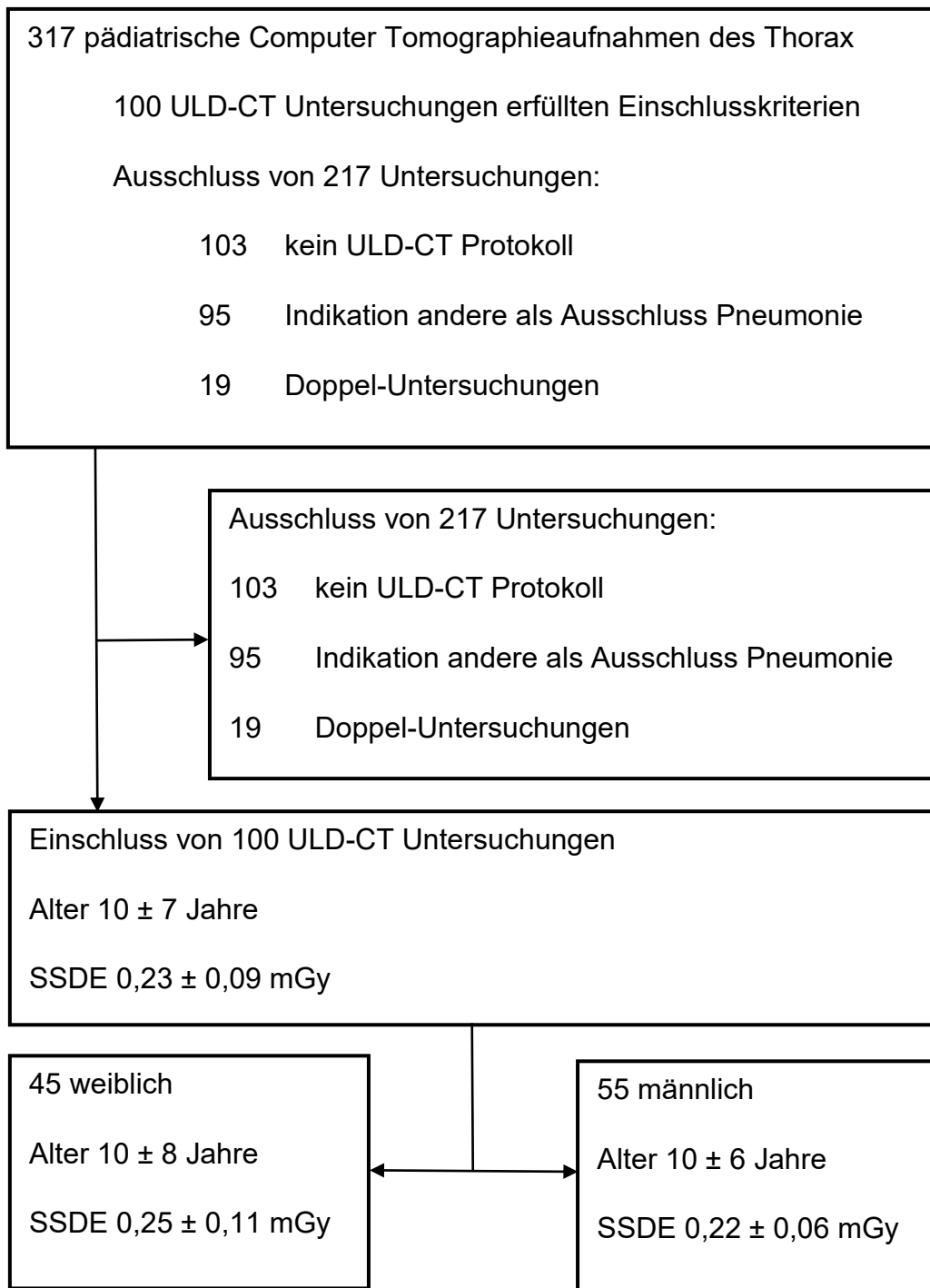


Abbildung 2: Aufgliederung der Patientenkohorte ²⁷

Der Untersuchungsgrund war stets der Verdacht auf eine Pneumonie. Die Ursache, auf der die Pneumonie entstand, konnte in verschiedene Gruppen eingeteilt werden. 60 Patienten erhielten ein CT-Screening aufgrund eines onkologischen Hintergrunds, davon 28 weiblich und 32 männlich. 20 Patienten wurden wegen einer Infektionskrankheit untersucht, davon 8 Patienten weiblich.

Weitere 15 Untersuchungen wurden mit kongenitalem Hintergrund durchgeführt (8 weiblich; 7 männlich). Und insgesamt fünf Patienten erhielten ein ULD-CT, weil eine chirurgische Ursache vorlag.²⁷

Untersuchungsgrund Pneumonie bei	Anzahl Untersuchungen		
	weiblich	männlich	gesamt
onkologischem Hintergrund	28	32	60
Infektion	8	12	20
kongenitalem Hintergrund	8	7	15
chirurgischer Ursache	0	5	5

Tabelle 2: Aufteilung der Studienkohorte nach Untersuchungsgrund bei Pneumonie ²⁷

Jede der ausgewählten 100 Untersuchungen wurde nach festem Schema, wie bereits in 2.2 Bildaufnahme und Rekonstruktionsparameter beschrieben, auf drei unterschiedliche Art und Weisen rekonstruiert und nachbearbeitet. Pro Untersuchung entstanden so drei Datensätze (wFBP, iterative Rekonstruktion (IR2), KI Rauschreduzierung). Anhand dieser drei Datensätze erfolgte die Untersuchung bezüglich objektiver und subjektiver Bildqualität, wie auch die Überprüfung der diagnostischen Genauigkeit und der Messung der Time to diagnosis.²⁷

3.2 Objektive Bildqualität

Die objektive Bildqualität wurde wie in 2.3 Objektive Bildqualität beschrieben ermittelt.

In der Patientenstudie sahen die Ergebnisse wie folgt aus.

Aufnahmen, die mit der wFBP rekonstruiert wurden, hatten einen Bildrauschen von durchschnittlich $76,9 \pm 9,62$ HU. Die iterative Rekonstruktion Stufe 2 folgte

mit einem Bildrauschen von $43,4 \pm 4,45$ HU. Der KI basierte Algorithmus reduzierte das Bildrauschen auf durchschnittlich $34,8 \pm 3,27$ HU.

Das höchste Bildrauschen wurde in wFBP-Aufnahmen gemessen. Mit absteigendem Bildrauschen wurden die wFBP-Aufnahmen von der IR 2 und der KI-Rekonstruktion gefolgt.²⁷

Eine Mixed-Effect Analyse hat einen signifikanten Unterschied bezüglich des Bildrauschens zwischen den einzelnen Gruppen dargelegt ($F(1,00; 99,0) = 4348$; $p < 0,001$).

Die untenstehende Boxplot Grafik verdeutlicht die Unterschiede der Rauschlevel der einzelnen Datensätze.

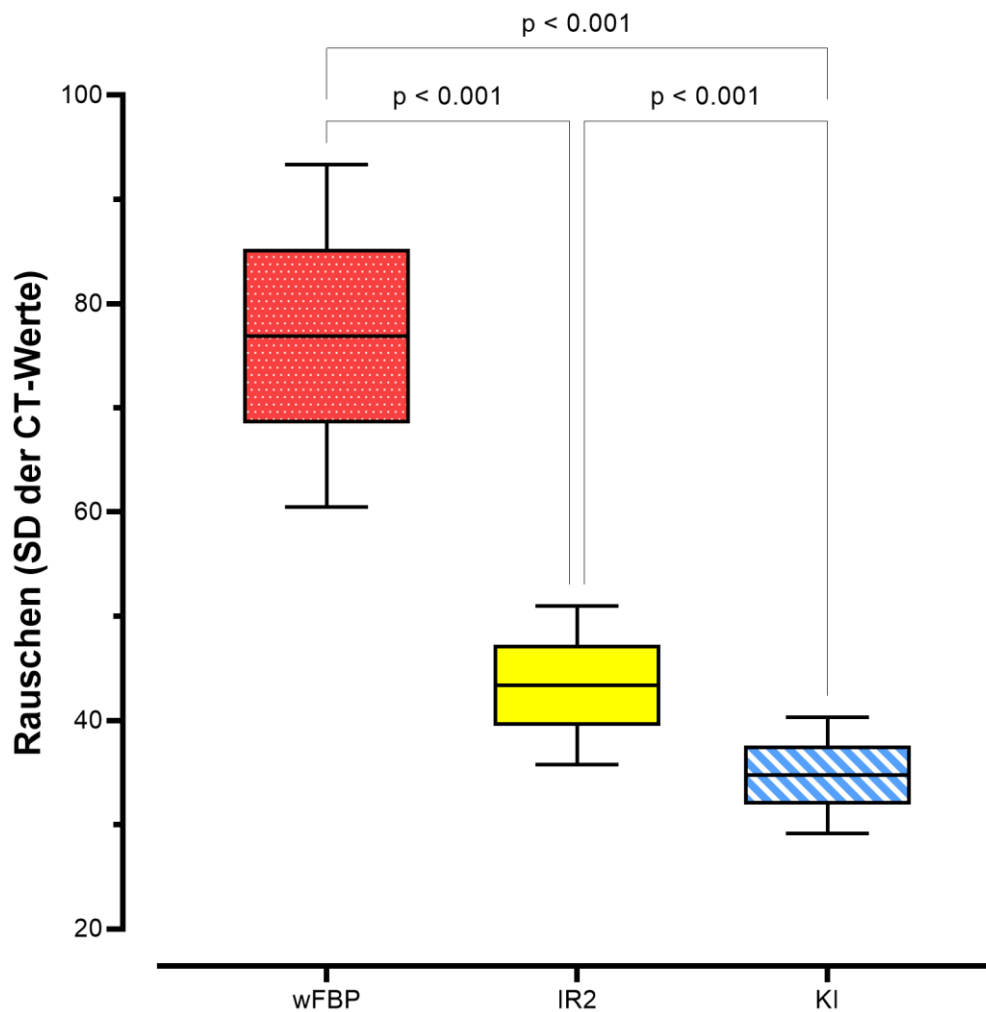


Abbildung 3: Bildrauschen der CT-Aufnahmen aus der Patientenstudie, separiert nach der Rekonstruktionsmethode²⁷

3.3 Subjektive Bildqualität

Die subjektive Bildqualität für die ULD-CT-Aufnahmen aus dem Datensatz wFBP wurden im Medianwert mit befriedigend (3) bewertet. Gefolgt wurden die wFBP-Aufnahmen in der Qualität von Aufnahmen, die mit IR2 rekonstruiert wurden. IR2 Aufnahmen wurde eine gute (4) Bildqualität zugeordnet. Eine ausgezeichnete (5) Bildqualität erhielten Aufnahmen in der Bewertung, welche mit KI nachbearbeitet wurden. Die höchste subjektive Bildqualität wurde bei den ULD-CT-Aufnahmen mit KI-Nachbearbeitung gemessen. Das iterative Rekonstruktionsverfahren IR2 folgte und wurde besser (4), als die Aufnahmen der wFBP bewertet. Die Unterschiede zwischen den Gruppen waren auf einem Niveau von $p < 0,001$ statistisch signifikant.²⁷

Die gemittelte subjektive Bildqualität im Vergleich zwischen den Gruppen kann in Abbildung 4 abgelesen werden.

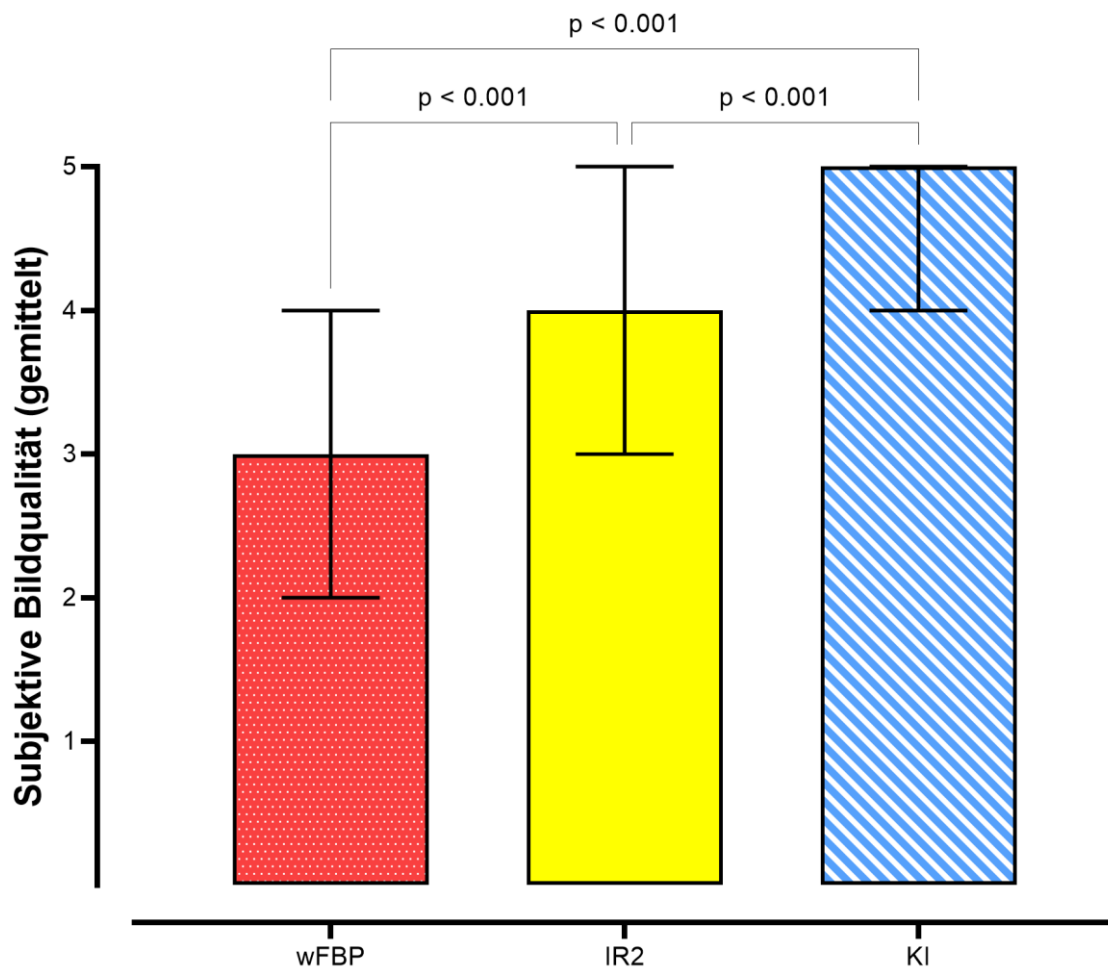


Abbildung 4: Verteilung der subjektiven gemittelten Bildqualität aufgeteilt in die einzelnen Rekonstruktionsmethoden²⁷

Die Interraterreliabilität wurde mit einer Spearman Korrelation überprüft. Diese war bei allen drei Datensätzen hoch bis sehr hoch.

Bei den wFBP-Aufnahmen betrug die Korrelation $r \geq 0.764$ bei $p \leq 0.001$. Eine höhere Interraterreliabilität fand sich bei den KI-Aufnahmen mit $r \geq 0.790$ bei $p \leq 0.001$. Angeführt mit der höchsten Interraterreliabilität ($r \geq 0.908$; $p \leq 0.001$) folgte der IR2 Datensatz. Zur genaueren Betrachtung der Korrelationen bezüglich der subjektiven Bildqualität zwischen den einzelnen Rater kann die untenstehende Tabelle 3 herangezogen werden.²⁷

Rating

Spearman Korrelationskoeffizient

		Median (IQR)	Reader 1	Reader 2	Reader 3	Reader 4	Reader 5	Reader 6	Reader 7	Reader 8
wFBP	Reader 1	3 (2–3)	1.000	0.843	0.975	0.834	0.962	0.823	0.953	0.813
	Reader 2	3 (2–4)	0.843	1.000	0.814	0.991	0.803	0.980	0.794	0.970
	Reader 3	3 (2–4)	0.975	0.814	1.000	0.805	0.988	0.794	0.979	0.784
	Reader 4	3 (2–4)	0.834	0.991	0.805	1.000	0.793	0.989	0.785	0.979
	Reader 5	3 (2–4)	0.962	0.803	0.988	0.793	1.000	0.782	0.991	0.772
	Reader 6	3 (2–4)	0.823	0.980	0.794	0.989	0.782	1.000	0.774	0.990
	Reader 7	3 (2–4)	0.953	0.794	0.979	0.785	0.991	0.774	1.000	0.764
	Reader 8	3 (2–4)	0.813	0.970	0.784	0.979	0.772	0.990	0.764	1.000
IR2	Reader 1	4 (3–5)	1.000	0.970	0.944	0.933	0.922	0.917	0.912	0.908
	Reader 2	4 (3–5)	0.970	1.000	0.971	0.957	0.945	0.939	0.933	0.928
	Reader 3	4 (3–5)	0.944	0.971	1.000	0.985	0.971	0.964	0.957	0.951
	Reader 4	4 (3–5)	0.933	0.957	0.985	1.000	0.985	0.978	0.971	0.964
	Reader 5	4 (3–5)	0.922	0.945	0.971	0.985	1.000	0.992	0.985	0.978
	Reader 6	4 (3–5)	0.917	0.939	0.964	0.978	0.992	1.000	0.992	0.985
	Reader 7	4 (3–5)	0.912	0.933	0.957	0.971	0.985	0.992	1.000	0.992
	Reader 8	4 (3–5)	0.908	0.928	0.951	0.964	0.978	0.985	0.992	1.000
KI	Reader 1	5 (4–5)	1.000	0.921	0.882	0.845	0.826	0.808	0.808	0.790
	Reader 2	5 (4–5)	0.921	1.000	0.958	0.918	0.898	0.878	0.878	0.858
	Reader 3	5 (4–5)	0.882	0.958	1.000	0.958	0.937	0.916	0.916	0.895
	Reader 4	5 (4–5)	0.845	0.918	0.958	1.000	0.978	0.956	0.956	0.935
	Reader 5	5 (4–5)	0.826	0.898	0.937	0.978	1.000	0.978	0.978	0.956
	Reader 6	5 (4–5)	0.808	0.878	0.916	0.956	0.978	1.000	1.000	0.977
	Reader 7	5 (4–5)	0.808	0.878	0.916	0.956	0.978	1.000	1.000	0.977
	Reader 8	5 (4–5)	0.790	0.858	0.895	0.935	0.956	0.977	0.977	1.000

Tabelle 3: Interraterreliabilität bei Bewertung der CT-Aufnahmen unterschiedlicher Rekonstruktionsmethoden ²⁷

3.4 Diagnostische Genauigkeit

Der gemittelte Severity-Score verteilte sich über die unterschiedlichen Gruppen wie folgt.

Bei den ULD-CT-Aufnahmen, die mit der wFBP rekonstruiert wurden, ergab sich ein gemittelter Severity-Score von 10 mit einer Abweichung von $\pm 6,4$ Punkten.²⁷

Die IR2 Aufnahmen verteilten sich ähnlich zu den wFBP-Aufnahmen mit einem gemittelten Severity-Score von 10 bei einer Abweichung von $\pm 6,4$ Punkten.²⁷

Die KI-Aufnahmen erhielten einen gemittelten Severity-Score von 10 mit einer Abweichung von $\pm 6,6$ Punkten.²⁷

Die diagnostische Genauigkeit unterscheidet sich statistisch auf nicht signifikantem Niveau bei Variation der Rekonstruktionsmethode.²⁷

Die Interraterreliabilität war bei den Aufnahmen wie folgt verteilt:

wFBP: $r \geq 0,764$

IR2: $r \geq 0,777$

KI: $r \geq 0,826$

Die Interraterreliabilität ist somit als solide zu beurteilen.²⁷ Zur genaueren Visualisierung dient untenstehende Tabelle 4.

	Severity-Score		Spearman Korrelationskoeffizient							
	Mean ± SD	Reader 1	Reader 2	Reader 3	Reader 4	Reader 5	Reader 6	Reader 7	Reader 8	
wFBP	Reader 1	11.90 ± 6.72	1.000	0.990	0.979	0.970	0.764	0.772	0.813	0.784
	Reader 2	11.70 ± 6.72	0.990	1.000	0.989	0.980	0.774	0.782	0.823	0.794
	Reader 3	11.60 ± 6.83	0.979	0.989	1.000	0.991	0.785	0.793	0.834	0.805
	Reader 4	11.50 ± 6.86	0.970	0.980	0.991	1.000	0.794	0.803	0.843	0.814
	Reader 5	9.35 ± 5.92	0.764	0.774	0.785	0.794	1.000	0.991	0.953	0.979
	Reader 6	9.15 ± 5.84	0.772	0.782	0.793	0.803	0.991	1.000	0.962	0.988
	Reader 7	9.03 ± 5.88	0.813	0.823	0.834	0.843	0.953	0.962	1.000	0.975
	Reader 8	9.03 ± 5.81	0.784	0.794	0.805	0.814	0.979	0.988	0.975	1.000
IR2	Reader 1	11.60 ± 6.72	1.000	0.996	0.993	0.989	0.782	0.778	0.783	0.777
	Reader 2	11.50 ± 6.76	0.996	1.000	0.998	0.994	0.784	0.779	0.785	0.780
	Reader 3	11.40 ± 6.76	0.993	0.998	1.000	0.997	0.790	0.787	0.794	0.786
	Reader 4	11.30 ± 6.71	0.989	0.994	0.997	1.000	0.797	0.795	0.802	0.794
	Reader 5	9.17 ± 5.7	0.782	0.784	0.790	0.797	1.000	0.996	0.985	0.989
	Reader 6	9.12 ± 5.78	0.778	0.779	0.787	0.795	0.996	1.000	0.986	0.991
	Reader 7	9.00 ± 5.85	0.783	0.785	0.794	0.802	0.985	0.986	1.000	0.990
	Reader 8	9.02 ± 5.81	0.777	0.780	0.786	0.794	0.989	0.991	0.990	1.000
KI	Reader 1	11.20 ± 6.45	1.000	0.998	0.996	0.996	0.830	0.831	0.861	0.826
	Reader 2	11.10 ± 6.49	0.998	1.000	0.997	0.996	0.831	0.832	0.864	0.827
	Reader 3	11.00 ± 6.43	0.996	0.997	1.000	1.000	0.845	0.845	0.878	0.840
	Reader 4	11.00 ± 6.45	0.996	0.996	1.000	1.000	0.846	0.846	0.879	0.841
	Reader 5	9.26 ± 5.93	0.830	0.831	0.845	0.846	1.000	0.999	0.989	0.998
	Reader 6	9.23 ± 5.92	0.831	0.832	0.845	0.846	0.999	1.000	0.990	0.999
	Reader 7	9.41 ± 5.93	0.861	0.864	0.878	0.879	0.989	0.990	1.000	0.988
	Reader 8	9.18 ± 5.90	0.826	0.827	0.840	0.841	0.998	0.999	0.988	1.000

Tabelle 4: Interraterreliabilität bei Ermittlung der diagnostischen Genauigkeit anhand eines Severity Scores²⁷

Zur Bewertung der diagnostischen Genauigkeit dienten verschiedene pathologische Korrelate einer Pneumonie. Dazu zählen unter anderem Milchglasinfiltrate oder peribronchiales Cuffing.²⁷

Die diagnostische Genauigkeit variierte je nach Korrelat, welches diagnostiziert werden sollte. Die durchgeführte Mixed-Effekt Analyse ergab zu diesem Zusammenhang ein Ergebnis von $F(3,68; 364) = 45,7$ bei $p < 0,001$.²⁷

Manche Korrelate einer Pneumonie konnten treffsicherer als andere gefunden werden. Peribronchiales Cuffing wurde beispielsweise sicherer erkannt als Cavitäten im Rahmen einer Pneumonie.²⁷

Unterschiede zwischen den einzelnen Rekonstruktionsalgorithmen innerhalb eines Bewertungskriteriums konnten jedoch nicht statistisch signifikant bewiesen werden. Die dazu durchgeführte Mixed-Effekt Analyse brachte ein Ergebnis von $F(5,71; 566) = 0,792$ bei $p = 0,570$.²⁷

Eine genauere Übersicht geben nachstehende Abbildungen. Sie zeigen die einzelnen Bewertungskriterien und Rekonstruktionsalgorithmen in Bezug auf die diagnostische Genauigkeit auf. Zudem finden sich einige Bildbeispiele zur Verdeutlichung, wie sich die einzelnen Rekonstruktionsalgorithmen auf die Bildqualität auswirken.

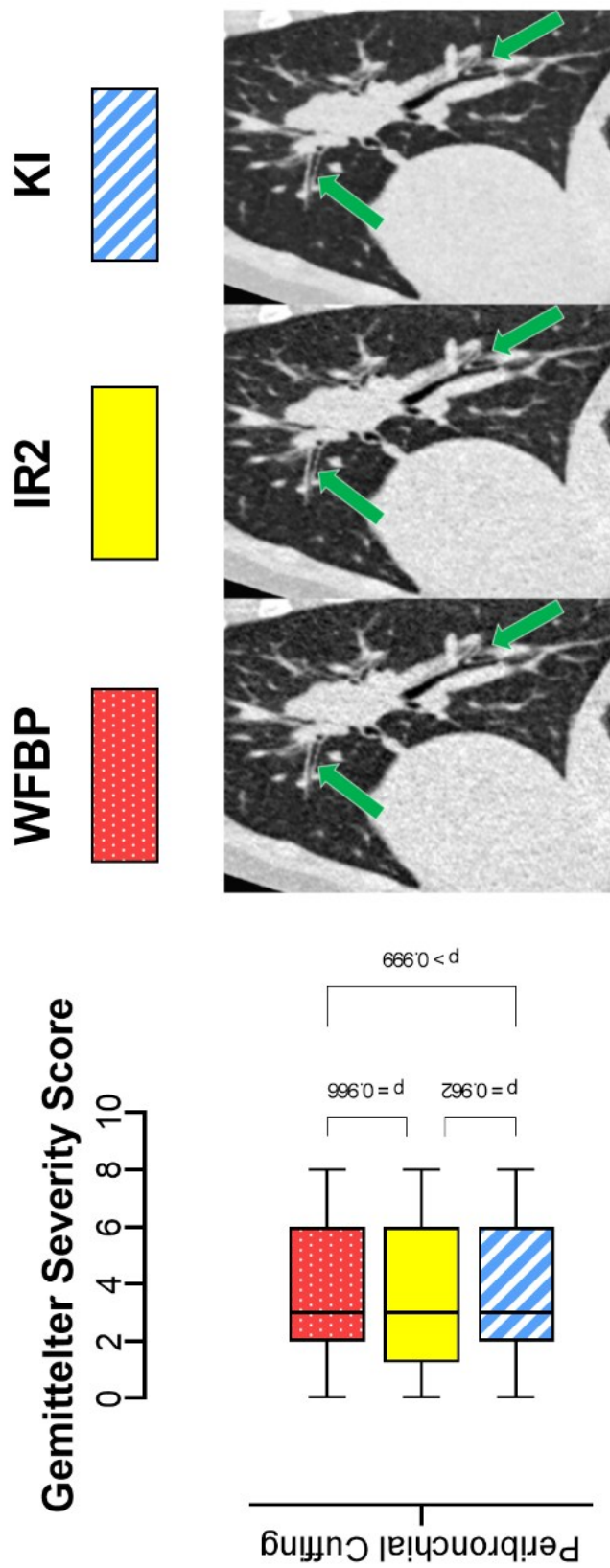


Abbildung 5: Visualisierung der diagnostischen Genauigkeit in Bezug auf peribronchiales Cuffing.²⁷

Abbildung 5 zeigt nebeneinander aufgereiht eine identische Schicht in sagittaler Schnittführung einer ULD-CT-Aufnahme eines Thorax, welche mittels der drei verschiedenen Rekonstruktionsmethoden (wFBP, IR2, KI) entstanden sind. In den Bildaufnahmen ist peribronchiales Cuffing zu erkennen, welches mit grünen Pfeilen markiert ist. Neben den ULD-CT-Aufnahmen ist der gemittelte Severity Score der einzelnen Rekonstruktionsmethoden dargestellt. Ebenso ist der p-Wert der durchgeführten Mixed-Effect-Analysen dargestellt. Der gemittelte Severity Score unterscheidet sich in Bezug auf das Kriterium Peribronchiales Cuffing zwischen den verschiedenen Rekonstruktionsmethoden auf nicht statistisch signifikanten Niveau.²⁷

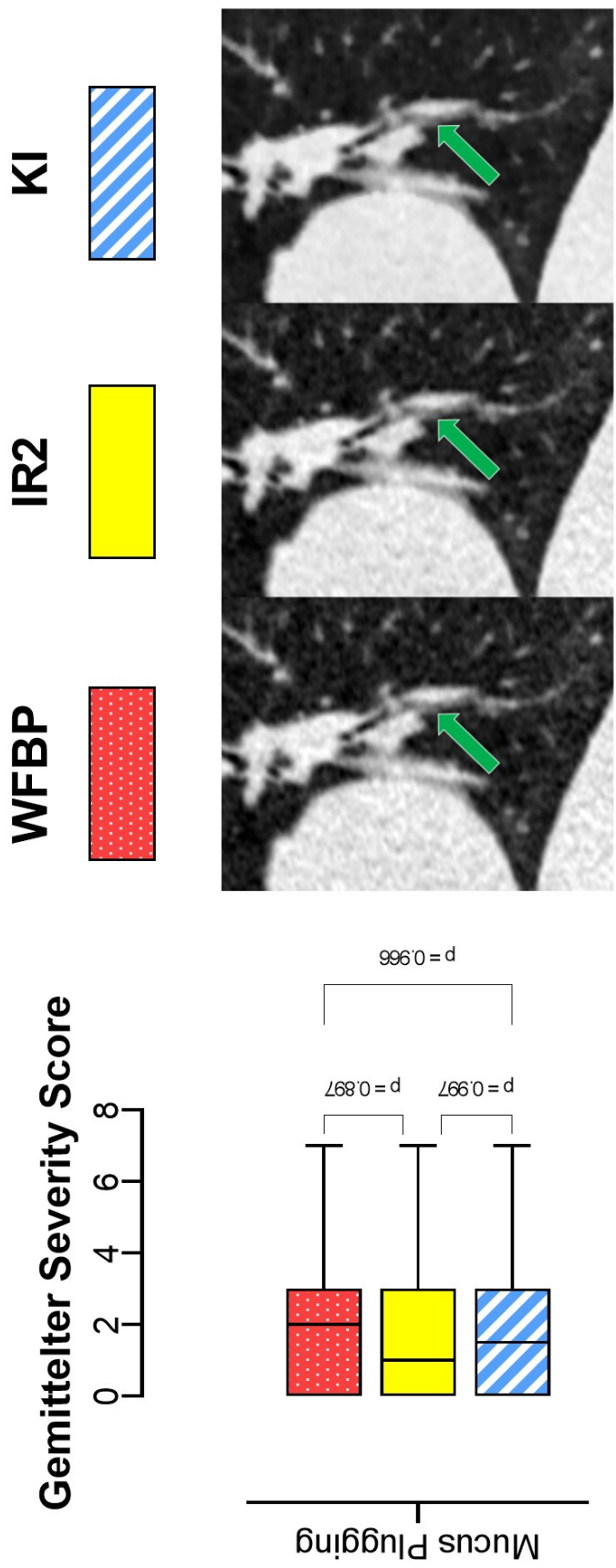
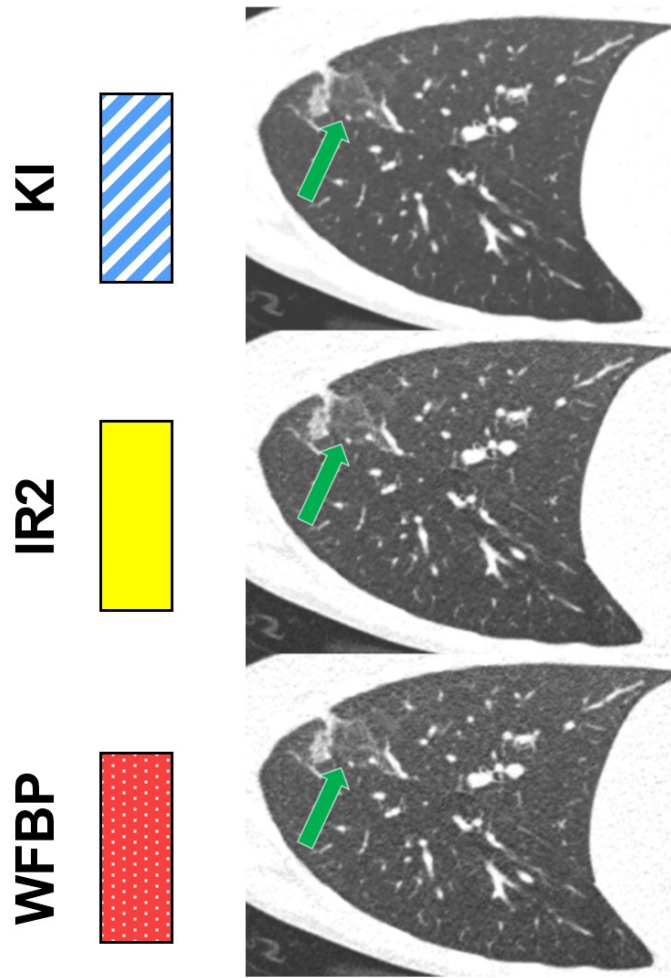


Abbildung 6: Visualisierung der diagnostischen Genauigkeit in Bezug auf Mucus Plugging.²⁷

Abbildung 6 zeigt ähnlich wie die vorangegangene Abbildung ein weiteres diagnostisch Kriterium einer Pneumonie. Im Falle der Abbildung 6 handelt es sich um das Kriterium Mucus Plugging. Es erfolgte erneut die Anordnung der Bilder der ULD-CT-Aufnahmen der unterschiedlichen Rekonstruktionsmethoden nebeneinander. Ebenso zeigen alle drei Bildausschnitte die identische Schicht in sagittaler Schnittführung. In den Bildausschnitten ist mit je einem grünen Pfeil Mucus Plugging markiert. Neben den ULD-CT-Aufnahmen ist der gemittelte Severity Score der einzelnen Rekonstruktionsmethoden dargestellt. Ebenso ist der jeweilige p-Wert der durchgeführten Mixed-Effect-Analysen dargestellt. Der gemittelte Severity Score unterscheidet sich bezüglich des Kriteriums Mucus Plugging zwischen den verschiedenen Rekonstruktionsmethoden auf nicht statistisch signifikanten Niveau.²⁷



Gemittelter Severity Score

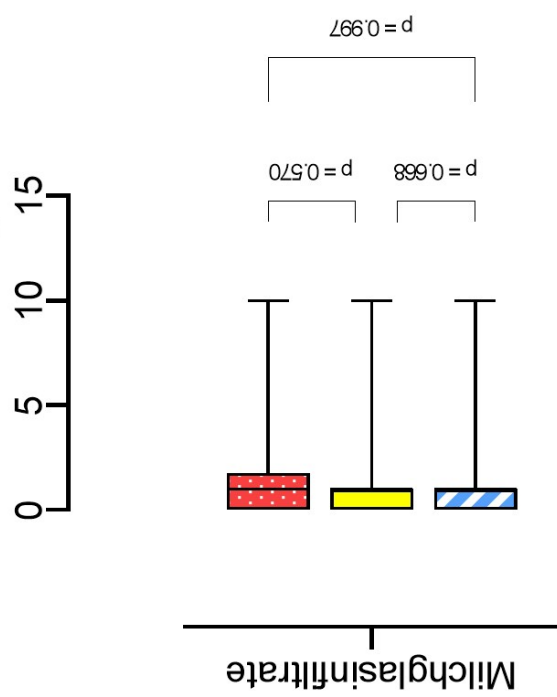


Abbildung 7: Visualisierung der diagnostischen Genauigkeit in Bezug auf Milchglasinfiltrate.²⁷

Auch Abbildung 7 zeigt nebeneinander aufgereiht eine identische Schicht in sagittaler Schnittführung einer ULD-CT-Aufnahme eines Thorax mit je verschiedener Rekonstruktionsmethode (wFBP, IR2, KI). In den Bildaufnahmen sind Milchglasinfiltrate zu erkennen, welche jeweils mit einem grünen Pfeil markiert sind. Neben den ULD-CT-Aufnahmen ist der gemittelte Severity Score der einzelnen Rekonstruktionsmethoden dargestellt. Ebenso ist der p-Wert der durchgeführten Mixed-Effect-Analysen dargestellt. Der gemittelte Severity Score unterscheidet sich zwischen den verschiedenen Rekonstruktionsmethoden in Bezug auf Milchglasinfiltrate auf nicht statistisch signifikanten Niveau.²⁷

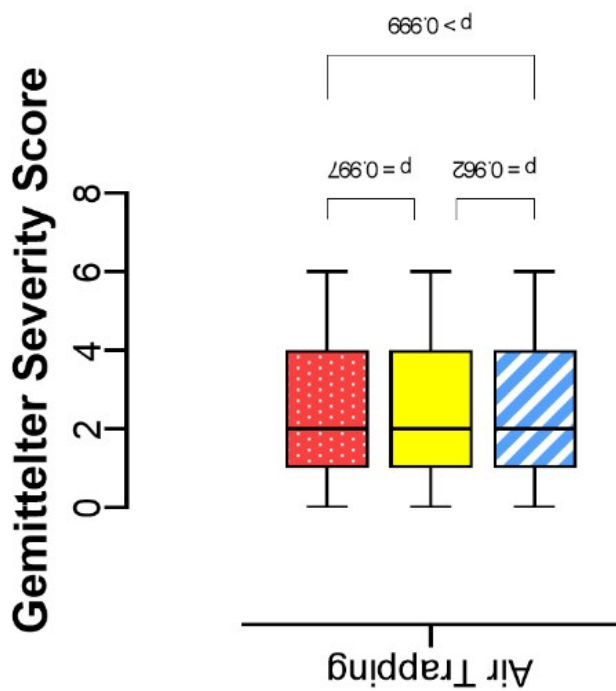
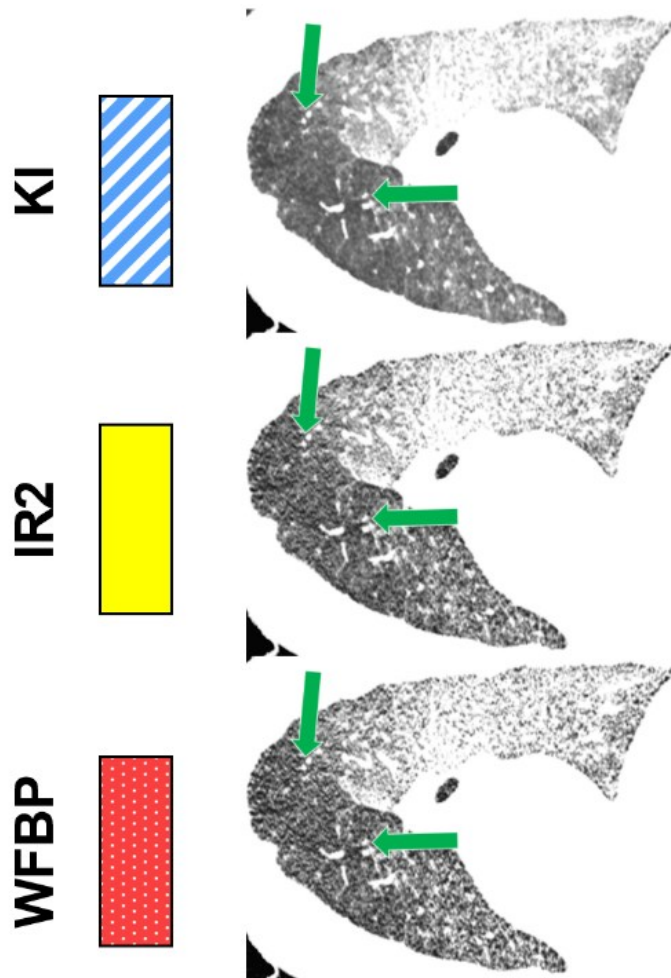


Abbildung 8: Visualisierung der diagnostischen Genauigkeit in Bezug auf Air Trapping.²⁷

Es folgt Abbildung 8: Angelehnt an die vorangegangenen Grafiken zeigt Abbildung 8 das Kriterium Air Trapping. Auch in Abbildung 8 erfolgte die Wahl auf einen sagittalen Schnitt durch eine ULD-CT-Thoraxaufnahme. Die Rekonstruktionsmethoden blieben die gleichen. Der gemittelte Severity Score in Bezug auf Air Trapping unterscheidet sich zwischen den verschiedenen Rekonstruktionsmethoden auf nicht statistisch signifikanten Niveau.²⁷

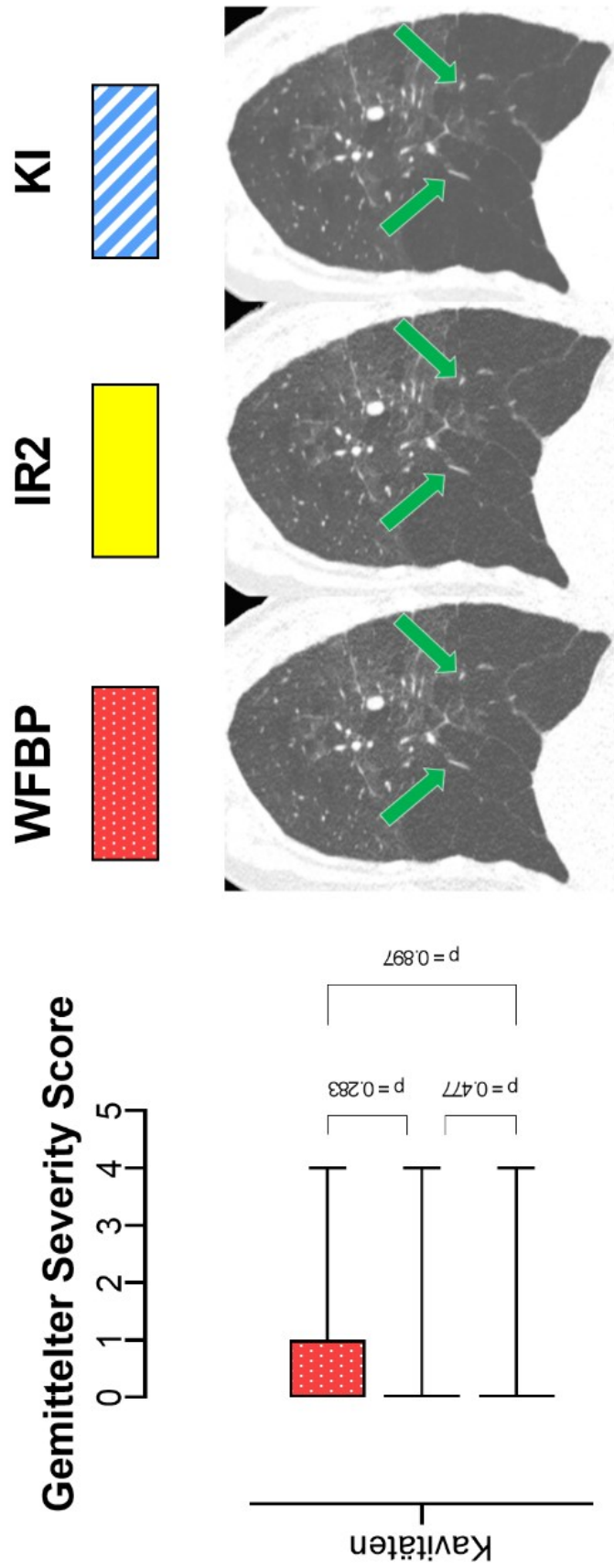


Abbildung 9: Visualisierung der diagnostischen Genauigkeit in Bezug auf Kavitäten.²⁷

Abbildung 9 besitzt den identischen Aufbau wie die vorangegangenen Grafiken und zeigt in sagittaler Schnittführung Kavitäten, welche mit grünen Pfeilen markiert sind. Ebenso findet sich der gemittelte Severity Score des Kriterium Kavitäten in Abbildung 9, wobei sich der Severity Score auf nicht statistisch signifikanten Niveau zwischen den verschiedenen Rekonstruktionsmethoden unterscheidet.²⁷

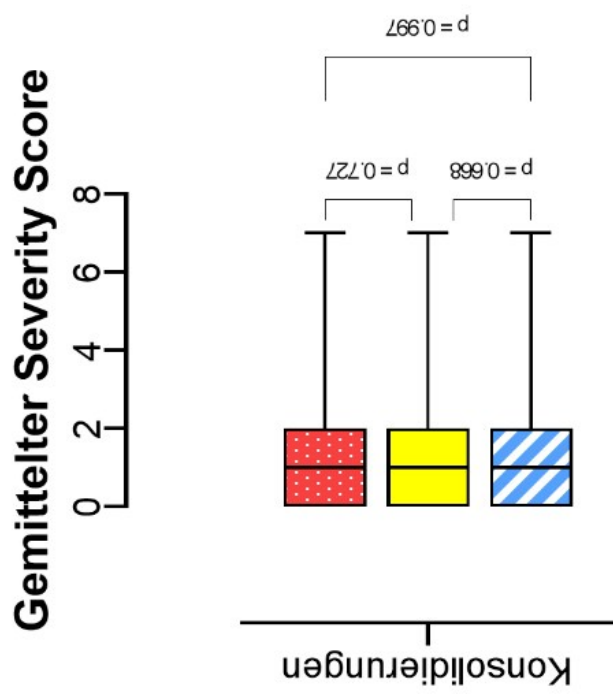
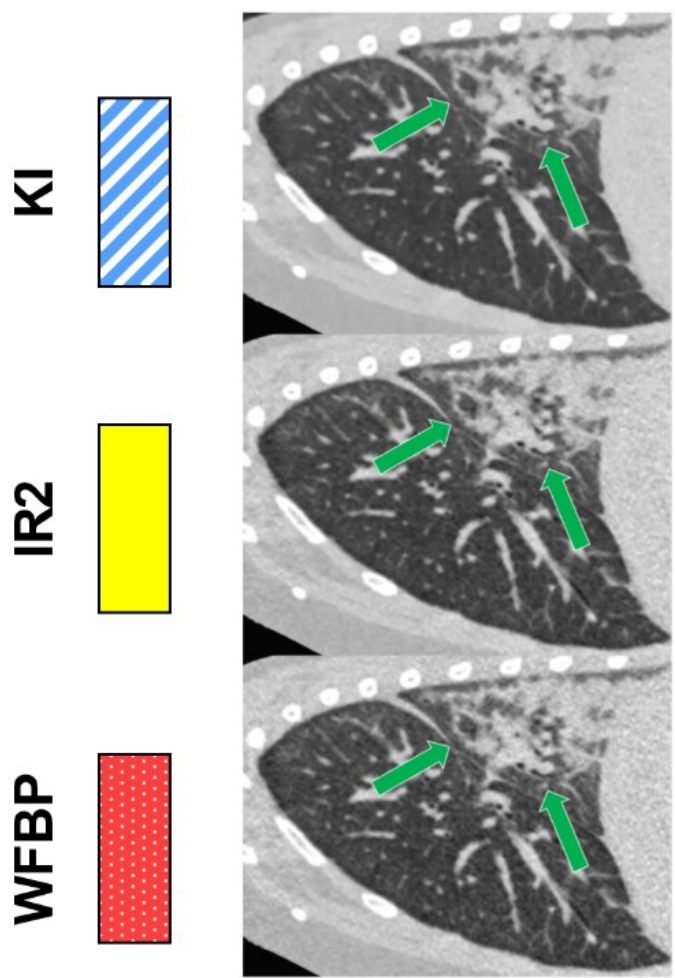


Abbildung 10: Visualisierung der diagnostischen Genauigkeit in Bezug auf Konsolidierungen.²⁷

Abbildung 10 reiht sich in die vorangegangenen Abbildungen ein und zeigt Konsolidierungen. Es erfolgt, wie bereits bekannt, der Vergleich zwischen den verschiedenen Rekonstruktionsmethoden. Der gemittelte Severity Score des Kriteriums Konsolidierungen unterscheidet sich zwischen den verschiedenen Rekonstruktionsmethoden auf nicht statistisch signifikanten Niveau.²⁷

Tabelle 5 zeigt zusammengefasst die einzelnen p-Werte des Severity Scores der einzelnen pathologischen Korrelate im Vergleich von wFBP- und IR2-Aufnahmen zu den KI-Aufnahmen.

	wFBP	IR2
	p im Vergleich zu KI	
Peribronchiales cuffing	> 0,999	0,962
Mucus Plugging	0,996	0,997
Milchglasinfiltrate	0,997	0,668
Air Trapping	> 0,999	0,962
Kavitäten	0,897	0,477
Konsolidierungen	0,997	0,668

Tabelle 5: p-Werte des Severity Scores der einzelnen pathologischen Korrelate im Vergleich zu den KI-Aufnahmen²⁷

3.5 Time to diagnosis

Die gemittelte Time to diagnosis der wFBP-Aufnahmen beträgt $2,66 \pm 2,31$ Minuten. Damit ist die Time to diagnosis der Aufnahmen, die mit der wFBP rekonstruiert werden, am höchsten. Die IR2 Aufnahmen folgten mit einer Time to diagnosis von $2,45 \pm 1,90$ Minuten. Für Aufnahmen, die mit KI rekonstruiert wurden, ermittelt sich eine durchschnittliche Time to diagnosis von $2,28 \pm 1,56$ Minuten. Damit werden KI-Aufnahmen, gemittelt über alle Ratergruppen hinweg, am schnellsten befundet.²⁷

Wenn man die einzelnen Subgruppen betrachtet, zeigte sich folgendes Ergebnis.

Am schnellsten befundeten Fachärzte CT-Aufnahmen. Durchschnittlich benötigte ein Facharzt der Radiologie $1,1 \pm 0,01$ Minuten, um die richtige Diagnose zu finden. Dabei konnte kein statistisch signifikanter Unterschied bezüglich der Time to diagnosis zwischen den einzelnen Rekonstruktionsmethoden festgestellt werden. Gefolgt wurden die Fachärzte von Ärzten in Weiterbildung, die durchschnittlich $3,05 \pm 0,32$ Minuten benötigten, um die richtige Diagnose zu stellen. Medizinstudierende waren am langsamsten bei der Diagnosefindung und benötigten durchschnittlich $6,46 \pm 1,04$ Minuten bis zur Diagnosefindung.²⁷

Besaß ein Bildbetrachter weniger Erfahrung, z.B. ein Medizinstudent oder Arzt im ersten Weiterbildungsjahr, so war ein statistisch signifikanter Unterschied bezüglich der Time to diagnosis zwischen den einzelnen Rekonstruktionsmethoden festzustellen. Hierbei fanden Medizinstudierende und Ärzte im ersten Weiterbildungsjahr die Diagnose am schnellsten mit den rauschreduzierten KI-Aufnahmen, gefolgt von IR2- und wFBP-Aufnahmen.²⁷

Eine detaillierte Darstellung findet sich in Abbildung 11, wobei die blaue Kurve den KI-Aufnahmen entspricht, die gelbe Kurve den IR2 Aufnahmen und die rote Kurve den wFBP-Aufnahmen entspricht.

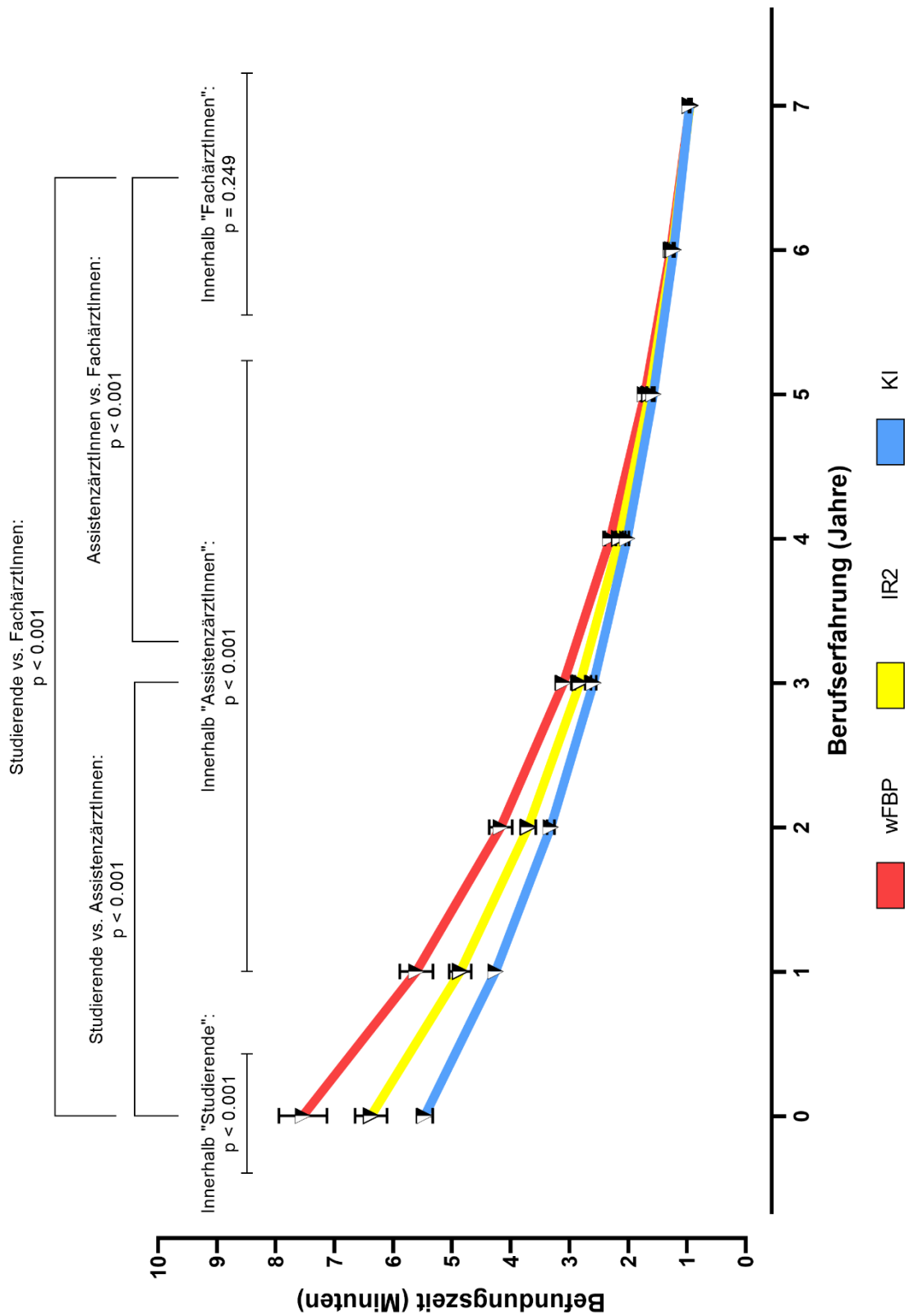


Abbildung 11: Time to diagnosis aufgetragen zur Berufserfahrung mit Aufteilung in die verschiedenen Rekonstruktionsmethoden²⁷

3.6 Phantomversuch

Die KI-Aufnahmen, die aus dem 100 % Scan generiert wurden, hatten ein gemittelt Bildrauschen von $33,7 \pm 1,75$ HU.

Das mittlere Bildrauschen der wFBP-Aufnahmen des 100 % Scans lag bei $76,9 \pm 9,73$ HU. Das Bildrauschen sank durch Dosiserhöhung stetig bis auf ein Niveau von $32,2 \pm 2,97$ bei Verwendung von 600 % Dosis im Vergleich zu einem 100 % Scan ab.

Aufnahmen, die mit IR2 rekonstruiert wurden, starteten mit einem mittleren Bildrauschen von $59,3 \pm 7,5$ bei 100 % Dosis. Durch Dosiserhöhung sank das mittlere Bildrauschen der IR2 Aufnahmen bis auf $24,8 \pm 2,28$ bei 600 % Dosis ab.

Die IR2 Aufnahmen erreichten bei 275 - 375 % Dosis ein mit den 100 % KI-Aufnahmen vergleichbares Bildrauschen. Wurde mit wFBP rekonstruiert, wurde 450 - 600 % Dosis benötigt, um äquivalentes Bildrauschen zur 100 % KI-Aufnahmen zu erhalten.

Die nachfolgende Tabelle zeigt die Zusammenhänge von Dosis und Bildrauschen der einzelnen Aufnahmevarianten. Ebenso finden sich in nachfolgender Tabelle die p-Werte der Mixed-Effekt Analyse. Abbildung 12 verdeutlicht den beschriebenen Zusammenhang bildlich.

		Vergleichswert		Mittelwert \pm SD	
		KI		33,7 \pm 1,75	
		wFBP		IR2	
		p im Vergleich zu 100 % KI	Mittelwert \pm SD	p im Vergleich zu 100 % KI	Mittelwert \pm SD
	100 %				
	Dosis				
	100%	<0.001	76,9 \pm 9,73	<0.001	59,3 \pm 7,5
	125%	<0.001	69 \pm 8,44	<0.001	53,2 \pm 6,5
	150%	<0.001	63,2 \pm 7,51	<0.001	48,7 \pm 5,78
	175%	<0.001	58,6 \pm 6,78	<0.001	45,1 \pm 5,23
	200%	<0.001	54,9 \pm 6,21	<0.001	42,3 \pm 4,79
	225%	<0.001	51,9 \pm 5,76	<0.001	40 \pm 4,43
	250%	<0.001	49,3 \pm 5,37	0.009	37,9 \pm 4,13
IR2 <i>vergleichbar mit KI</i>	275%	<0.001	47,1 \pm 5,03	>0.999	36,2 \pm 3,89
	300%	<0.001	45,1 \pm 4,76	>0.999	34,7 \pm 3,66
	325%	<0.001	43,4 \pm 4,51	>0.999	33,4 \pm 3,47
	350%	<0.001	41,9 \pm 4,29	>0.999	32,2 \pm 3,3
	375%	<0.001	40,5 \pm 4,09	0.283	31,2 \pm 3,15
	400%	<0.001	39,2 \pm 3,92	0.002	30,2 \pm 3,02
	425%	0.001	38,1 \pm 3,76	<0.001	29,3 \pm 2,9
wFBP <i>vergleichbar mit KI</i>	450%	0.051	37,1 \pm 3,62	<0.001	28,5 \pm 2,79
	475%	>0.999	36,1 \pm 3,49	<0.001	27,8 \pm 2,68
	500%	>0.999	35,2 \pm 3,37	<0.001	27,1 \pm 2,59
	525%	>0.999	34,4 \pm 3,25	<0.001	26,5 \pm 2,5
	550%	>0.999	33,6 \pm 3,15	<0.001	25,9 \pm 2,42
	575%	>0.999	32,9 \pm 3,05	<0.001	25,3 \pm 2,35
	600%	>0.999	32,2 \pm 2,97	<0.001	24,8 \pm 2,28

Tabelle 6: Ergebnisse der Mixed-Effekt Analyse bei Vergleich des Bildrauschens der verschiedenen Phantoms cans mit unterschiedlicher Strahlendosis im Vergleich mit den 100 % KI-Aufnahmen

Phantom Scan: Rauschen

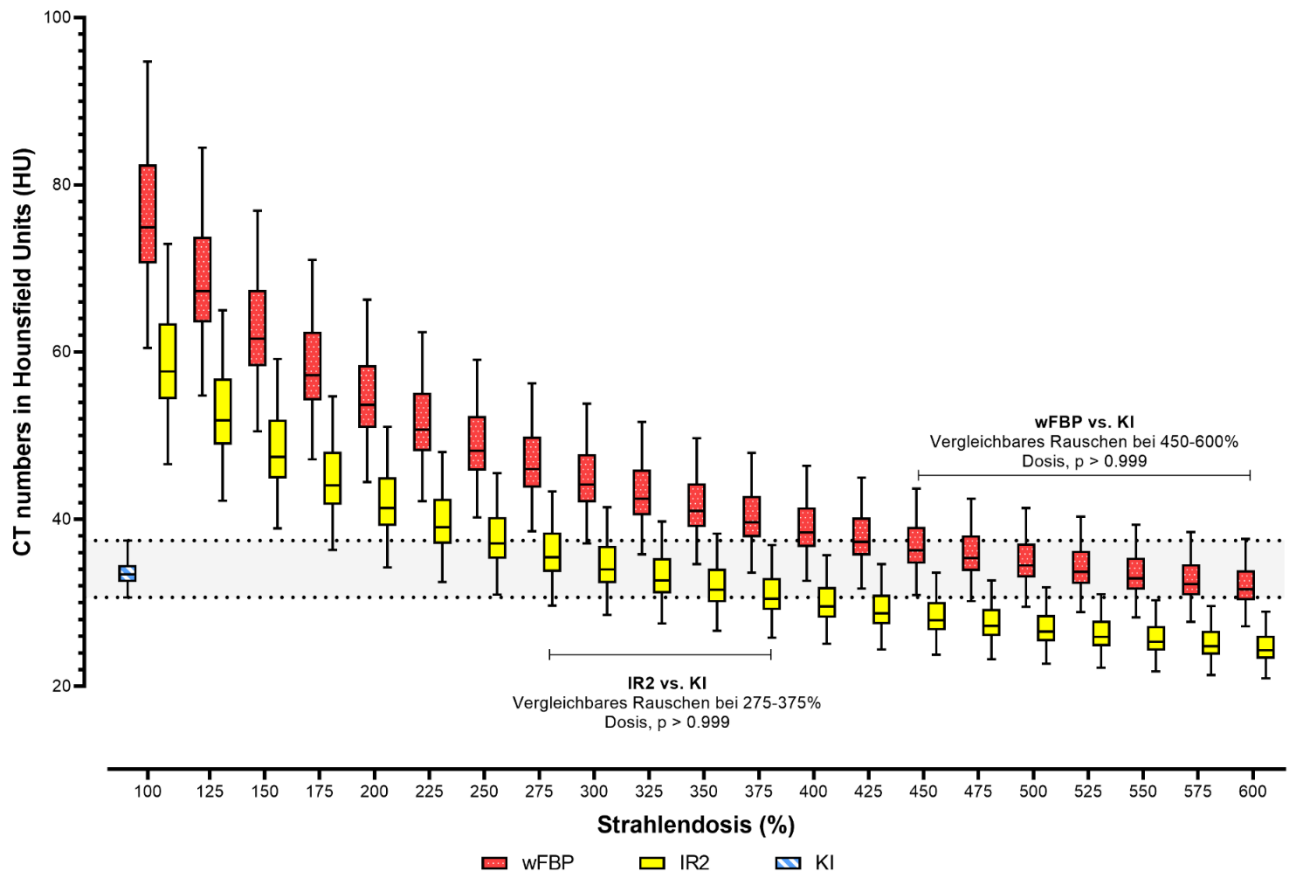


Abbildung 12: Darstellung des mittleren Bildrauschens der Phantomaufnahmen bei unterschiedlicher Strahlendosis und Rekonstruktionsmethode im Vergleich zu den 100 % KI-Aufnahmen

4 Diskussion

4.1 Objektive Bildqualität

Die CT ist bei pädiatrischen Fragestellungen zum Ausschluss einer Pneumonie ein häufig verwendetes Verfahren, da sie eine hohe diagnostische Genauigkeit aufweist und mit geringem Zeitaufwand durchzuführen ist. Jedoch ist ein CT-Scan mit einer relevanten Dosis-Exposition verbunden. Da die Langzeitfolgen von applizierter Strahlung zur Bildgebung, besonders bei Kindern, nicht abschließend geklärt ist, bestehen Bemühungen die Strahlendosis durch ULD-Protokolle zu minimieren. Auch die Anwendung von Spectral Shaping gehört

dazu. Dabei leidet jedoch die Bildqualität. Es besteht jedoch die Möglichkeit, CT-Aufnahmen mit unterschiedlichen Algorithmen zu rekonstruieren. In dieser Studie wurden verschiedene Algorithmen verglichen.

Bereits Suntharalingam et al. 2018 beschrieben eine Verschlechterung der objektiven Bildqualität aufgrund von höherem Bildrauschen in CT-Aufnahmen, die Spectral Shaping verwendeten.²⁹ Die Studie setzte jedoch keinen Schwerpunkt auf die Nachbearbeitung der Aufnahmen. Wetzl et al. 2020 führten Versuche mit Spectral Shaping CT-Aufnahmen mittels iterativem Rekonstruktionsalgorithmus durch und konnten so eine Steigerung der Bildqualität beobachten.³⁰ In der von uns durchgeführten Studie kann der KI-basierte Algorithmus die Bildqualität nochmals im Vergleich zu den iterativ rekonstruierten CT-Aufnahmen mittels IR2 steigern. Dabei wird bei den KI-Aufnahmen das Bildrauschen stärker reduziert als bei der Nachbearbeitung mittels eines wFBP oder iterativen Rekonstruktionsalgorithmus. Die objektive Bildqualität der KI-Aufnahmen ist im Vergleich somit am höchsten zu bewerten.²⁷

4.2 Subjektive Bildqualität

Die Auswirkungen auf die subjektive Bildqualität von CT-Aufnahmen durch den Einsatz von KI basierten Rekonstruktionsalgorithmen war bereits Bestandteil anderer Studien (z.B. Brendlin et al. 2022).²⁸ Vorangegangene Studien beschäftigten sich jedoch hauptsächlich mit der Fragestellung, wieviel Dosisreduktion bei einem CT-Scan bei gleichbleibender Bildqualität durch KI-basierte Bildnachbearbeitung zu erreichen ist. Dabei wurden simulierte niedrig Dosis Aufnahmen verwendet, welche anschließend mit KI-basierter Nachbearbeitung auf ursprüngliches Bildrauschen gebracht wurden.²⁸ Manche Studien beobachten im Vergleich zu den ursprünglichen Volldosisaufnahmen eine Verbesserung der objektiven Bildqualität bei Verschlechterung der subjektiven Bildqualität der nachbearbeiteten Niedrigdosis Aufnahmen.²⁸ Andere berichten von Verbesserung der subjektiven Bildqualität durch Nachbearbeitung in Form von Entrauschung der Aufnahmen.²⁸ Diese Ansätze basieren auf Dosisreduktion durch das Strom-Zeit Produkt der Röntgenröhre. Spectral

Shaping Aufnahmen wurden unseres Wissenstands nach noch nicht mit KI-basierten Rekonstruktionsalgorithmen nachbearbeitet und auf Auswirkungen auf die Bildqualität untersucht. In dieser Arbeit wurde die Nachbearbeitung von Spectral Shaping ULD-CT-Aufnahmen durchgeführt.²⁷

Die Durchführung der Studie ergab, dass die subjektive Bildqualität mit Verwendung des KI-basierten Algorithmus im Vergleich zu den wFBP und IR2 Aufnahmen gesteigert werden kann. Die KI-Nachbearbeitung produziert in den Versuchen dieser Arbeit die Aufnahmen mit der höchsten Bildqualität. Es ist davon auszugehen, dass sich die subjektive Bildqualität von Spectral Shaping CT-Aufnahmen durch die KI-Rekonstruktion steigern lässt.²⁷

4.3 Diagnostische Genauigkeit

Bei der Verwendung von KI-basierten Algorithmen in der Bildverarbeitung in der Medizin besteht oftmals die Sorge, dass Bildaufnahmen in einer Weise verändert werden, dass relevante Bildinformationen verloren gehen oder fälschlicherweise verändert werden.³¹ Im ungünstigsten Fall handelt es sich um Bildinformationen, die ausschlaggebend zur Diagnosefindung beitragen könnten. Da eine Veränderung von Bildmaterial unter Umständen gravierende Auswirkungen auf weiteres klinisches Handeln haben kann, gilt es für die jeweilige gewünschte klinische Fragestellung zu klären, ob die Anwendung eines Rekonstruktionsalgorithmus als unbedenklich einzustufen ist. Dies erfolgte in der Studie, in der die diagnostische Genauigkeit des KI basierten Algorithmus mit herkömmlichen Bildrekonstruktionsverfahren verglichen wurde.

Die diagnostische Genauigkeit ist bei den KI-Aufnahmen unverändert hoch und ist als gleichwertig zu den wFBP und IR2 Aufnahmen anzusehen. Effekte, wie sie Yang et al. 2017 beschrieben, dass durch die Verwendung einer KI-gestützten Rauschreduzierung eine bessere Läsionsabgrenzung erreicht werden könnte,³² haben sich in unserem Studiendesign nicht nachweisen lassen.²⁷

Es lässt sich sagen, dass die hier getestete KI-gestützten Rauschreduzierung bei der Rekonstruktion von ULD-CT-Thoraxaufnahmen zum Ausschluss von

Pneumonien verwendbar ist, ohne dass sich die diagnostische Genauigkeit im Vergleich zu wFBP oder IR2 rekonstruierten Aufnahmen mindert.²⁷

4.4 Time to diagnosis

Es gibt bereits Studien, die den Einsatz von künstlicher Intelligenz in der Radiologie untersucht haben, ob und wie sich der Workflow durch supportive Maßnahmen, wie z.B. Computer unterstützte Diagnosen, verbessern lässt.³³⁻³⁵

In der hier durchgeführten Arbeit wird nicht durch aktive Unterstützung versucht den Workflow zu erhöhen, sondern es wird untersucht, ob die Time to diagnosis durch die Art der Bildrekonstruktion beeinflusst wird. Weitergehend wird untersucht, ob die Verwendung von einem KI-basierten Bildnachbearbeitungsalgorithmus, der Bildaufnahmen mit höherer Bildqualität generiert als herkömmliche Bildrekonstruktionsverfahren, die Time to diagnosis positiv beeinflussen kann.

Es zeigte sich folgender Zusammenhang. Die Time to diagnosis wird vor allem bei Medizinstudierende und Ärzten in Weiterbildung durch die Rekonstruktionsmethode beeinflusst. Bei radiologischen Fachärzten mit mehrjähriger Berufserfahrung ändert die Bildrekonstruktionsmethode die Time to diagnosis nicht statistisch signifikant nachweisbar.²⁷

Man kann aus diesem Zusammenhang die Hypothese aufstellen, dass die KI-Rekonstruktion von der Diagnose ablenkende Bildelemente reduziert und es so unerfahreneren Bildbetrachtern erleichtert, die richtige Diagnose zu finden.

Gemittelt über alle Betrachtergruppen ist die Time to diagnosis bei den KI-Aufnahmen am geringsten. Der Algorithmus wirkt sich positiv auf den Workflow aus.²⁷

4.5 Phantomversuch

In diesem Teil der Arbeit wurde versucht herauszufinden, in welchem Ausmaß der KI-basierte Algorithmus eine Entrauschung an CT-Aufnahmen durchführen kann. Dazu dienten CT-Scans von einem Phantomkörper. Als Vergleichsaufnahmen zu den KI-Aufnahmen haben CT-Scans gleicher und höherer Dosis mit anderen Rekonstruktionsalgorithmen gedient. Brendlin et al. 2022 konnten bereits zeigen, dass die Nachbearbeitung von CT-Aufnahmen mit KI bei Volldosisprotokollen die Bildqualität steigern kann. Ebenso wurde aufgezeigt, dass aus simulierten Niedrigdosis-CT-Aufnahmen, erstellt aus Volldosisaufnahmen, die hohe Bildqualität von Volldosisaufnahmen durch Bildnachbearbeitung mit KI nahezu wieder erreicht werden kann.²⁸

In dieser Arbeit erfolgte die Untersuchung an CT-Aufnahmen, die mittels ULD-CT-Protokollen erstellt wurden.

Der allseits bekannte Zusammenhang zwischen Dosis und Bildrauschen ist deutlich in den Scans des Phantommesskörpers zu erkennen. Auch die Art der Rekonstruktionsmethode wirkt sich auf das Bildrauschen verschieden aus.

IR2 Aufnahmen benötigen im Vergleich zu den 100 % KI-Aufnahmen ca. dreimal so viel Dosis, um ein vergleichbares Bildrauschen vorzuweisen. Beim Vergleich zwischen KI und wFBP-Aufnahmen ist der Unterschied größer. Wird mit wFBP rekonstruiert, wird ca. 5-mal so viel Dosis benötigt, um äquivalentes Bildrauschen, wie in den 100 % KI-Aufnahmen zu erhalten.

Der KI-basierte Algorithmus kann das Bildrauschen in CT-Aufnahmen stärker reduzieren als herkömmliche Bildrekonstruktionsmethoden. Je nach Anwendung lässt sich der Ansatz verfolgen die Bildqualität bei gleicher Dosis durch eine KI-basierte Bildrekonstruktion zu erhöhen oder die benötigte Dosis bei gleichbleibender Bildqualität zu senken. Je nach Anwendung und verwendetem Aufnahmeprotokoll kann die Dosis nicht weiter gesenkt werden, sodass in diesem Falle von höherer Bildqualität durch Rauschreduzierung profitiert werden kann. Kann an Strahlendosis bei einem Aufnahmeprotokoll gespart werden, so kann die Bildqualität durch den KI-basierten Algorithmus erhöht werden.

4.6 Limitationen

Die Studie verwendet ein retrospektives Design. Trotz einer vorher durchgeführten a-priori-Power-Analyse zur Bestimmung der Anzahl von nötigen Versuchspersonen, ist nur eine begrenzte Übertragbarkeit durch retrospektiv gewonnene Ergebnisse auf künftige klinische Fälle möglich.

Des Weiteren wurden nur pädiatrische CT-Aufnahmen mit Verdacht auf Pneumonie ausgewertet. KI-gestützte Algorithmen sollten mit Bedacht auf die jeweilige Fragestellung eingesetzt werden.

Zudem ist in Bezug auf den Phantomversuch hinzuzufügen, dass eine Übertragbarkeit der Ergebnisse von den Messungen am Phantomkörper auf einen realen Patienten unter Umständen nicht oder nur eingeschränkt möglich ist.

Zuletzt muss erwähnt werden, dass alle CT-Aufnahmen an einem modernen CT-Scanner der dritten Generation erstellt wurden. Die Ergebnisse sind auf das verwendete Setup spezifisch.

4.7 Schlussfolgerung

KI-Aufnahmen konnten in allen untersuchten Punkten herkömmliche Bildrekonstruktionsverfahren überbieten oder waren ebenbürtig. Dies zeigt das Potential neuartiger KI-basierter Ansätze zur Bildnachbearbeitung im medizinischen Bereich. Wie in bereits mehreren hier zitierten Artikeln erwähnt, ist der Einsatz KI-basierter Software in der Bildnachbearbeitung immer kritisch zu hinterfragen, bevor eine klinische Anwendung stattfindet. Es gilt das vorhandene Potential neuartiger Bildnachbearbeitungsmethoden zu nutzen, aber dabei sollte stets ein bedachtes Vorgehen gewählt werden, ohne sich von guten Bildeindrücken nachbearbeiteter Bilder täuschen zu lassen.

5 Zusammenfassung

Die durchgeführte Promotionsarbeit hat den Einfluss eines KI-basierten Entrauschungsalgorithmus auf ULD-CT-Aufnahmen des pädiatrischen Thorax untersucht und dabei das mögliche Potential eines solchen Algorithmus im Nachbearbeitungsprozess von ULD-CT-Aufnahmen demonstriert und mit herkömmlichen Bildrekonstruktionsalgorithmen aus der CT-Bildgebung, die einen analytischen Ansatz besitzen und im heutigen klinischen Alltag etabliert sind, verglichen.

In allen untersuchten Bereichen zeigte sich eine Überlegenheit oder mindestens eine Ebenbürtigkeit der KI-basierten Bildverarbeitung im Vergleich zu den analytischen Bildrekonstruktionsverfahren. Im Bereich objektive sowie subjektive Bildqualität erzielten die KI-nachbearbeiteten Aufnahmen die besten Ergebnisse. Auch die diagnostische Genauigkeit der KI-nachbearbeiteten CT-Aufnahmen stellte sich vergleichbar zu herkömmlichen Rekonstruktions- und Nachbearbeitungsverfahren dar. Ferner profitierten weniger erfahrene Bildbetrachter, wie Medizinstudierende oder Assistenzärzte am Anfang ihres beruflichen Werdegangs, von einer kürzeren Time to diagnosis durch die KI-basierte Nachverarbeitung.

Im Versuchsabschnitt des Phantomscans zeigte sich das Potential der KI-basierten Entrauschung. Das Bildrauschen konnte durch die Verwendung der KI-basierten Entrauschung stark gemindert werden. Moderne iterative Rekonstruktionsverfahren benötigten am Phantomkörper ungefähr die dreifache Strahlendosis, um äquivalentes Bildrauschen wie in den KI-Aufnahmen zu erreichen.

Zusammenfassend lässt sich sagen, dass die Verwendung von KI in der Nachbearbeitung für ULD-CT-Aufnahmen des pädiatrischen Thorax Verbesserungen in der Bildqualität und vergleichbare Sicherheit in der Diagnosefindung in dem in der Studie verwendeten Setting mit sich brachte. Einer klinischen Anwendung der KI-Aufnahmen in der Konstellation aus Scanner, Bilderhebungsparametern und Bildnachbearbeitung, wie in der durchgeführten Studie verwendet, steht angesichts der Ergebnisse nichts im Wege. Weitere

Anwendungen von KI-basierter Bildnachbearbeitung in der CT wird Teil zukünftiger Forschung sein. Aufgrund des großen Potentials gilt es, die Entwicklungen gut zu beobachten sowie voranzutreiben.

6 Literaturverzeichnis

Bücher

2. M. R, F. K, Debus J. (2017): *Duale Reihe Radiologie*. Stuttgart: Thieme
8. Ulzheimer S, Bongers M, Flohr T. (2019): Multi-slice CT: Current Technology and Future Developments. In: Nikolaou K, Bamberg F, Laghi A, Rubin GD, eds. *Multislice CT*. Cham: Springer International Publishing; Seiten: 3-34.
13. Alkadhi H, Leschka S, Stolzmann P, Scheffel H. (2011): *Wie funktioniert CT?* Heidelberg: Springer Berlin
17. Grunert J-H. (2019): *Strahlenschutz für Röntgendiagnostik und Computertomografie: Grundkurs und Spezialkurse*. Springer-Verlag

Artikel

1. Trugenberger-Schnabel A, Merchel D, Lönke-Reinl A, Peter J. *Umweltradioaktivität und Strahlenbelastung: Jahresbericht 2019*. 2022. Accessed 06.10.2022. <http://nbn-resolving.de/urn:nbn:de:0221-2022041232235>
3. Hahn C, Siepenkötter C, Tewes S. *Orientierungshilfe für bildgebende Verfahren*. 2019. 978-3-943422-51-1.
4. Strahlenschutzkommission. Bildgebende Diagnostik beim Kind – Strahlenschutz, Rechtfertigung und Effektivität. *Band 61: Empfehlungen und Stellungnahmen der Strahlenschutzkommission 2006*. 2006;
5. Alzen G, Benz-Bohm G. Radiation protection in pediatric radiology. *Deutsches Ärzteblatt*. 2011;108(24):407-14.
6. Staatz GZ, Fred. Neue Entwicklungen in der pädiatrischen bildgebenden Untersuchung. *Monatsschrift Kinderheilkunde*. 2020;168(5):393-394.
7. M. Rose JL, M. Barker. Management der ambulant erworbenen Pneumonie bei Kindern und Jugendlichen (pädiatrische ambulant erworbene Pneumonie, pCAP). *AWMF online*. 2017;
9. Johnston JH, Podberesky DJ, Yoshizumi TT, et al. Comparison of radiation dose estimates, image noise, and scan duration in pediatric body imaging for volumetric and helical modes on 320-detector CT and helical mode on 64-detector CT. *Pediatric Radiology*. 2013;43(9):1117-1127.

10. Tsiflikas I, Thomas C, Ketelsen D, et al. High-Pitch Computed Tomography of the Lung in Pediatric Patients: An Intraindividual Comparison of Image Quality and Radiation Dose to Conventional 64-MDCT. High-pitch Computertomografie der Lunge bei Kindern: Ein intraindividueller Vergleich der Bildqualität und Strahlendosis zur konventionellen 64-MDCT. *roefo*. 2014;186(06):585-590.
11. Flohr T. CT Systems. *Current Radiology Reports*. 2013;1(1):52-63.
12. SiemensHealthineers. Photon-counting CT - A quantum leap in computed tomography. Accessed 25.08.2022, <https://www.siemens-healthineers.com/de/computed-tomography/technologies-and-innovations/photon-counting-ct>
14. Zenger I. Die Geschichte der Computertomographie bei Siemens Healthineers. Siemens Healthcare GmbH; 2021.
15. McLeavy CM, Chunara MH, Gravell RJ, et al. The future of CT: deep learning reconstruction. *Clin Radiol*. 2021;76(6):407-415.
16. Thibault J-B, Sauer KD, Bouman CA, Hsieh J. A three-dimensional statistical approach to improved image quality for multislice helical CT. *Medical Physics*. 2007;34(11):4526-4544.
18. Frasch G, Kammerer L, Karofsky R, Schlosser A, Spiesl J, Stegemann R. *Die berufliche Strahlenexposition des fliegenden Personals in Deutschland 2004 -2009 : Bericht des Strahlenschutzregisters*. 2011. Accessed 06.10.2022. <http://nbn-resolving.de/urn:nbn:de:0221-201108016029>
19. Hendee WR, Marc Edwards F. ALARA and an integrated approach to radiation protection. *Seminars in Nuclear Medicine*. 1986;16(2):142-150.
20. Mulkens TH, Bellinck P, Baeyaert M, et al. Use of an Automatic Exposure Control Mechanism for Dose Optimization in Multi-Detector Row CT Examinations: Clinical Evaluation. *Radiology*. 2005;237(1):213-223.
21. La Riviere PJ, Crawford CR. From EMI to AI: a brief history of commercial CT reconstruction algorithms. *J Med Imaging (Bellingham)*. 2021;8(5):052111.
22. Mazurowski MA, Buda M, Saha A, Bashir MR. Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on MRI. *J Magn Reson Imaging*. 2019;49(4):939-954.
23. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*. 1998;86(11):2278-2324.
24. Deng L, Yu D. Deep Learning: Methods and Applications. *Found Trends Signal Process*. 2014;7(3-4):197-387.

25. Hoeschen C. Einsatz künstlicher Intelligenz für die Bildrekonstruktion. *Der Radiologe*. 2020;60(1):15-23.
26. Umehara K, Ota J, Ishida T. Application of Super-Resolution Convolutional Neural Network for Enhancing Image Resolution in Chest CT. *Journal of Digital Imaging*. 2018;31(4):441-450.
27. Brendlin A, Schmid U, Plajer D, et al. AI Denoising Improves Image Quality and Radiological Workflows in Pediatric Ultra-Low-Dose Thorax Computed Tomography Scans. *Tomography*. 2022;8(4):1678-1689.
28. Brendlin AS, Plajer D, Chaika M, et al. AI Denoising Significantly Improves Image Quality in Whole-Body Low-Dose Computed Tomography Staging. *Diagnostics (Basel)*. 2022;12(1)
29. Suntharalingam S, Mikat C, Wetter A, et al. Whole-body ultra-low dose CT using spectral shaping for detection of osteolytic lesion in multiple myeloma. *European Radiology*. 2018;28(6):2273-2280.
30. Wetzl M, May MS, Weinmann D, et al. Dual-source computed tomography of the lung with spectral shaping and advanced iterative reconstruction: potential for maximum radiation dose reduction. *Pediatr Radiol*. 2020;50(9):1240-1248.
31. McCollough CH, Leng S. Use of artificial intelligence in computed tomography dose optimisation. *Ann ICRP*. 2020;49(1_suppl):113-125.
32. Yang Q, Yan P, Kalra MK, Wang G. CT image denoising with perceptive deep neural networks. *arXiv preprint arXiv:170207019*. 2017;
33. Dikici E, Bigelow M, Prevedello L, White R, Erdal B. Integrating AI into radiology workflow: levels of research, production, and feedback maturity. *Journal of Medical Imaging*. 2020;7(1):016502.
34. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nature Reviews Cancer*. 2018;18(8):500-510.
35. Kapoor N, Lacson R, Khorasani R. Workflow Applications of Artificial Intelligence in Radiology and an Overview of Available Tools. *Journal of the American College of Radiology*. 2020;17(11):1363-1370.

7 Erklärung zum Eigenanteil

Die Arbeit wurde in der Radiologischen Universitätsklinik in der Abteilung für Diagnostische und Interventionelle Radiologie unter Betreuung von PD Dr. med. Ilias Tsiflikas durchgeführt.

Die Konzeptionierung der Studie erfolgte durch PD Dr. med. Andreas Brendlin, PD Dr. med. Saif Afat und PD Dr. med. Ilias Tsiflikas.

Die Datenkuration erfolgte durch mich.

Die statistische Auswertung erfolgte in Zusammenarbeit von PD Dr. med. Andreas Brendlin und mir.

Die Versuche der Abschnitte subjektive Bildqualität, diagnostische Genauigkeit und Time to diagnosis wurden in Zusammenarbeit von PD Dr. med. Andreas Brendlin, Dr. med. David Plajer, Dr. med. Marianna Chaika, Markus Mader, Dr. med. Robin Wrazidlo, Dr. med. Jakob Spogis, Dr. med. Dr. med. univ. Arne Estler, PD Dr. med. Michael Esser und mir durchgeführt.

Die Versuche des Abschnitts objektive Bildqualität und des Phantomversuchs wurden durch PD Dr. med. Andreas Brendlin und mir durchgeführt.

Ich versichere das Manuskript selbständig nach Anleitung durch PD Dr. med. Ilias Tsiflikas und PD Dr. med. Andreas Brendlin verfasst zu haben und keine weiteren als die von mir angegebenen Quellen verwendet zu haben.

Tübingen, den 24.11.2024

Ulrich Schmid

8 Veröffentlichungen

Die Abschnitte „objektive Bildqualität“, „subjektive Bildqualität“, „diagnostische Genauigkeit“ und „Time to diagnosis“ wurden inhaltlich bereits im Journal „Tomography“ im Jahr 2022 veröffentlicht und sind in der Publikation mit dem Titel „AI Denoising Improves Image Quality and Radiological Workflows in Pediatric Ultra-Low-Dose Thorax Computed Tomography Scans“ zu finden.²⁷

Die Beiträge der beteiligten Autoren setzten sich wie folgt zusammen:

Konzeptionierung: Brendlin, Afat und Tsiflikas;

Datenkuration: Schmid;

Statistische Auswertung: Brendlin und Schmid;

Durchführung der Experimente: Brendlin, Schmid, Plajer, Chaika, Mader, Wradzidlo, Spogis, Estler und Esser;

Methoden: Brendlin, Männlin und Afat;

Projektverwaltung: Tsiflikas;

Ressourcen: Schäfer, Afat und Tsiflikas;

Software: Brendlin und Tsiflikas;

Supervision: Afat;

Validierung: Brendlin und Tsiflikas;

Grafiken: Brendlin;

Verfassung des Manuskripts: Brendlin;

Review und Bearbeitung des Manuskripts: Schmid, Plajer, Chaika, Mader, Wradzidlo, Männlin, Spogis, Estler, Esser, Afat und Tsiflikas

9 Danksagung

Hiermit möchte ich mich bei meinem Doktorvater Herrn PD Dr. Ilias Tsiflikas und Betreuer PD Dr. Andreas Brendlin für die Überlassung des Promotionsthemas meiner Dissertation, die tatkräftige Unterstützung und hervorragende Betreuung in allen Phasen der Promotion bedanken.

Abschließend danke ich meinen Eltern und meiner Familie für die Unterstützung während meiner Dissertation und des gesamten Studiums.