
Deep-Learning-basierte Materialbestimmung für die industrielle Computertomographie

der Fakultät für Elektrotechnik, Informationstechnik und Medientechnik
der Bergischen Universität Wuppertal
vorgelegte

Dissertation

zur Erlangung des akademischen Grades
eines Doktors der Ingenieurwissenschaften

von
M. Sc. Moritz Weiß
aus
Essen

Wuppertal 2025

DANKSAGUNG

Mein Dank richtet sich zunächst an meine Familie, meine Freunde und meine Kollegen, die mich auf dem Weg zu dieser Dissertationsschrift bestärkt und unterstützt haben.

Besonderer Dank gilt Prof. Dr.-Ing. Tobias Meisen, der meine Forschung in den letzten Jahren betreut hat. Neben seiner fachlichen Kompetenz, die stets ein wertvoller Input für meine Forschung war, ist es ihm gelungen, meine Ideen in strukturierte Bahnen zu leiten, sodass ich weit über meine eigenen Grenzen hinauswachsen konnte.

Weiterhin danke ich Dr. Mirko von Schmid, der stets ein offenes Ohr für meine Fragen hatte. Ohne seine Hilfe hätte mein fachlicher Einstieg in die Welt der Computertomographie viele zusätzliche Jahre in Anspruch genommen.

Abschließend danke ich meiner Partnerin Inga, die mit unerschöpflicher Geduld und Verständnis über die vergangenen Jahre an meiner Seite stand. Ich widme ihr diese Arbeit:

Für Inga.

ABSTRACT

Die industrielle Computertomographie (CT) ist eine zerstörungsfreie Prüftechnik, die dreidimensionale, digitale Abbilder, die sogenannten CT-Volumen, von Prüfteilen mithilfe von Röntgenstrahlung erzeugt. Diese Dissertation untersucht den Einsatz von Deep Learning (DL) zur Bestimmung der Kernladungszahl und Dichte von Prüfteilen anhand von industriellen CT-Scans. Mit dem so erlangten Wissen über die Materialzusammensetzung eines Prüfteils können vielfältige Prüfszenarien, insbesondere an sicherheitskritischen Prüfteilen, realisiert sowie der Recyclingprozess verbessert werden.

Der Einsatz eines DL-basierten Ansatzes verlangt neben einer passenden DL-Architektur Trainingsdaten, die das zu lernende Problem, die Materialbestimmung, hinreichend gut darstellen. Somit besteht ein zentraler Beitrag dieser Dissertation in der Entwicklung einer anforderungsgerechten CT-Simulation zur Erzeugung annotierter Datensätze. Weiterhin wird in dieser Dissertation erstmalig eine Convolution-basierte Architektur mit einer Transformer-basierten Architektur für die Materialbestimmung anhand von CT-Volumen quantitativ verglichen. Die wichtigsten Einflussfaktoren, die bei dem Transfer der entwickelten Methode zur Materialbestimmung in die reale CT-Anwendung auftreten, werden identifiziert und quantitativ analysiert. Anhand eines realen CT-Scans einer Lithium-Ionen-Batterie wird gezeigt, dass die entwickelte Methode geeignet ist, um sowohl in Bezug auf Präzision als auch Laufzeit in der in-line Fertigung von Lithium-Ionen-Batterien eingesetzt zu werden. Abschließend werden abseits der Materialbestimmung weitere Anwendungspotentiale aus der entwickelten Methodik abgeleitet und exemplarisch erprobt.

INHALTSVERZEICHNIS

1	Einleitung	1
1.1	Motivation	3
1.2	Forschungsfragen	5
1.3	Struktur der Thesis	6
2	Grundlagen der industriellen Computertomographie	9
2.1	Wechselwirkung von Photonen mit Materie	10
2.2	Quelle: Erzeugung von Röntgenstrahlung	14
2.3	Detektor: Quantitative Detektion von Röntgenstrahlung	16
2.3.1	Bildkorrekturen	17
2.3.2	Bildrauschen	18
2.4	Rekonstruktion	18
2.4.1	CT-Artefakte	20
2.5	Trajektorien	22
2.6	Dual-Energy CT und klassische Materialzerlegung	23
2.7	Vorstellung reale CT-Anlage	25
3	Grundlagen des Deep Learnings für die Computertomographie	27
3.1	Einführung: Deep Learning	27
3.2	Training neuronaler Netze	29
3.3	Beurteilung des Trainings	31
3.3.1	Weitere Verbesserungen des Trainings	33
3.4	Deep Learning für die Bildverarbeitung	34
3.4.1	U-Net	35
3.4.2	Swin Transformer	36
4	Verwandte Forschung	39
4.1	Methodik der Literaturrecherche	40
4.2	Analyse des Korpus	43
4.2.1	Beschleunigungsspannung und Detektortechnologie	44
4.2.2	Materialien	46
4.2.3	Klassische Ansätze zur Materialbestimmung	47
4.2.4	Datengetriebene Ansätze	48
4.3	Erkenntnisgewinn aus der Literatur	55
4.3.1	Erzeugung der Trainingsdaten	55
4.3.2	Architekturen	55
5	Simulation der Computertomographie	57
5.1	Anforderungsanalyse: CT-Simulation	57
5.1.1	CT-spezifische Anforderungen	58
5.1.2	Deep-Learning-spezifische Anforderungen	59
5.1.3	Nicht-funktionale Anforderungen	61
5.1.4	Überblick über die Anforderungen	61

5.2	Existierende CT-Simulationen	62
5.2.1	aRTist	63
5.2.2	gVXR	64
5.2.3	Diskussion der Anforderungen	64
5.3	Eigene Simulation	65
5.4	Evaluation der implementierten Simulation	68
5.4.1	Exemplarischer Datensatz & Datenverteilung	68
5.4.2	Vergleich mit aRTist und mit einem realen Scan	72
5.4.3	Beobachtungen mit (simulierter) Dual-Energy CT	77
5.4.4	Geschwindigkeit und Skalierbarkeit	79
5.5	Zusammenfassung	80
6	Deep-Learning-Architektur für die Materialbestimmung	83
6.1	Industrielle Rahmenbedingungen für DL in der CT	83
6.2	Simuliertes Phantom zur quantitativen Auswertung	88
6.3	Einfluss des globalen Bildkontextes	91
6.4	Verschiedene Architekturen: Convolution oder Transformer?	93
6.4.1	U-Net	94
6.4.2	Swin Transformer	98
6.4.3	Vergleich	102
6.5	Inferenz mit dem U-Net	103
6.6	Zusammenfassung der Ergebnisse	104
7	Transfer in die CT-Praxis	105
7.1	Größe der Materialregionen	106
7.2	Kantenschärfe	111
7.3	Größe des simulierten Phantoms	113
7.3.1	Fine-Tuning zur Adaption der Kantenlänge	114
7.4	Relative Dichten	116
7.5	Vorhersagen außerhalb der CT-Mittelschicht	119
7.6	Anwendung auf reale CT-Daten	122
7.7	Zusammenfassung der Ergebnisse	126
8	Anwendungspotential auf andere Problemstellungen in der CT	129
8.1	Reduktion der Strahllaufhärtungsartefakte	129
8.2	Rauschreduktion	132
8.3	Detektion von Poren und Partikeln	133
8.4	Zusammenfassung der Ergebnisse	134
9	Kritische Würdigung	135
10	Zusammenfassung	141

ABKÜRZUNGSVERZEICHNIS

CT	Computertomographie	2
EDX	Energy-Dispersive X-ray Spectroscopy	2
XRF	X-ray Fluorescence	2
DL	Deep Learning	4
SOTA	State-Of-The-Art	4
CV	Computer Vision	4
DECT	Dual-Energy Computed Tomography	10
EID	Energy-Integrating Detector	17
PCD	Photon-Counting Detector	24
MLP	Multi-Layer-Perceptron	27
CNN	Convolutional Neural Network	34
NLP	Natural Language Processing	36
ViT	Vision Transformer	37
GAN	Generative Adversarial Network	51
LIB	Lithium-Ionen-Batterie	123
NMC	Lithium-Nickel-Mangan-Cobalt-Oxid	124

1

EINLEITUNG

Mit der Entdeckung der Röntgenstrahlung im Jahr 1895 und dem damit verbundenen ersten Röntgenbild, das die Hand von Wilhelm Conrad Röntgens Frau zeigt, wurde ein neuartiges Teilgebiet der angewandten Physik eröffnet. Wilhelm Conrad Röntgen veröffentlichte im Jahr 1896 seine Beschreibung einer neuen Strahlenart, die heutzutage als *Röntgenstrahlung* bekannt ist [1]. Für seine Entdeckung erhielt er im Jahr 1901 den ersten Nobelpreis für Physik [2]. Im Vergleich zu sichtbarem Licht bietet Röntgenstrahlung eine stärkere Durchdringung von Materie, sodass Objekte durchleuchtet werden können. In diesen sogenannten *Röntgenprojektionsbildern* (im Folgenden: *Projektionen*) wird die totale Abschwächung der Röntgenstrahlung durch eine Probe ortsaufgelöst gemessen. Sie ist das Produkt aus dem material- und energieabhängigen linearen Abschwächungskoeffizienten $\mu_{L,m}(E)$ ¹ und einer Durchstrahlungslänge x_m . Analog zu sichtbarem Licht kann für Röntgenstrahlung das **LAMBERT-BEER-Gesetz**

$$I(E) = I_0(E) \cdot \exp^{-\mu_{L,m}(E) \cdot x} \quad (1.1)$$

mit dem Spektrum der einfallenden Photonen $I_0(E)$ und dem gedämpften Spektrum $I(E)$ angewandt werden [3, S. 54]. Daraus folgt, dass Objekte aus verschiedenen Materialien mit verschiedenen Durchstrahlungslängen die gleiche Dämpfung aufweisen können und somit anhand einer Röntgenprojektion ununterscheidbar sind. Abbildung 1.1 zeigt die Projektion von zwei Objekten. Das linke Objekt ist ein Würfel aus Aluminium mit einer



Abbildung 1.1: Röntgenprojektion von zwei Objekten unter monochromatischer Strahlung bei 6 MeV. Gleiche Größe, gleiche Graustufe, gleiches Objekt? Links: 10 cm Durchstrahlung durch Aluminium. Rechts: 8.8 mm Durchstrahlung durch Wolfram. Die genannten Materialien wurden aufgrund ihrer stark unterschiedlichen Abschwächung der Röntgenstrahlung gewählt.

Kantenlänge von 10 cm. Das rechte Objekt ist ein 8.8 mm dünnes Wolframeblech. Zunächst wurde die Durchstrahlungsprüfung für medizinische Zwecke genutzt, um beispielsweise im Jahr 1903 ein diagnostisches Verfahren zur Erkennung von Tuberkulose anhand der Zwerchfell-Bewegungen zu realisieren [4]. Wenige Jahre später erkannte Moseley eine Proportionalität zwischen dem Quadrat der (effektiven) Kernladungszahl Z_{eff}^2 eines Elements und der Frequenz der ausgesandten (später: *charakteristischen*) Strahlung ν nach vorheriger Anregung durch Röntgenstrahlung [5]. In den folgenden Jahrzehnten

¹ Die in der Literatur häufig verwendete Notation des energieabhängigen linearen Abschwächungskoeffizienten $\mu_L(E)$ wird im Folgenden mit einem Materialindex m für die Notation von Multimaterialien ergänzt.

mündete diese Erkenntnis in die Entwicklung elementspezifischer Analyseverfahren wie X-ray Fluorescence (XRF) und Energy-Dispersive X-ray Spectroscopy (EDX), die beide auf der ausgesandten, charakteristischen Röntgenstrahlung der Elemente beruhen. Aufbauend auf der bereits bekannten Beugung von Licht an Kristallgittern zeigte Max von Laue, dass Röntgenstrahlung ebenfalls Beugungsmuster an Kristallen hervorbringt. Hiermit bewies er, dass Röntgenstrahlung als Welle dargestellt werden kann und erhielt im Jahr 1914 den Nobelpreis [6]. Basierend auf der Röntgenbeugung wurde das Gebiet der Kristallstrukturanalyse (englisch: *X-ray Crystallography*) begründet, wodurch im Jahr 1953 die molekulare Struktur der DNA entdeckt wurde [7]. Der nächste technologische Durchbruch mithilfe von Röntgenstrahlung gelang Sir Godfrey Hounsfield im Jahr 1971, indem er den ersten Computertomographiescan durchführte und wenige Jahre später zusammen mit Allan Cormack mit dem Nobelpreis für Medizin ausgezeichnet wurde [8]. Im Gegensatz zur zweidimensionalen Durchstrahlungsprüfung wird bei der Computertomographie (CT) eine Serie an Projektionen aus verschiedenen Blickwinkeln aufgenommen. Vergleichbar mit dem stereoskopischen Sehen wird die Tiefeninformation eines betrachteten Objektes im Kontext der aufgenommenen Bildserie aus verschiedenen Perspektiven verfügbar. Durch Rekombination (im Folgenden: *Rekonstruktion*) der Bilder wird ein dreidimensionales Abbild, ein sogenanntes *Volumen* oder auch *CT-Volumen*, rekonstruiert, anhand dessen die Unterscheidung zwischen den Objekten aus Abbildung 1.1 trivial möglich ist. Abbildung 1.2a zeigt das Rendering des Volumens nach erfolgtem CT-Scan und dessen Rekonstruktion. Der Schnitt durch die horizontal orientierte Mittelschicht ist in Abbildung 1.2b dargestellt. Es ist zu beobachten, dass Aluminium



(a) Rendering des rekonstruierten Volumens

(b) Schnittbild aus der Draufsicht

Abbildung 1.2: Rekonstruiertes Volumen des CT-Scans eines Aluminiumwürfels mit der Kantenlänge 10cm und einem Wolframblech mit einer Dicke von 8.8mm, wobei eine einzelne Projektion in Abbildung 1.1 gezeigt wurde. Dargestellt werden ein Rendering in (a) und ein Schnittbild in horizontaler Ebene in (b) – die Draufsicht.

und Wolfram im Schnittbild unterschiedliche Grauwerte aufweisen. Im Vergleich zu der in Abbildung 1.1 dargestellten Projektion besitzen die Volumenpixeln (auch: *Voxel*) eine physikalische Einheit, die die Dämpfung der Röntgenstrahlung pro Weglänge durch ebendieses Voxel beschreibt. Sie entspricht dem linearen Abschwächungskoeffizienten $\mu_{L,m}(E)$ aus Gleichung 1.1, sodass die CT als ortsaufgelöstes Messinstrument für Abschwächungskoeffizienten der Röntgenstrahlung fungiert.

1.1 MOTIVATION

Jedes chemische Element, beziehungsweise Material m , besitzt eine charakteristische Elektronenkonfiguration² und somit einen charakteristischen Abschwächungskoeffizienten $\mu_{L,m}(E)$, der für die Unterscheidung zwischen verschiedenen Materialien herangezogen werden kann. Im Folgenden werden materialauflösende Algorithmen basierend auf CT-Daten (auch: *Material-CT*) wie folgt unterschieden:

- Die **Materialbestimmung** ist die Berechnung einer effektiven Kernladungszahl Z und einer absoluten Dichte ρ (in g/cm³). Beide Größen sind somit Mittelwerte der elementaren Mischung und die zugrundeliegenden Elemente können ohne weiteres Wissen nicht eindeutig bestimmt werden.
- Die **Materialzerlegung** setzt einen festen Satz an Basismaterialien voraus, die aus mathematischer Sicht eine linear unabhängige Basis bilden. Es werden somit relative Anteile der Basismaterialien an einem Voxel berechnet und das Material effektiv in die Basismaterialien zerlegt.

Beide Methoden bieten vielfältige Anwendungen für den industriellen und medizinischen Sektor. Im medizinischen Umfeld wird der gemessene Abschwächungskoeffizient verwendet, um beispielsweise die Knochendichte eines Patienten abzuschätzen [9] oder Lungenkrankheiten durch Gewebeveränderung zu erkennen [10]. Für industrielle Zwecke kann die Materialbestimmung beziehungsweise -zerlegung ebenso vielfältig genutzt werden. Als zerstörungsfreie Prüftechnik für nahezu jedes technisch relevante Material findet die Material-CT beispielsweise Anwendung in der Erkennung gefährlicher oder illegaler Stoffe, die in Containern an Flughäfen, Bahnhöfen oder Häfen transportiert werden [11]. Die Untersuchung von Alterungsprozessen in technischen Materialien, zum Beispiel Beschichtungen in Lithium-Ionen-Batterien, stellt ein weiteres Anwendungsfeld dar [12]. Alle genannten Fälle haben die Gemeinsamkeit, dass die Voxel des Volumens im Bestfall in intuitive und zugängliche Größen, meist eine Dichte ρ und eine effektive Kernladungszahl Z , umgerechnet werden. Algorithmisch wird die materialauflösende CT seit fast 50 Jahren erforscht. Lange Zeit wurden zu diesem Zweck Algorithmen händisch entwickelt, die auf den physikalischen Grundlagen der CT basieren (im Folgenden: *klassische Algorithmen*). Die einflussreichsten Faktoren bei der CT-Bildgebung sind umfassend erforscht und können entsprechend bei der Modellierung eines klassischen Algorithmus berücksichtigt werden. Dennoch besteht eine weitverbreitete Einschränkung klassischer Algorithmen darin, dass sie punktweise (pro Voxel) definiert sind und Kontextinformationen aus benachbarten Voxeln nicht miteinbeziehen [13, 14, 15]. CT-Daten unterliegen aufgrund verschiedener Einflüsse, die in Kapitel 2 genauer ausgeführt werden, den sogenannten *CT-Artefakten*, die zwar lokal erhebliche Abweichungen der rekonstruierten Abschwächungskoeffizienten induzieren können, jedoch unter Einbezug eines größeren

² Im für die meisten CT-Anwendungen relevanten Energiebereich $E < 1 \text{ MeV}$ ist der direkte Einfluss der Atomkerne auf die Abschwächung der Röntgenstrahlung vernachlässigbar. Ausschlaggebend sind die Bindungsenergien der Elektronen sowie die (räumliche) Elektronendichte. In Kapitel 2 werden die zugrundeliegenden Wechselwirkungen für die Entstehung des charakteristischen Abschwächungskoeffizienten detailliert beschrieben.

(Bild-)Kontextbereiches eindeutig als Artefakte identifizierbar und somit potentiell korrigierbar sind. Moderne Ansätze der klassischen Materialzerlegung setzen daher auf ein zweistufiges Verfahren, wobei nach der klassischen Zerlegung ein Cluster-Verfahren angewandt wird, um Artefakte in semantisch zusammenhängenden Regionen herauszufiltern [16].

Formell betrachtet ist die Materialbestimmung anhand von rekonstruierten Volumen eine sogenannte *Bild-zu-Bild-Berechnung* – eine Aufgabe aus dem Bereich Computer Vision (CV), der in den letzten Jahren durch Methoden basierend auf Deep Learning (DL) deutlich weiterentwickelt wurde. DL-basierte Methoden gelten derzeit als State-Of-The-Art (SOTA) für viele CV-Anwendungen und dominieren die gängigsten Benchmarks in der Objekterkennung [17], der Klassifizierung [18], der Segmentierung [19] und der Segmentierung klinischer CT-Daten [20, 21], sodass traditionelle Verfahren weitestgehend verdrängt worden sind. Ein Teil dieses Erfolges basiert auf dem Konzept des *rezeptiven Feldes* – einer Methode zur Verarbeitung potentiell hilfreicher, räumlicher Kontextinformationen. Der Wandel hin zu DL-basierten Verfahren dringt zunehmend in die CT-Anwendung, wie wir im Jahr 2024 anhand einer stark steigenden Zahl an Publikationen, die DL als Werkzeug für die Material-CT verwenden, zeigen konnten [22]. Grundsätzlich benötigen DL-Verfahren einen vorgegebenen Datensatz, die sogenannten *Trainingsdaten*, der die zu lernenden Zusammenhänge implizit darstellt. Material-CT-Anwendungen werden mit rekonstruierten Volumen oder Schnittbildern aus den rekonstruierten Volumen als Eingabedaten (auch: *Inputs*) und den sogenannten *Materialkarten* als Ausgabedaten (auch: *Ground Truths*) trainiert.³ Die Materialkarten geben bei der Materialbestimmung für jedes Voxel, beziehungsweise auf Bildebene *Pixel*, in den Inputs eine Kernladungszahl Z und eine Dichte ρ an. Zur Erzeugung der Trainingsdaten besteht somit die Möglichkeit, sogenannte *Phantome* aus bekannten Materialien zu fertigen und CT-Scans an einer realen Anlage durchzuführen. Die rekonstruierten Volumen können anschließend händisch annotiert werden (Kernladungszahl und Dichte) und als Trainingsdaten verwendet werden. Eine Alternative bietet die Erzeugung der Trainingsdaten mithilfe einer Simulation. Obwohl beide Ansätze effektiv verwendet werden können [22], bietet die Simulation einige Vorteile. Simulativ erzeugte Phantome sind, unabhängig von ihrer morphologischen Komplexität, per Definition annotiert – das Phantom entspricht den Materialkarten und Annotationsfehler treten nicht auf. Zudem erlaubt die Simulation die zufällige Erzeugung eines breiten Spektrums an Formen und Materialzusammensetzungen, sodass die Varianz der Trainingsdaten enorm gesteigert wird. Der wichtigste Nachteil einer Simulation ist stets, dass sie die reale CT-Messung nicht fehlerfrei nachbildet. Diese Abweichung von der Realität, die sogenannte *Sim-To-Real-Gap*, hat einen Einfluss auf die Transferfähigkeit eines jeden auf simulierten Daten trainierten DL-Modells in die reale Anwendung. Abhängig von der Größe der Sim-To-Real-Gap sinkt die Vorhersagegenauigkeit des Modells auf Realdaten signifikant ab, sodass ihr Einfluss für jede Anwendung quantitativ untersucht werden muss.

³ Verfahren, die Projektionen in sogenannte *Material-Projektionen* übersetzen, werden im Rahmen dieser Dissertationsschrift lediglich ergänzend genannt, da es sich stets um kalibrierte Verfahren mit vielen Einschränkungen handelt. Weitere Argumente für die Wahl zwischen den Datendomänen *Volumen* oder *Projektion* als Input für die Material-CT werden in Kapitel 2 aufgeführt.

Im Rahmen dieser Dissertationsschrift wird die Anwendbarkeit DL-basierter Ansätze für die Materialbestimmung in industriellen CT-Scans evaluiert und vorangetrieben. Zunächst werden die Anforderungen an eine Simulation zur Erzeugung der Trainingsdaten beschrieben und zur Auswahl einer passenden Methodik verwendet. Aufbauend wird ein entsprechend simulierter Datensatz verwendet, um verschiedene DL-Architekturen, die mit der Fragestellung der Material-CT kompatibel sind, hinsichtlich ihrer Eignung für die Materialbestimmung zu evaluieren. Die Transferfähigkeit der trainierten Modelle in die reale Anwendung wird untersucht, um Einschränkungen bedingt durch die Trainingsdatenerzeugung, die DL-Architektur oder die generelle Methodik zu identifizieren und zu beheben. Hierbei bestehen weiterhin offene Fragen, die in der Literatur zur Materialbestimmung mit DL bisher nicht ausreichend untersucht worden sind, obwohl sie für den Transfer in die reale Anwendung von zentraler Bedeutung sind. Abschließend werden die entwickelten Methoden dieser Arbeit auf andere Fragestellungen der CT-Anwendung projiziert, um weitere Anwendungsfelder zu erschließen und um zukünftige Weiterentwicklung zu eröffnen.

1.2 FORSCHUNGSFRAGEN

Im Folgenden werden vier aus der Motivation abgeleitete Forschungsfragen formuliert und erläutert.

Forschungsfrage 1: Wie lässt sich eine Simulation zur Trainingsdatenerzeugung für die Materialbestimmung in der Computertomographie realisieren?

Im ersten Schritt wird eine Anforderungsanalyse an eine Methodik zur Trainingsdatenerzeugung für die DL-basierte Materialbestimmung durchgeführt. Anhand der abgeleiteten Anforderungen werden etablierte CT-Simulationsprogramme bewertet sowie eine Methodik zur Trainingsdatenerzeugung realisiert. Abschließend müssen die mit der ausgewählten Methodik erzeugten Daten quantitativ mit einer der etablierten Simulationen verglichen und mindestens exemplarisch mit den Messdaten einer realen CT-Anlage gegenübergestellt werden, um die Sim-To-Real-Gap abzuschätzen.

Forschungsfrage 2: Welche Deep-Learning-Architekturen eignen sich für die Materialbestimmung in CT-Bildern unter Berücksichtigung industrieller Rahmenbedingungen?

Zunächst werden die industriellen Rahmenbedingungen, die von einer DL-Architektur in Bezug auf Kompatibilität, Laufzeit und Hardwareanforderungen erfüllt werden müssen, abgeleitet. Die Problemstellung der Materialbestimmung ist im Kern eine Bild-zu-Bild-Berechnung beziehungsweise eine Pixel-zu-Pixel-Berechnung (auch: *Dense Prediction*). Es werden bekannte Architekturen für Dense Prediction einerseits aus der Literatur zur Material-CT und andererseits aus anderen CV-Anwendungen implementiert, für die Materialbestimmung optimiert und quantitativ verglichen. Zur quantitativen Bewertung werden die anfangs abgeleiteten industriellen Rahmenbedingungen herangezogen.

Forschungsfrage 3: Welche Einschränkungen ergeben sich für die reale Anwendung aus der gewählten Methodik zur Trainingsdatenerzeugung und der Deep-Learning-Architektur?

Abseits der technischen Machbarkeit spielt die praktische Anwendbarkeit eines Produktes eine zentrale Rolle für die Realisierung der Methode. Die Methodik der Trainingsdatenerzeugung sowie die DL-Architektur treffen, begründet durch die Anforderungsanalyse und die industriellen Rahmenbedingungen, verschiedene Annahmen und Einschränkungen, die einen negativen Einfluss auf die universelle Anwendbarkeit der Modelle in einer praktischen Anwendung haben können. Ziel ist es, ebendiese Einflüsse zunächst zu identifizieren und ihre Auswirkungen auf die absolute Vorhersagegenauigkeit der entwickelten DL-Modelle im Kontext der Materialbestimmung zu quantifizieren. Diese Vorhersagegenauigkeit wird anhand von zuvor beschriebenen Fehlertoleranzen, abgeleitet aus praktischen CT-Anwendungen, jeweils für die Kernladungszahl und die Dichte beurteilt. Weiterhin werden Lösungsansätze diskutiert oder umgesetzt, um die gefundenen Abweichungen zu korrigieren und somit die Vorhersagegenauigkeit bei der Anwendung in realen CT-Scans zu verbessern.

Forschungsfrage 4: Können mit den entwickelten Methoden weitere Anwendungsfelder für Deep Learning in der Computertomographie erschlossen werden?

Die im Rahmen der ersten drei Forschungsfragen entwickelten Methoden und Daten können möglicherweise auf verwandte Fragestellungen der CT angewandt werden. Insbesondere Methoden zur Verbesserung der Bildqualität können bereits durch die gewählte Methodik der Material-CT implizit vom DL-Modell erlernt worden sein. Darüber hinaus werden kleine Weiterentwicklungsmöglichkeiten der Simulation diskutiert, die zur Erzeugung geeigneter Trainingsdaten für andere CT-Fragestellungen verwendet werden können.

1.3 STRUKTUR DER THESIS

Diese Thesis besteht aus zwei Teilen, wobei der erste Teil eine gemeinsame Basis für das Verständnis der Computertomographie legt (Kapitel 2) und Deep Learning (Kapitel 3) als vielseitiges Werkzeug für die Verarbeitung von CT-Daten vorstellt. Im Rahmen der theoretischen Beschreibung der CT wird die Materialbestimmung eingeführt und in den physikalischen Kontext eingebettet. Der Hauptteil befasst sich mit dem Forschungsbeitrag dieser Dissertationsschrift in den Kapitel 4 bis 9.

Als systematischer Einstieg in den Hauptteil fungiert Kapitel 4, das die verwandte Forschung zu dieser Dissertationsschrift zusammenfasst. Besonderer Fokus liegt hierbei auf dem Wandel, der Forschende auch im Umfeld der Computertomographie zunehmend motiviert, Problemstellungen mit Deep Learning zu bearbeiten. Anschließend werden die Anforderungen an eine Simulation zur Trainingsdatenerzeugung, physikalisch motiviert durch Kapitel 2 und ergänzt durch die DL-basierten Anforderungen aus Kapitel 3, im Rahmen der ersten Forschungsfrage in Kapitel 5 abgeleitet. Basierend auf ebendiesen

Anforderungen werden mehrere existierende CT-Simulationsprogramme bewertet sowie die fehlenden Funktionalitäten identifiziert und implementiert. Die mit der entwickelten Methode simulierten CT-Bilder werden mit den CT-Bildern einer realen CT-Anlage verglichen.

In Kapitel 6 werden die industriellen Rahmenbedingungen für die Nutzung eines DL-basierten Ansatzes zur Materialbestimmung abgeleitet und eine passende DL-Architektur gesucht. Die Entscheidung zwischen einer Convolution-basierten und einer Transformer-basierten Architektur stellt den zentralen Bestandteil dieses Kapitels dar.

Aufbauend auf der Trainingsdatenerzeugung sowie der ausgewählten DL-Architektur wird in Kapitel 7 der Transfer in die reale CT-Anwendung untersucht, wobei Limitierungen aus den beiden vorangehenden Forschungsfragen quantitativ eingeordnet werden. Einige der gefundenen Limitierungen können teilweise durch eine angepasste Trainingsstrategie oder eine angepasste CT-Messstrategie reduziert oder vollständig aufgehoben werden.

Abschließend werden weitere Anwendungsmöglichkeiten des entwickelten Verfahrens für offene Fragestellungen der Computertomographie, abseits der Materialbestimmung, in Kapitel 8 identifiziert und diskutiert. Der Fokus liegt auf Anwendungen, die direkt aus der entwickelten Trainingsdatenerzeugung in Kombination mit der ausgewählten DL-Architektur realisiert werden können, ohne komplexe Weiterentwicklungen zu erfordern.

Kapitel 9 reflektiert die Forschungsergebnisse kritisch und stellt mögliche Verbesserungen in Aussicht, die in erster Linie die Trainingsdatenerzeugung in Kapitel 5, die DL-Architektur in Kapitel 6 und den Transfer in die CT-Anwendung in Kapitel 7 betreffen. Eine Zusammenfassung der erzielten Ergebnisse dieser Dissertationsschrift erfolgt in Kapitel 10.

2

GRUNDLAGEN DER INDUSTRIELLEN COMPUTERTOMOGRAPHIE

Abgeleitet vom Begriff *Schnitt* (altgriechisch: *tomē*) handelt es sich bei der Computertomographie (CT) um ein Schnittbild-Verfahren, das einzelne Schichten eines Objektes sichtbar macht. Insbesondere innenliegende Strukturen einer Probe werden durch die Verwendung von Röntgenstrahlung bei der Bildaufnahme erfasst. Zunächst ähnelt die CT einer klassischen Durchstrahlungsprüfung, wie sie auch in der klinischen Diagnostik angewandt wird. Abbildung 2.1 zeigt den Aufbau einer sogenannten *Kegelstrahl-CT*. Die

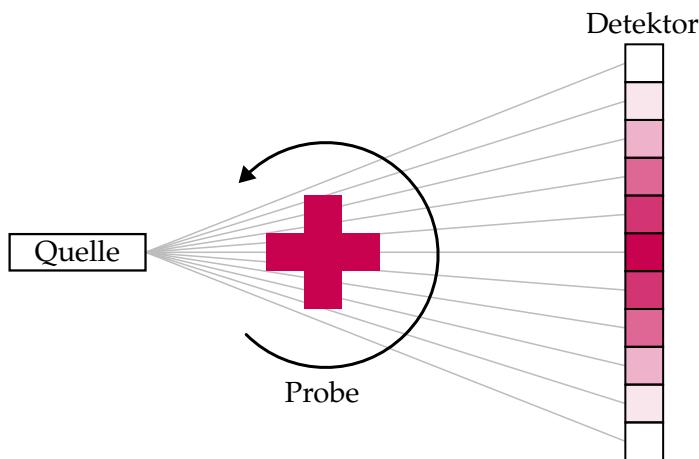


Abbildung 2.1: Aufbau einer Kegelstrahl-CT bestehend aus einer (Röntgen-) Quelle, einem Detektor und einer zu scannenden Probe. Die Pixel des Detektors werden schematisch durch die Kacheln dargestellt.

Probe wird vor der idealisierten, punktförmigen Röntgenstrahlungsquelle (im Folgenden kurz: *Quelle*) positioniert und von der kegelförmig auslaufenden Röntgenstrahlung durchleuchtet. Durch die Wechselwirkung der Röntgenstrahlung mit dem Material der Probe wird die Strahlung abgeschwächt. Hinter der Probe befindet sich ein Detektor, der die transmittierte Strahlung detektiert und in Form einer Projektion aufzeichnet. Erweiternd zur einfachen Durchstrahlungsprüfung wird die Probe nach der Aufnahme einer Projektion gedreht, sodass sich die Abbildungsperspektive relativ zur Probe ändert. Anschaulich ergibt sich wie beim räumlichen Sehen mit zwei (oder mehr) Augen eine Tiefeninformation, indem die Bilder aus verschiedenen Perspektiven zusammengerechnet werden. Diese Berechnung wird in der CT als *Rekonstruktion* bezeichnet und ergibt für einen Satz zweidimensionaler Projektionen ein dreidimensionales Volumen, welches daher *rekonstruiertes Volumen* genannt wird. In praktischen CT-Anwendungen werden mehrere hundert bis hin zu einigen tausend Projektionen aus verschiedenen Perspektiven, meist verteilt auf den Vollkreis, aufgenommen.

Aus der Kegelstrahl-Geometrie ergibt sich, dass Details der Probe vergrößert auf den

Detektor projiziert werden. Die geometrische Vergrößerung M (englisch: *Magnification*) lässt sich mittels Strahlensatz durch

$$M = \frac{SDD}{SOD} \quad (2.1)$$

mit dem Abstand zwischen Quelle und Detektor SDD (englisch: *Source-Detector-Distance*) und dem Abstand zwischen Quelle und Probe SOD (englisch: *Source-Object-Distance*) berechnen. Verschiedene Vergrößerungen erlauben somit das Scannen unterschiedlich dimensionierter Proben, beziehungsweise Scanvolumina, bei maximaler Ausnutzung der Detektorfläche.

In den folgenden Abschnitten werden die wichtigsten Bestandteile eines CT-Scanners vorgestellt und die zugrundeliegende Funktionsweise beschrieben. Für die Beantwortung der ersten und dritten Forschungsfrage ist ein grundlegendes Verständnis der Computertomographie aus theoretischer Sicht zentral. Insbesondere für die Modellierung des CT-Messprozesses im Kontext der ersten Forschungsfrage werden die zugrundeliegenden Wechselwirkungsprozesse von Photonen mit Materie in Abschnitt 2.1, die (technische) Erzeugung von Röntgenstrahlung in Abschnitt 2.2 sowie die Detektion ebendieser Röntgenstrahlung in Abschnitt 2.3 beschrieben. Die Rekonstruktion wird in Abschnitt 2.4 eingeführt, um die resultierenden CT-Artefakte zu verstehen, deren Einfluss einer der Hauptgründe für die Nutzung eines DL-basierten Ansatzes in dieser Arbeit ist. In Abschnitt 2.5 werden verschiedene Strategien zur Datenaufnahme, sogenannte *Trajektorien*, vorgestellt. Das Konzept der Dual-Energy Computed Tomography (DECT) wird in Abschnitt 2.6 gemeinsam mit praktischen Anwendungsmöglichkeiten, beispielsweise der Materialerkennung, eingeführt. Abschließend wird in Abschnitt 2.7 eine reale CT-Anlage vorgestellt, die in den folgenden Kapiteln für Vergleichsmessungen verwendet wird.

2.1 WECHSELWIRKUNG VON PHOTONEN MIT MATERIE

Die Energie eines Photons mit der Frequenz ν lässt sich durch

$$E = h \cdot \nu$$

mit dem Planckschen Wirkungsquantum $h = 4.135 \cdot 10^{-15} \text{ eV} \cdot \text{s}$ beschreiben. Abhängig von dieser Energie interagieren Photonen unterschiedlich mit Materie. Im Rahmen dieser Arbeit werden lediglich Photonen mit Energien im Bereich einiger Kiloelektronenvolt (keV) bis Megaelektronenvolt (MeV) betrachtet¹, sodass die folgenden Wechselwirkungsprozesse dominieren: der photoelektrische Effekt, die Compton-Streuung, die Rayleigh-Streuung und die Paarbildung.

¹ Das genannte Energieband ist der effektiv für die CT nutzbare, messtechnisch sinnvolle Energiebereich. Photonen einer niedrigeren Energie, im Bereich einzelner keV , können die meisten makroskopischen Proben nicht durchdringen und sind somit in der CT nicht allgemein nutzbar. Sehr hochenergetische Photonen weisen stark abnehmende Wirkungsquerschnitte einerseits mit den Probenmaterialien und andererseits mit den Detektormaterialien auf, sodass der Bildkontrast absinkt.

Photoelektrischer Effekt

Der *photoelektrische Effekt* (kurz: *Photoeffekt*) beschreibt die Absorption eines Photons durch ein gebundenes Elektron. Hierbei muss die Energie des einfallenden Photons E_γ mindestens die Bindungsenergie des Elektrons E_B aufbringen, sodass das Elektron den Bindungszustand verlassen kann. Überschüssige Energie wird in kinetische Energie des Elektrons umgewandelt, sodass

$$\underbrace{E_\gamma = h \cdot v}_{\text{Photon}} = \underbrace{-E_B + E_{\text{kin}}}_{\text{Elektron}}$$

gilt. In der Literatur werden Bindungsenergien E_B teils negativ angegeben, sodass das negative Vorzeichen zum Ausgleich der Energiebilanz benötigt wird [23, S. 257f]. Der Photoeffekt wird schematisch in Abbildung 2.2 dargestellt.

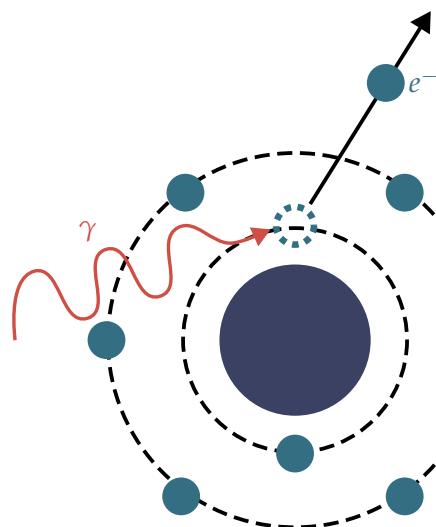


Abbildung 2.2: Schematische Darstellung des photoelektrischen Effektes. Ein Elektron e^- wird durch Absorption eines Photons γ aus seinem Bindungszustand gelöst. Der Atomkern ist dunkelblau dargestellt. Die Abbildung ist nicht maßstabsgetreu.

Compton-Streuung

Bei der *Compton-Streuung* führt ein Photon einen inelastischen Stoß mit einem quasifreien Elektron aus [23, S. 257]. Das Photon überträgt einen Teil seiner Energie an das Elektron, sodass sich die Wellenlänge λ des Photons auf λ' vergrößert. Die effektive Zunahme der Wellenlänge $\Delta\lambda$ beträgt mit der Compton-Wellenlänge des Elektrons $\lambda_C = 2.426 \times 10^{-12} \text{ m}$ und dem Streuwinkel φ

$$\Delta\lambda = \lambda' - \lambda = \lambda_C(1 - \cos \varphi) \quad . \quad (2.2)$$

Abbildung 2.3 stellt die Compton-Streuung konzeptionell dar. Die Zunahme der Wellenlänge, und damit der absolute Energieübertrag auf das Elektron, ist demnach vollständig durch den Streuwinkel φ beschrieben [3, S. 50f.]. Die Verteilung dieser Streuwinkel wird

durch die Klein-Nishina-Gleichung beschrieben und ist für steigende Photonenenergien zunehmend vorwärts (in die ursprüngliche Bewegungsrichtung des Photons) gerichtet [3, S. 50f.]. Demnach besteht ein gerichteter Röntgenstrahl nach Durchgang durch Materie bei zunehmenden Photonenenergien zunehmend aus bereits gestreuten Photonen, die einer anderen Energieverteilung folgen.

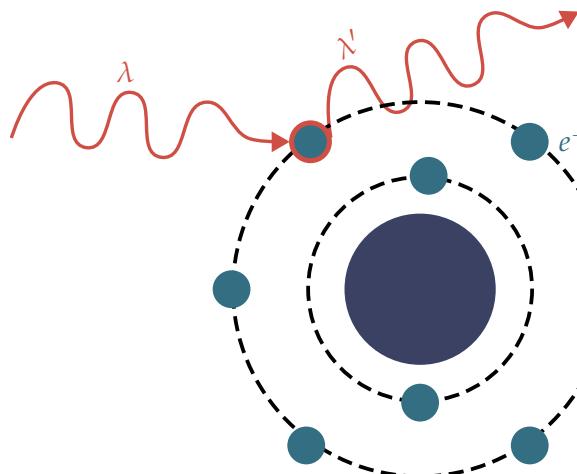


Abbildung 2.3: Schematische Darstellung der Compton-Streuung: Ein einfallendes Photon wird an einem schwach gebundenen bzw. freien Elektron gestreut, wobei es seine Wellenlänge λ auf λ' vergrößert. Die Abbildung ist nicht maßstabsgetreu.

Rayleigh-Streuung

Bei der *Rayleigh-Streuung*, im Folgenden auch *kohärente Streuung*, wechselwirkt ein Photon mit dem Feld der Elektronenhülle eines Atoms, ohne dass Energie effektiv übertragen wird, da die Energie des Photons zu niedrig für einen Ionisations- oder Anregungsprozess ist [3, S. 53] – lediglich die Richtung des Photons ändert sich, sodass eine primäre, gerichtete Strahlung räumlich zerstreut wird. Die Rayleigh-Streuung ist somit ein elastischer und kohärender Streuprozess.

Paarbildung

Die letzte relevante Wechselwirkung ist die *Paarbildung*, die ab Photonenenergien oberhalb von 1022 keV auftritt. Ab dieser Energie kann sich das Photon in ein Elektron-Positron-Paar umwandeln, sofern ein Atomkern in der Nähe ist, um die Impulserhaltung zu gewährleisten. Da diese Dissertationsschrift auf Röntgenstrahlung unterhalb dieser Energieschwelle aufbaut, wird auf eine detaillierte Beschreibung der Paarbildung bewusst verzichtet. Weiterführende Informationen können bei den Autoren Demtröder [23, S. 257f] und Knoll [3, S. 51f] eingesehen werden.

Massenschwächungskoeffizient

Die aus allen genannten Wechselwirkungsprozessen aufsummierte Abschwächung der Röntgenstrahlung wird im sogenannten *Massenschwächungskoeffizienten* μ_m zusammengefasst [3, S. 54; G. 2.22], der für ein Material m mit einer Dichte ρ_m die Beziehung

$$\mu_{L,m}(E) = \rho_m \cdot \mu_m(E) \quad (2.3)$$

mit dem bereits in der Einleitung genannten linearen Abschwächungskoeffizienten $\mu_{L,m}(E)$ erfüllt. Abbildung 2.4 zeigt die verschiedenen Anteile der Wechselwirkungsprozesse am linearen Abschwächungskoeffizienten von Kupfer. Der grau hinterlegte Bereich markiert einen typischen Energiebereich für die industrielle CT-Bildgebung. Für das gezeigte Material Kupfer sind somit der Photoeffekt und die Compton-Streuung die wichtigsten Dämpfungsprozesse von Röntgenstrahlung im markierten Energiebereich.

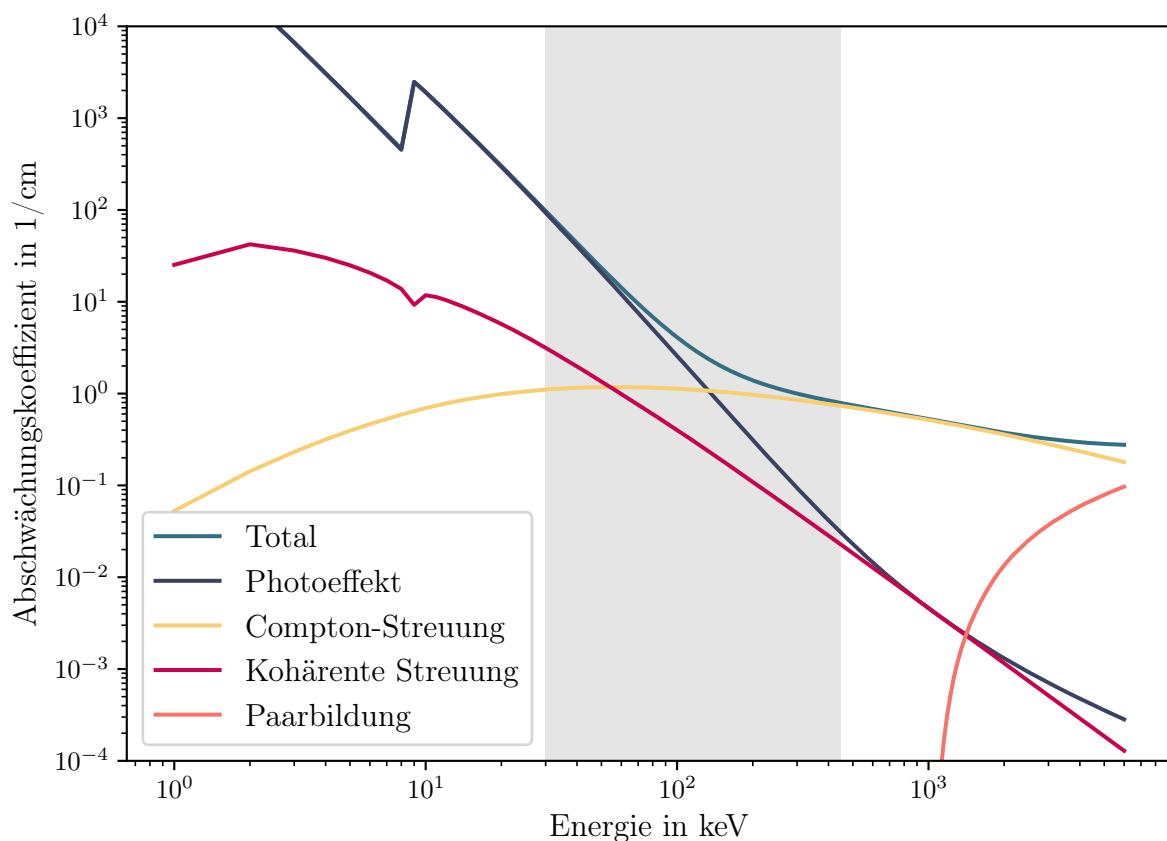


Abbildung 2.4: Linearer Abschwächungskoeffizient von Kupfer *Total* mit den zugrundeliegenden Abschwächungseffekten: *Photoeffekt*, *Compton-Streuung*, *Kohärente Streuung* und *Paarbildung*. Die Abschwächungskoeffizienten wurden mit dem Programm *aRTist* [24] simuliert. Ein typischer Energiebereich für die CT-Bildgebung ist im Plot grau hinterlegt und liegt zwischen 30 keV und 450 keV.

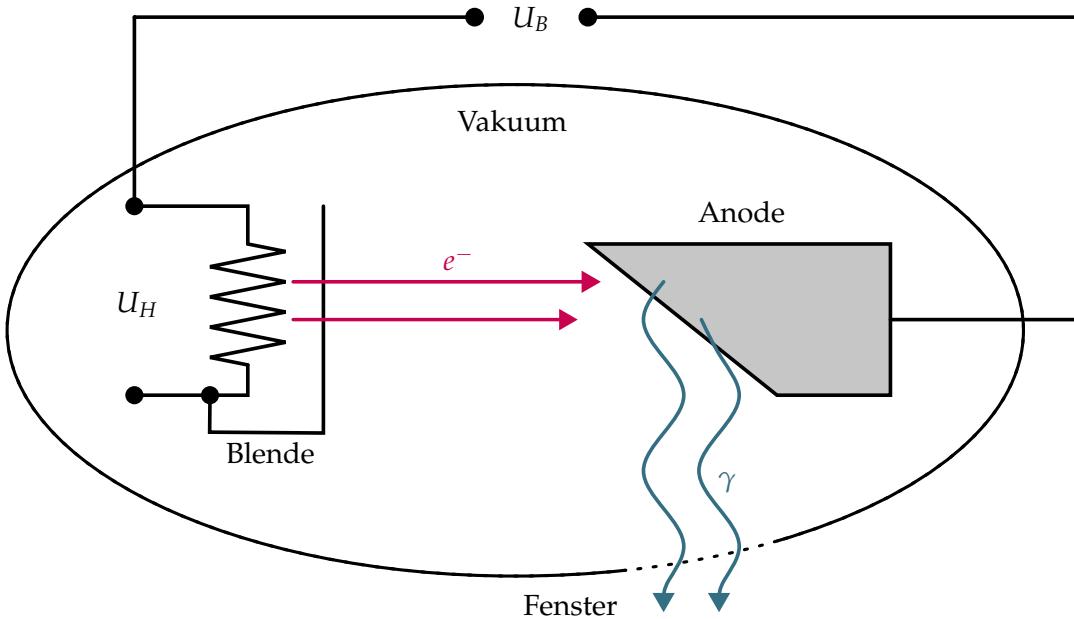


Abbildung 2.5: Schematischer Aufbau einer Röntgenquelle mit der Heizspannung U_H zur Erzeugung freier Elektronen e^- und der Beschleunigungsspannung U_B zur Beschleunigung der freien Elektronen. Durch Interaktion mit dem Anodenmaterial wird die kinetische Energie der freien Elektronen in Photonen γ umgewandelt.

2.2 QUELLE: ERZEUGUNG VON RÖNTGENSTRAHLUNG

Die zur Bildgebung benötigte Röntgenstrahlung wird häufig durch sogenannte *Röntgenquellen* (historisch: *Röntgenröhren*) technisch erzeugt.² Abbildung 2.5 stellt den Aufbau einer Röntgenquelle dar. Eine Glühkathode, häufig in Form eines Wolframdrahtes, wird durch das Anlegen einer Heizspannung U_H und dem daraus resultierenden Heizstrom erwärmt, sodass die Elektronen aus dem Wolframdraht zunehmend ins Vakuumniveau übergehen. Die nun freien Elektronen werden mithilfe eines elektrischen Feldes E_e , dem Resultat der sogenannten *Beschleunigungsspannung* U_B , in Richtung der Anode beschleunigt. Die Zunahme der kinetischen Energie der Elektronen mit der Elementarladung q_e beträgt $\Delta E_{\text{kin}} = U_B \cdot q_e$. Zentral ist, dass der gesamte Prozess im Vakuum stattfindet, so dass die mittlere freie Weglänge der Elektronen sehr viel größer als die Strecke zwischen Glühkathode und Anode ist. Die beschleunigten Elektronen erreichen die Anode und regen die Anodenatome an. Während der Relaxation erzeugen die Anodenatome ein für das Anodenmaterial charakteristisches Emissionsspektrum (auch: *Quellspektrum*). Das Spektrum setzt sich aus zwei wesentlichen Bestandteilen zusammen: der *charakteristischen Röntgenstrahlung* und der *Bremsstrahlung*. Im folgenden Beispiel wird Wolfram als Anodenmaterial angenommen, um exemplarische Quellspektren zu simulieren und zu beschreiben.

² Neben technisch erzeugter Röntgenstrahlung können auch radioaktive Strahler zur Bildgebung verwendet werden [25]. Radioaktive Strahler werden vorrangig für Anwendungen im mobilen Einsatz, beispielsweise zur Prüfung des Betons eines Brückenpfeilers, eingesetzt.

Charakteristische Röntgenstrahlung

Die Atome im Anodenmaterial können durch die einfallenden Elektronen ionisiert werden. Da die unbesetzten Energiezustände in den ionisierten Atomen durch Elektronen aus höheren Energieniveaus, oder dem Vakuumniveau, besetzt werden, geben die Elektronen jeweils beim Übergang in den unbesetzten Zustand ein Photon ab. Bedingt durch die diskreten Energiezustände der Elektronen in einem Atom, gibt es ebenso diskrete Übergänge, die die sogenannte *charakteristische Röntgenstrahlung* im Emissionsspektrum bilden [23, S. 254ff]. Die Bindungszustände K 1s mit 69.525 keV und L₁ 2s mit 12.1 keV sind an der Ausbildung der charakteristischen Linien von Wolfram beteiligt [26, S. 1–6]. Folgende Hauptlinien (historisch auch: *Serien*) werden bedingt durch Elektronenübergänge in den genannten Bindungszuständen erwartet: vom Vakuumniveau in die K-Schale mit etwa 69 keV und von der L-Schale in die K-Schale mit etwa 57 keV. In ebendiesen Größenordnungen können die erwarteten charakteristischen Emissionslinien in Abbildung 2.6 beobachtet werden. Außerhalb des dargestellten, praktisch für die industrielle CT nutzbaren Energiebereiches in Abbildung 2.6 tritt zudem der Übergang vom Vakuumniveau in die L-Schale mit etwa 12 keV auf. Die Berechnung des Spektrums berücksichtigt auch Übergänge aus anderen Bindungszuständen, sodass weitere (Neben-) Linien zu den oben beschriebenen Hauptlinien sichtbar sind.

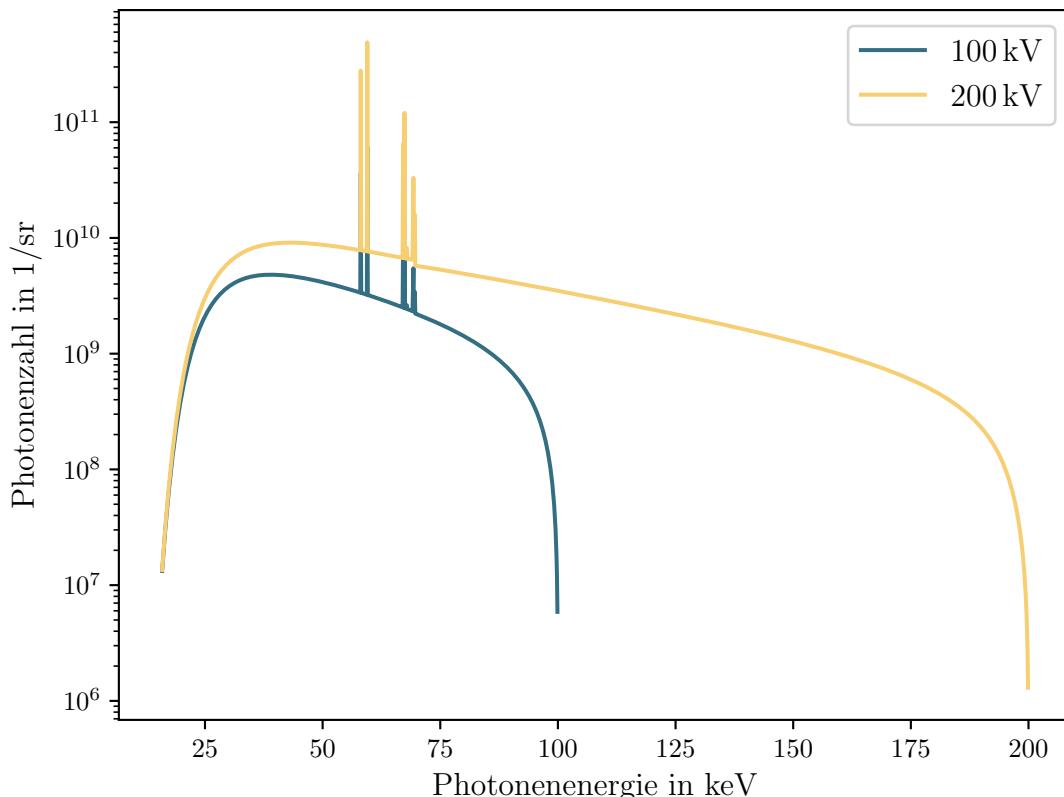


Abbildung 2.6: Röntgenspektren für zwei verschiedene Beschleunigungsspannungen $U_B = (100, 200)$ kV. Beide Spektren weisen die gleichen charakteristischen Emissionslinien aufgrund der quantisierten Bindungszustände im Wolframatom aus. Die Bremsstrahlung verteilt sich jeweils über den gesamten Energiebereich der einfallenden Elektronen.

Bremsstrahlung

Neben der charakteristischen Röntgenstrahlung besteht der Großteil der Spektren in Abbildung 2.6 aus der sogenannten *Bremsstrahlung*. Die einfallenden Elektronen werden durch die Coulomb-Felder der Atomkerne im Anodenmaterial beschleunigt und somit von ihrer Bahn abgelenkt. Bei der Beschleunigung eines geladenen Teilchens wird Energie in Form von Photonen abgestrahlt [23, S. 254]. Aufgrund dieser kontinuierlichen Wechselwirkung ist das Spektrum der Bremsstrahlung kontinuierlich (auch: *Röntgenkontinuum*) und über den gesamten Energiebereich der einfallenden Elektronen verteilt.

2.3 DETEKTOR: QUANTITATIVE DETEKTION VON RÖNTGENSTRÄHLUNG

Dieser Abschnitt basiert auf Knoll (2000) [3] und beschreibt die physikalischen Grundlagen eines Detektors für Röntgenstrahlung im messtechnisch relevanten Energiebereich für die Computertomographie zwischen wenigen keV bis hin zu einigen MeV. Im Rahmen dieser Thesis werden vorrangig Detektoren verwendet, die auf dem Prinzip der Szintillation basieren. Schematisch ist das Funktionsprinzip in Abbildung 2.7 dargestellt. Die einfallenden Photonen γ wechselwirken mit dem sogenannten *Szintillator*, der im

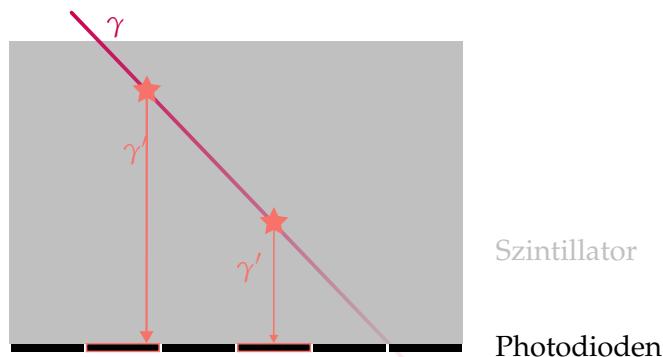


Abbildung 2.7: Funktionelle Skizze eines auf Szintillation basierenden Detektors in seitlicher Ansicht.

Folgenden auch als *aktives Medium* bezeichnet wird. Häufig werden Szintillatoren aus Elementen mit hoher Kernladungszahl gefertigt, um die Wirkungsquerschnitte mit den einfallenden Photonen γ zu maximieren. Durch die Wechselwirkung wird der Szintillator angeregt und gibt die überschüssige Energie in Form neuer Photonen γ' im sichtbaren Bereich wieder ab (Szintillation). Der Szintillator ist für die Photonen γ' transparent, sodass diese sich im Idealfall ohne weitere Interaktion fortbewegen können. Unter dem Szintillator befinden sich die eigentlichen Pixel des Detektors in Form von Photodioden, die die niederenergetischen Photonen γ' effizient detektieren können. Im Vergleich zu den Photodioden, die auch in dünner Bauweise eine hohe Effizienz bei der Detektion der niederenergetischen Photonen γ' erreichen, werden die Szintillatoren dicker ausgelegt, um die Interaktionswahrscheinlichkeit mit den hochenergetischen Photonen γ zu erhöhen.

Insgesamt wird die im Szintillator deponierte Energie in elektrische Ladung umgesetzt,

die in einem Kondensator zeitlich aufsummiert und anschließend nach der sogenannten *Integrationszeit* ausgelesen wird. Ein Detektor nach diesem Konzept wird Energy-Integrating Detector (EID) genannt, da die deponierte Energie über die Zeit integriert wird. Der Szintillator sollte die von den Photonen γ deponierte Energie idealerweise linear in niederenergetische Photonen γ' umsetzen [3, S. 219], sodass die im Pixel ausgelesene Ladung proportional zur ursprünglich deponierten Energie im Szintillator ist. Neben der Gesamtzahl der Pixel d_x und d_y , angeordnet in einem zweidimensionalen Array, ist der Abstand zwischen zwei Pixelmittelpunkten, der sogenannte *Pixel Pitch*, eine wichtige Eigenschaft des Detektors. Die in Abbildung 2.7 dargestellte Situation wird *Schrägdurchstrahlung* genannt und ist zugunsten der Abbildung überspitzt dargestellt. Im Bestfall liegt die Detektorebene senkrecht zum einfallenden Röntgenstrahl, sodass ein einzelnes Photon γ ein räumlich präzise lokalisiertes Signal in einem einzigen Pixel hinterlässt.

2.3.1 Bildkorrekturen

Die vom Detektor aufgenommenen Projektionen müssen korrigiert werden, da verschiedene Effekte die Bildqualität maßgeblich beeinträchtigen.

Erstens wird für jedes Pixel eine sogenannte *Flatfield-Korrektur* durchgeführt, um das Untergrundsignal, den *Offset*, bedingt durch den Dunkelstrom des Detektors aus dem Signal zu entfernen und um den Verstärkungsfaktor, den *Gain*, für jedes Pixel zu ermitteln. Hierzu werden ein Offset-Bild P_{Offset} (auch: *Dunkelbild*) ohne Röntgenstrahlung und ein Gain-Bild P_{Gain} (auch: *Hellbild*) im Freistrahl ohne Probe aufgenommen. Ein unkorrigiertes Pixel $P^{x,y}$ kann durch

$$P_{\text{GOC}}^{(x,y)} = \frac{P^{(x,y)} - P_{\text{Offset}}^{(x,y)}}{P_{\text{Gain}}^{(x,y)} - P_{\text{Offset}}^{(x,y)}} \cdot \overline{P_{\text{Gain}}} \quad (2.4)$$

in ein korrigiertes Pixel $P_{\text{GOC}}^{x,y}$ umgerechnet werden, wobei (x, y) die Position des Pixels auf dem Detektor innerhalb des Pixelarrays beschreibt. $\overline{P_{\text{Gain}}}$ ist der Mittelwert des Gain-Bildes und fungiert als Normierung.

Zweitens gibt es Pixel, die so stark beschädigt sind, dass ihr Signal nicht verwendet werden kann. Die genauen Gründe für den Ausfall einzelner Pixel reichen von mechanischer Beschädigung bis hin zu elektrischen Problemen beim Auslesen und sind für die Korrektur in erster Instanz nicht relevant. Diese Pixel werden *Bad Pixel* (auch: *tote Pixel*) genannt und können mithilfe eines Prüfverfahrens aus der Norm ASTM E2597/E2597M - 14 [27] identifiziert werden. Neben bilinearer Interpolation [28, S. 65 f.] der Nachbarpixel kann ebenso ein Median-Filter [28, S. 326] verwendet werden, um basierend auf den Nachbarpixeln einen Ersatzwert für das beschädigte Pixel neu zu berechnen. Problematisch wird dieser Lösungsansatz, sobald größere Ansammlungen beschädigter Pixel, sogenannte *Cluster*, auftreten, da die benachbarten und ebenfalls beschädigten Pixel nicht zur Korrektur verwendet werden können. Daher müssen für die Korrektur von Clustern komplexere Algorithmen, die einen größeren Bildbereich (auch: Kontext) um das beschädigte Pixel berücksichtigen, verwendet werden. Neben klassischen Bildrestau-

rationsverfahren existieren auch DL-basierte Ansätze, die für die Restauration toter Pixel eingesetzt werden [29]. Ebenso können ganze Zeilen oder Spalten des Detektors ausfallen – eine häufige Folge elektrischer Defekte. Analog zur Argumentation bei Clusterdefekten ist die bilineare Interpolation für die Korrektur ungeeignet. Jedoch kann mit dem Wissen der Ausrichtung der fehlerhaften Zeile (oder Spalte) senkrecht zum Liniendefekt linear interpoliert oder der Median-Filter angewandt werden. Eine Übersicht der klassischen Korrekturverfahren für tote Pixel ist in Abbildung 2.8 dargestellt.

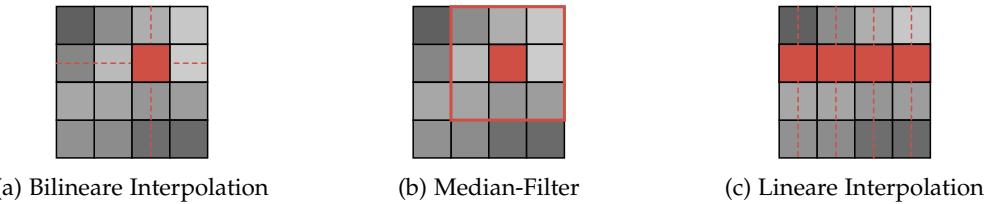


Abbildung 2.8: Übersicht gängiger Korrekturverfahren zur Ausbesserung toter Pixel. Einzelne, isolierte, tote Pixel können mit bilinearer Interpolation (a) oder einem Median-Filter (b) korrigiert werden. Liniendefekte werden mit linearer Interpolation senkrecht zum Defekt (c) korrigiert.

2.3.2 Bildrauschen

Die Projektionen unterliegen einem Bildrauschen, das sich aus dem physikalischen Bildgebungsprozess ableiten lässt. Das von der Quelle erzeugte Spektrum besteht insgesamt aus N_0 Photonen, die mit einer Wahrscheinlichkeit p_a durch die Probe transmittieren und den Detektor erreichen. Durch Anwendung der Zählstatistik ergibt sich, dass im Rahmen einer Standardabweichung $N_0 \cdot p_a \pm \sqrt{N_0 \cdot p_a}$ Photonen den Detektor erreichen [30, S. 170 f.] und zu einem Bildrauschen, basierend auf einem Poisson-Prozess, führen (*Poisson Noise* oder *Shot Noise*). Zusätzlich unterliegen die Detektoren einem elektronischen und thermischen Grundrauschen. Eine weitere Quelle des Bildrauschens kann in der zeitlichen Instabilität der erzeugten Röntgenstrahlung identifiziert werden. Zeitliche Instabilitäten des von der Röntgenquelle emittierten Spektrums beziehen sich in den meisten Fällen auf die Flussdichte des Elektronenstroms auf die Anode (vergleiche Abbildung 2.5), sodass die Leistung des emittierten Spektrums insgesamt variiert, die spektrale Verteilung aber weitestgehend konstant bleibt. Somit können einzelne Projektionen während des CT-Scans zwar andere Zählraten anzeigen, jedoch bleiben die Verhältnisse zwischen den projizierten Objekten stets gleich. Durch die Normierung der Projektionen vor der Rekonstruktion werden erstens diese Schwankung und zweitens räumliche Inhomogenität des Quellspektrums ausgeglichen.

2.4 REKONSTRUKTION

Die mathematische Grundlage zur Rekonstruktion, und damit für die Computertomographie, lieferte Johann Radon im Jahr 1917 [31], indem er zeigte, dass eine zweidimensionale Funktion durch die Werte aller Linienintegrale des Definitionsbereiches ebendieser Funktion eindeutig dargestellt werden kann [30, S. 150]. Der eigentliche Messprozess bei

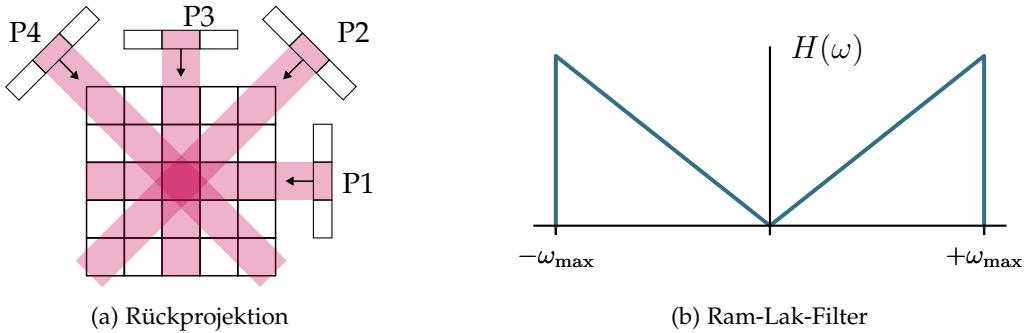


Abbildung 2.9: Schematische Darstellung der Rekonstruktion durch Rückprojektion der Parallelstrahl-Geometrie (a) mit einem möglichen Filter (b) im Frequenzraum $H(\omega)$ für die Anwendung der gefilterten Rückprojektion.

der CT, das heißt die Abschwächung der Röntgenstrahlung durch die Probe gemäß Gleichung 1.1 sowie das Aufzeichnen des integrierten Signals, folgt dem nach Radon beschriebenen Modell und wird auch als *Radon-Transformation* bezeichnet. Die Inverse der Radon-Transformation liefert somit aus einem Satz gemessener Projektionen und einer bekannten Abbildungsgeometrie eine Rekonstruktion, die in einfachster Form durch das *Verschmieren* der Projektionen entlang ihrer Aufnahmerichtung (auch *ungefilterte Rückprojektion*) erfolgt. Eine ausführliche Betrachtung dieser Rekonstruktionsmethode kann bei den Autoren Maier [30, S. 150-156] und Kak [32, S. 56-68] eingesehen werden. Abbildung 2.9a skizziert die Rückprojektion eines Satzes von Projektionen durch ein Zielvolumen unter der Annahme paralleler Röntgenstrahlen (auch: *Parallelstrahl-Geometrie*). Das rekonstruierte Volumen eines gescannten Zylinders wird abhängig von der Anzahl der verwendeten Projektionen in der oberen Zeile von Abbildung 2.10 dargestellt. Mit dem Wissen, dass es sich bei dem in Abbildung 2.10 rekonstruierten Objekt um einen Zylinder, beziehungsweise in der Schnittebene um einen Kreis, handelt, liefert die ungefilterte Rückprojektion auch unter Verwendung von 1000 Projektionen ein unscharfes Resultat. Diese Unschärfe folgt aus der unzureichenden Abtastung hoher Frequenzanteile durch den CT-Messprozess [30, S. 156ff] und kann mithilfe eines Filters, der auf die Projektionen angewandt wird, korrigiert werden.³ Häufig wird der in Abbildung 2.9b geplottete *Ram-Lak-Filter* verwendet, der hohe Frequenzanteile bis zu einer Grenzfrequenz ω_{\max} verstärkt und niederfrequente Anteile unterdrückt. Oberhalb der Grenzfrequenz wird das Signal abgeschnitten, um das hochfrequente Rauschen im Bild zu reduzieren [30, S. 158ff].

In klinischen und industriellen Anwendungen ist die Parallelstrahl-Geometrie zugunsten der Kegelstrahl-CT wenig verbreitet. Daher basieren die meisten Rekonstruktionsalgorithmen auf dem sogenannten *FDK-Algorithmus*, der nach Feldkamp, Davis und Kress benannt wurde [33] und erstmals eine effiziente Möglichkeit beschrieb, dreidimensionale Volumen aus zweidimensionalen Projektionen unter Berücksichtigung der geometrischen Vergrößerung zu rekonstruieren. Alle weiteren Rekonstruktionen in dieser Arbeit werden mit dem FDK-Algorithmus berechnet. Ergänzend sind iterative Rekonstruktionsalgo-

³ Durch Darstellung der Röntgenprojektionen im Fourierraum kann anschaulich gezeigt werden, dass die niedrigen Frequenzanteile des Bildes geometrisch bedingt feiner abgetastet werden als die hochfrequenten Bildanteile. Eine zu dieser Erläuterung ergänzende Abbildung kann bei Kak [32, S. 59, Fig. 3.7] eingesehen werden.

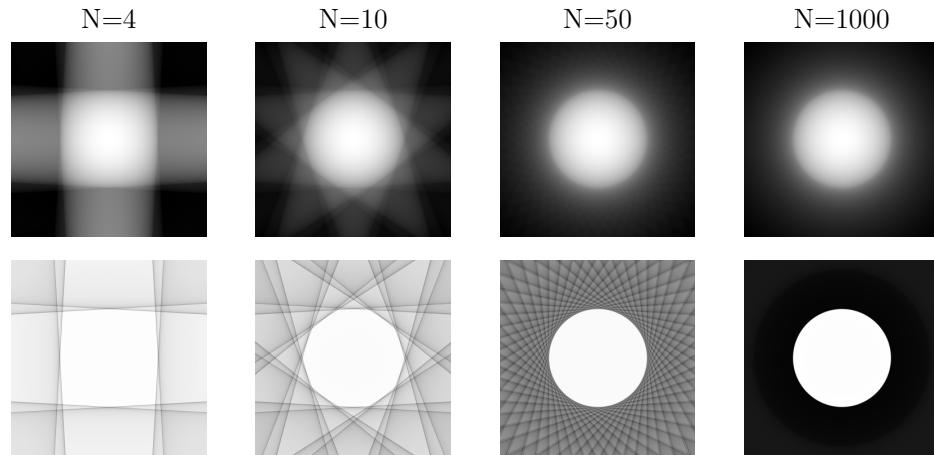


Abbildung 2.10: Ungefilterte Rekonstruktion (obere Reihe) und gefilterte Rekonstruktion (untere Reihe) für verschiedene Projektionszahlen N gleichverteilt über den Vollkreis.

rithmen zu nennen, die im Vergleich zur FDK-Rekonstruktion deutlich rechenintensiver sind. Ihre Vorteile sind im Allgemeinen auf die Reduktion verschiedener CT-Artefakte beschränkt, die im folgenden Abschnitt beschrieben werden. Im weiteren Verlauf dieser Thesis werden keine iterativen Rekonstruktionsalgorithmen verwendet.

2.4.1 CT-Artefakte

Die gefilterte Rückprojektion mit 1000 Projektionen in Abbildung 2.10 (untere Zeile, ganz rechts) ist qualitativ betrachtet artefaktfrei. In praktischen Anwendungsfällen kommt es aufgrund verschiedener Einflüsse dennoch häufig zu sogenannten *CT-Artefakten*, die einen maßgeblichen Einfluss auf die Bildqualität haben können. Aus dem vorangehenden Abschnitt ist das *Undersampling* (deutsch: *Unterabtastung*) in Abbildung 2.10 in Form von linienartigen Artefakten insbesondere bei niedrigen Projektionszahlen zu erkennen. Eine niedrige Projektionszahl führt zu einer Unterabtastung der Probe und somit zu einer verminderten Rekonstruktionsqualität. Die Literatur empfiehlt für einen Detektor mit d_x Pixeln in der Breite⁴ mindestens d_x Projektionen gleichverteilt auf einen Vollkreis für die Rekonstruktion zu verwenden [32, S. 186]. Im Folgenden wird ohne weiteren Kommentar diese Faustformel angewandt. Eine weitere Artefaktquelle besteht in der Aufhärtung der polychromatischen Röntgenstrahlung durch die energieabhängige Dämpfung innerhalb einer Probe [30, S. 171 f.]. Abbildung 2.4 zeigt die Energieabhängigkeit des linearen Abschwächungskoeffizienten für Kupfer. Im für die CT-Bildgebung relevanten, grau schattierten Bereich sinkt der lineare Abschwächungskoeffizient streng monoton, sodass die niederenergetischen Anteile des einfallenden Spektrums eine stärkere Dämpfung erfahren. Das Spektrum wird somit durch die Wechselwirkung mit der Probe nicht nur insgesamt gedämpft, sondern auch aufgehärtet.⁵ Dieser Effekt nimmt mit der Dicke einer gescannten Probe zu, sodass in den inneren Schichten eines homogenen Zylin-

⁴ Gemeint ist die Pixelausbreitungsrichtung \vec{x} , die senkrecht zur Rotationsachse (CT-Achse) und senkrecht zur Strahlrichtung steht.

⁵ Historisch bedingt wird von *weicher* (niederenergetischer) und *harter* (hochenergetischer) Röntgenstrahlung gesprochen.

ders niedrigere Abschwächungskoeffizienten aufgrund des aufgehärteten Spektrums rekonstruiert werden. Abbildung 2.11 zeigt den beschriebenen Effekt und die daraus resultierende, charakteristische Form im Schnittbild (rechte Spalte) – das sogenannte *Cupping*. Zur Reduktion der Strahlaufhärtungsartefakte kann eine Vorfilterung des Röntgenspektrums verwendet werden [30, S. 172]. Die Vorfilterung härtet das Spektrum vor der Interaktion mit der Probe auf, sodass insgesamt weniger Photonen aus den niedrigeren Energiebereichen für den Messprozess verwendet werden. Somit sinkt die Stärke der Strahlaufhärtungsartefakte deutlich ab. In der unteren Zeile von Abbildung 2.11 wird das gleiche Quellspektrum mit Vorfilterung durch ein 1 mm dickes Kupferblech zur Bildaufnahme verwendet. Im Vergleich zur Bildaufnahme ohne Vorfilterung (obere Zeile) ist zu erkennen, dass die Strahlaufhärtung reduziert werden konnte. Zudem sind die absolut rekonstruierten Abschwächungskoeffizienten in Abbildung 2.11 (rechte Spalte) durch die Vorfilterung gesunken, da die effektive Energie des einfallenden Spektrums gestiegen ist und der effektive Abschwächungskoeffizient gemäß Abbildung 2.4 in diesem Energiebereich monoton fällt.

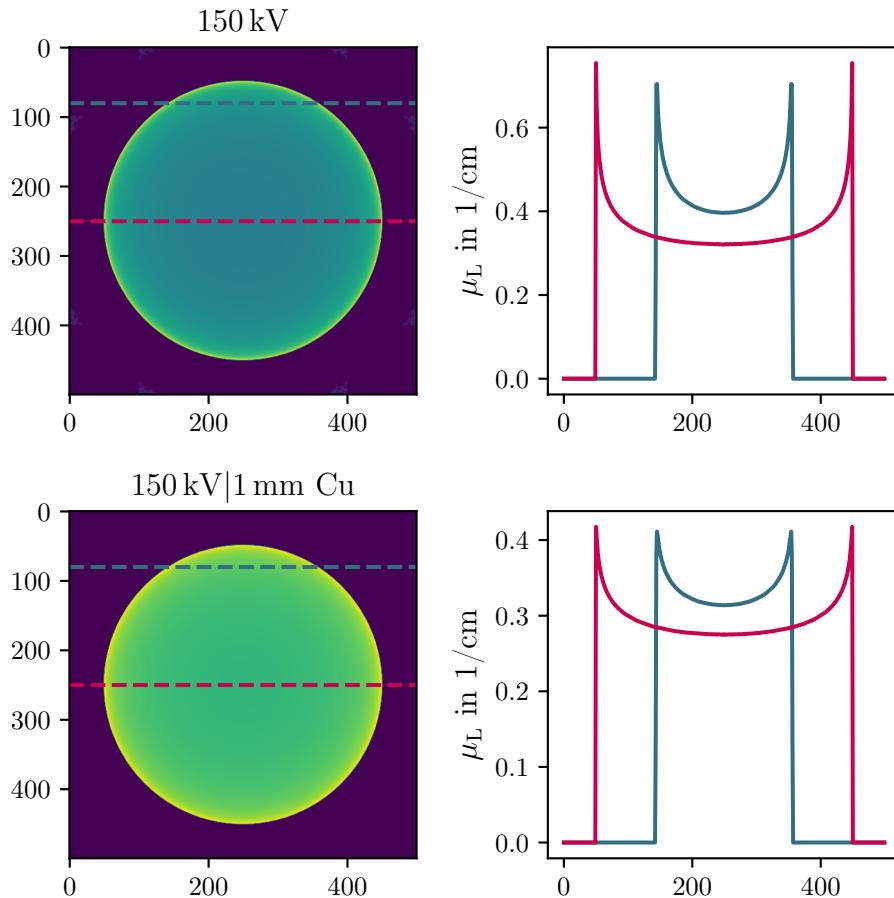
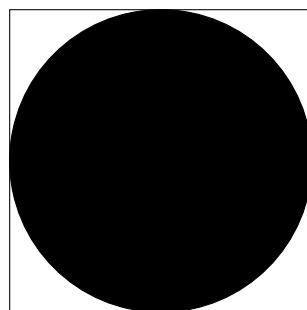


Abbildung 2.11: Rekonstruierte Volumen eines Zylinders mit verschiedenen Röntgenspektren: ungefiltertes Spektrum (obere Zeile) und durch 1 mm Kupfer vorgefiltertes Spektrum (untere Zeile) bei gleicher Beschleunigungsspannung $U_B = 150 \text{ kV}$. Die Schnitte durch die Volumen entlang der gestrichelten Linien werden in der rechten Spalte geplottet. Sofern nicht anders beschriftet, handelt es sich bei den Achsen um numerierte Voxel.

Wie in Abschnitt 2.1 erwähnt, liefert die Compton-Streuung teilweise den stärksten

Beitrag zur Abschwächung der Strahlung im CT-relevanten Energiebereich (vergleiche Abbildung 2.4). Da es sich um einen Streuprozess mit einer bestimmten Winkelverteilung handelt, wird ein Teil der Streustrahlung ebenfalls in Richtung des Detektors ausgesandt und trägt zum gemessenen Signal bei. In der klassischen FDK-Rekonstruktion, beschrieben in Kapitel 2.4, wird dieser Streuprozess nicht berücksichtigt, sodass mitunter von der Realität abweichende Abschwächungskoeffizienten rekonstruiert werden.⁶ Ebenso werden Abbildungsunschärfen, zum Beispiel durch eine nicht ideal punktförmige Röntgenstrahlungsquelle, nicht berücksichtigt, sodass die Detailerkennbarkeit weiter abnimmt. Während die Kreisprojektion die Mittelschicht der Probe⁷ ideal abbildet (zweidimensional auch: *Fächerstrahl-Geometrie*), werden die Schichten im Randbereich geometrisch verzerrt und diagonal durchstrahlt abgebildet, sodass sich die sogenannten *Feldkamp-Artefakte* herausbilden. Sie haben einen direkten Einfluss auf die absolut rekonstruierten Abschwächungskoeffizienten und somit auch auf die im Folgenden vorgestellte Materialbestimmung, die auf ebendiesen Abschwächungskoeffizienten basiert. Das letzte hier genannte CT-Artefakt ist der *Partial Volume Effect*, der den partiellen, inhomogenen Inhalt eines rekonstruierten Voxels beschreibt. Anhand des folgenden Beispiels in Abbildung 2.12 kann der Partial Volume Effect veranschaulicht werden. Der Effekt tritt somit an



(a) Probe kleiner als Voxel



(b) rekonstruiertes Voxel

Abbildung 2.12: Ausgangssituation einer runden Probe, die den Durchmesser des umliegenden Voxels aufweist (a). Aufgrund der Voxelgröße kann das runde Objekt nicht korrekt rekonstruiert werden. Das rekonstruierte Voxel (b) zeigt somit den Mittelwert des Inhalts von (a), also eine Mischung aus der Probe und ihrer Umgebung – den Partial Volume Effect.

allen räumlichen Eigenschaften einer Probe auf, die kleiner sind als die Voxelgröße selbst (häufig: Kanten und Partikel).

2.5 TRAJEKTORIEN

Neben der bereits am Anfang dieses Kapitels beschriebenen Kreisprojektion in Abbildung 2.13a wird die Helixprojektion, dargestellt in Abbildung 2.13b, häufig verwendet. Die Nutzung der Helixprojektion bietet den Vorteil, dass jedes Volumenelement der Probe zu einem bestimmten Zeitpunkt in der Mittelebene der Abbildungsgeometrie liegt und

⁶ Konkret führt die Streustrahlung zu einer Reduktion der rekonstruierten Abschwächungskoeffizienten innerhalb einer Probe.

⁷ Gemeint ist die mittlere Schnittebene durch die Probe, die in derselben Ebene wie die eingezeichnete Trajektorie liegt. Sie befindet sich in der Mitte der Abbildungsgeometrie und wird daher auch *Mittelebene* genannt.

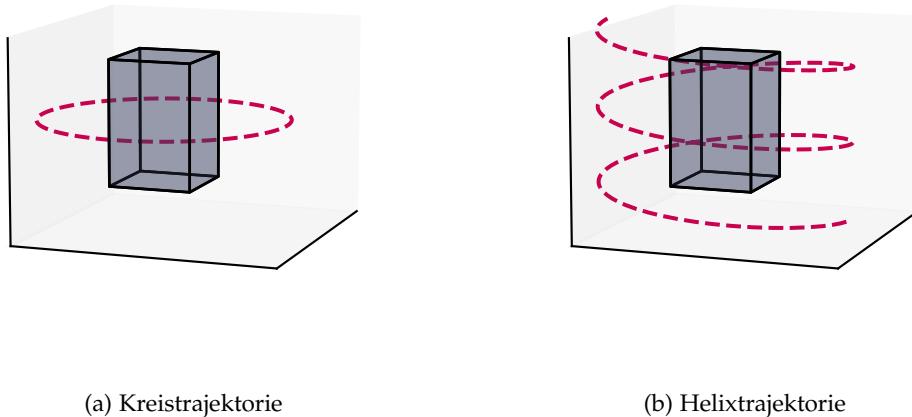


Abbildung 2.13: Renderings einer Kreisrajektorie (a) und einer Helixrajektorie (b) zum Scan der quaderförmigen Probe. Die Quelle bewegt sich entlang der roten Linie. Gegenüberliegend zu jeder Quellposition befindet sich der Detektor, sodass beide gemeinsam um die Probe rotieren (siehe Abbildung 2.1). Zur Verbesserung der Darstellung werden die Positionen des Detektors nicht eingezeichnet.

somit keine Feldkamp-Artefakte auftreten. Zudem können längliche Proben mit einer höheren Vergrößerung als bei der Kreistrajektorie abgebildet werden.

2.6 DUAL-ENERGY CT UND KLASISCHE MATERIALZERLEGGUNG

Bei der Computertomographie mit polychromatischer Röntgenstrahlung und einem Energy-Integrating Detector (EID) werden die Abschwächungskoeffizienten einer Probe ortsaufgelöst und energieintegriert gemessen. Das Spektrum der Röntgenstrahlung bestimmt somit den Energiebereich in dem der Abschwächungskoeffizient des Probenmaterials effektiv gemessen wird. Nach Sprawls besteht eine gängige Hilfsgröße zur Beschreibung dieses effektiven Energiebereiches durch die sogenannte *effektive Energie* [34]. Sie ist ebenjene Photonenergie, die ein monochromatischer Röntgenstrahl benötigt, um die gleiche Penetration wie ein gegebener polychromatischer Röntgenstrahl an einer festen Probe aufzuweisen [34]. Die im Folgenden genannten Abschwächungskoeffizienten werden daher zusätzlich mit dem Index *eff* versehen, um zu verdeutlichen, dass es sich um die effektiven Abschwächungskoeffizienten unter Verwendung polychromatischer Röntgenstrahlung handelt. Bei der Durchführung mehrerer CT-Scans mit unterschiedlichen Röntgenspektren, beziehungsweise unterschiedlichen effektiven Energien, kann der Abschwächungskoeffizient in verschiedenen Energiebereichen gemessen und für weiterführende Analysen verwendet werden. Im Fall von zwei unterschiedlichen Energien wird der gesamte Scan auch *DECT* genannt. Die Aufnahme einer DECT kann durch verschiedene Strategien realisiert werden:

- Beim *Multi-Scan* werden zwei separate CT-Scans nacheinander mit verschiedenen Parametern der Röntgenquelle, zum Beispiel Beschleunigungsspannung oder Vorfilterung, durchgeführt [35, S. e893].

- Mehrschichtige Detektoren, sogenannte *Sandwich Detectors*, können während eines einzigen CT-Scans verschiedene Energiekanäle durch die unterschiedlichen Detektorschichten registrieren [35, S. e890 ff.].
- Moderne *Photonenzähler* (auch *Photon-Counting Detectors (PCDs)*) können einzelne Photonen energieaufgelöst detektieren [35, S. e892 f.].
- Durch schnelles Alternieren der Beschleunigungsspannung in der Röntgenröhre, das *Rapid Tube Potential Switching* oder *Fast-kV Switching*, können verschiedene Röntgenspektren zur Aufnahme aufeinanderfolgender Bilder verwendet werden [35, S. e893 ff.].
- Die Verwendung mehrerer, unabhängiger Systeme, bestehend aus Röntgenquelle und Detektor, die den gleichen Prüfraum beleuchten, stellt eine weitere Möglichkeit zur Aufnahme von DECT dar [35, S. e896 f.].

Bereits im Jahr 1976 publizierte Alvarez einen Algorithmus zur Materialzerlegung basierend auf DECT [13]. Eine gängige Darstellung von DECT zur Materialdifferenzierung wird in Abbildung 2.14 für die Materialien Aluminium, Eisen und Kupfer dargestellt. Für zwei gegebene Röntgenspektren, in diesem Beispiel definiert durch die Energie der

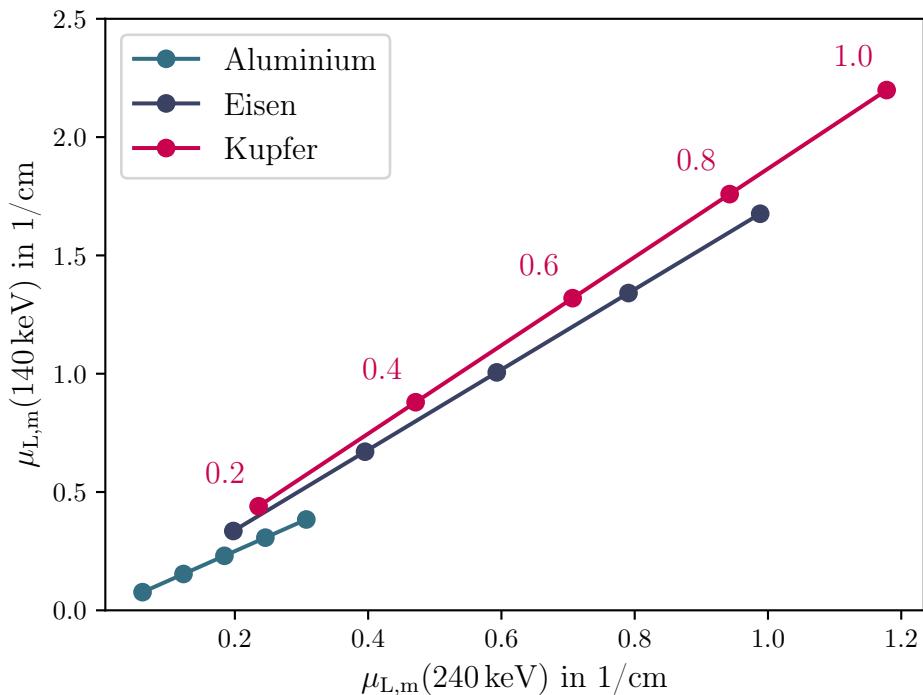


Abbildung 2.14: Ideal rekonstruierte Dual-Energy-Abschwächungskoeffizienten μ_L verschiedener Materialien m basierend auf monochromatischer Röntgenstrahlung mit 140 keV und 240 keV. Die Abschwächungskoeffizienten tragen keinen Index zur Anzeige der effektiven Energie (eff), da es sich zur Vereinfachung um monochromatische Röntgenstrahlung handelt. Jede Linie beschreibt ein Material in verschiedenen relativen Dichten. Die relativen Dichten sind beispielhaft an den Stützpunkten der Kupferlinie annotiert.

monochromatischen Strahlung, werden die rekonstruierten Abschwächungskoeffizienten $\mu_{L,m}$ graphisch aufgetragen. Aus der Definition des linearen Abschwächungskoeffizienten

$$\mu_{L,m}^{\text{eff}} = \rho_m \cdot \mu_m \quad (2.5)$$

mit einer absoluten Dichte ρ_m und dem Massenschwächungskoeffizienten μ_m folgen unmittelbar die Datenpunkte für die relativen Dichten zwischen 0.2 und 1.0. Eine unbekannte Probe kann somit durch einen DECT-Scan eindeutig anhand der rekonstruierten Abschwächungskoeffizienten einer Gerade und damit einem Material und einer Dichte zugeordnet werden. Analog können für polychromatische Röntgenstrahlung, mit Verwendung der effektiven Energie, in realen Anwendungen Kalibrationsmessungen durchgeführt werden, um Materialien anhand der rekonstruierten Abschwächungskoeffizienten erkennen zu können. Ein Problem bei diesem sehr direkten Ansatz ist in Abbildung 2.11 aus dem vorangehenden Abschnitt über CT-Artefakte ersichtlich: die rekonstruierten Abschwächungskoeffizienten sind in einer homogenen Probe bei Verwendung von polychromatischer Röntgenstrahlung aufgrund der Strahlaufhärtung nicht konstant. Daraus folgt, dass der oben beschriebene Ansatz zur Materialerkennung nur in absoluter Abwesenheit der Strahlaufhärtungsartefakte, also beispielsweise bei der Verwendung einer monochromatischen Röntgenquelle, erfolgsversprechend ist.

2.7 VORSTELLUNG REALE CT-ANLAGE

An einigen Stellen dieser Arbeit bedarf es eines Abgleiches mit realen CT-Daten, die mit der in diesem Abschnitt vorgestellten CT-Anlage aufgenommen werden. Eine Fotografie der verwendeten CT-Anlage vom Typ *diondo d2* ist in Abbildung 2.15 dargestellt. Die Anlage ist auf einem Granit-Manipulator aufgebaut, um den Einfluss der thermischen Ausdehnung und generell durch Vibrationen zu unterdrücken. Die Röntgenquelle kann mit einer maximalen Beschleunigungsspannung $U_B = 240 \text{ kV}$ betrieben werden. In der Röntgenquelle befindet sich eine scheibenförmige Anode, die in Transmission, das heißt die Strahlrichtung der Elektronen entspricht der Richtung des primären Photonenstrahls, betrieben wird (auch: *Transmissionsquelle*). Transmissionsquellen sind aufgrund ihres kleinen Brennflecks⁸ im Vergleich zu Reflexionsquellen (vergleiche Abbildung 2.5) für hochauflöste CT-Scans besser geeignet, wobei dies auf Kosten der abgestrahlten Leistung in Form der Röntgenstrahlung geschieht.

Der Detektor basiert auf einem Caesiumiodid-Szintillator und wird energieintegrierend betrieben (vergleiche *EID* in Abschnitt 2.3). Der Pixel Pitch beträgt $139 \mu\text{m}$ auf einer Fläche mit insgesamt 3000×3000 Pixeln. Zur Bedienung der gesamten CT-Anlage wird die Software *diControl* verwendet. Weiterführende Informationen bezüglich der vorgestellten CT-Anlage können bei der Firma *diondo* in Hattingen angefordert werden.

Einige im weiteren Verlauf dieser Dissertationsschrift gewählte Parameter, beispielsweise die Beschleunigungsspannungen der Röntgenquelle, werden in Anlehnung an die *diondo d2* gewählt, um den angesprochenen quantitativen Vergleich der Daten zu erleichtern.

⁸ Gemeint ist der Bereich aus dem die Röntgenstrahlung innerhalb der Anode emittiert wird.



Abbildung 2.15: Fotografie der verwendeten CT-Anlage *diondo d2* im Dienstleistungsbereich der Firma *diondo*. Rechts im Bild ist die Röntgenquelle zu sehen, links befindet sich der Detektor.

3

GRUNDLAGEN DES DEEP LEARNINGS FÜR DIE COMPUTERTOMOGRAPHIE

Dieses Kapitel dient als Einstieg für CT-Experten in die Methodik des Deep Learnings als Werkzeug für die Bildverarbeitung und insbesondere für die Verarbeitung von CT-Bilddaten im Rahmen dieser Dissertationsschrift. Daher werden vorrangig die zugrunde liegenden Konzepte einiger Bausteine des Deep Learnings im Folgenden phänomenologisch und nicht in voller Tiefe erklärt. Versierte Deep-Learning-Entwickler können zur Abkürzung direkt die Abbildung 3.5 sowie die Publikationen von Ronneberger et al. [36] und Liu et al. [37, 38] einsehen, um die in den folgenden Kapiteln verwendeten Architekturen nachzuvollziehen. Zunächst wird Deep Learning formell in Abschnitt 3.1 anhand eines einfachen neuronalen Netzes eingeführt. Darauf aufbauend werden das Training und die Bewertung des Trainingserfolges in den Abschnitten 3.2 und 3.3 beschrieben. Da im Rahmen dieser Dissertationsschrift mit CT-Bildern gearbeitet wird, vertieft Abschnitt 3.4 die Anwendung des Deep Learnings für die Verarbeitung von Bildern. Zwei im weiteren Verlauf dieser Arbeit genutzte und konzeptionell verschiedene Architekturen zur Bildverarbeitung werden in den Abschnitten 3.4.1 und 3.4.2 vorgestellt: das U-Net [36] und der Swin Transformer [38].

3.1 EINFÜHRUNG: DEEP LEARNING

Gesucht sei eine Funktion f , die eine Variable x (auch: *Input*) in eine Variable y (auch: *Ground Truth*) übersetzt:

$$f : x \mapsto y . \quad (3.1)$$

Weiterhin sei eine Serie, bestehend aus beispielhaften Tupeln (x, y) , den sogenannten *Trainingsdaten*, gegeben. Unter diesen Bedingungen wird vom sogenannten *Supervised Learning* gesprochen, da x und y bekannt sind. Mit dem entsprechenden Vorwissen über die von der Funktion f darzustellende Abbildung, beispielsweise durch Domänenwissen, kann eine Funktionsgleichung explizit aufgestellt und an die Trainingsdaten angepasst werden. Besteht kein derartiges Wissen, kann alternativ nach Hornik ein künstliches neuronales Netz angesetzt werden, das bei geeigneter Auslegung die Fähigkeit besitzt, mathematische Funktionen f anhand der gegebenen Trainingsdaten zu approximieren [39]. Neuronale Netze bestehen aus *Neuronen*, die mithilfe sogenannter *Verbindungen* untereinander vernetzt werden [40, S. 185 f.]. Die Neuronen werden typischerweise in mehreren Schichten (auch: *Layer*) angeordnet. In Abbildung 3.1 wird ein Multi-Layer-Perceptron (MLP) dargestellt, das aus insgesamt drei Schichten besteht, die mithilfe der Verbindungen verknüpft werden. Der Input x ist in diesem Beispiel ein Vektor mit drei Komponenten (x_1, x_2, x_3) und bildet die Neuronen der *Input Layer*. Allgemein werden

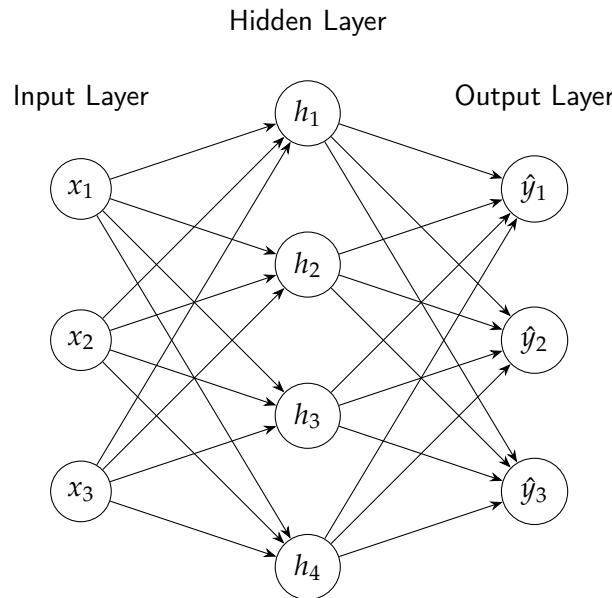


Abbildung 3.1: Architekturdiagramm eines einfachen MLPs mit den Neuronen der Input Layer x , der Hidden Layer h und der Output Layer \hat{y} .

die von neuronalen Netzen verarbeiteten Daten auch *Tensoren*¹ genannt – eine Verallgemeinerung der zuvor beschriebenen Vektordarstellung des Inputs x . Die Neuronen der Hidden Layer h_i werden durch die Linearkombination des Inputs x mit sogenannten *Weights* w , der Addition des sogenannten *Bias* b und der Anwendung der sogenannten *Aktivierungsfunktion* ϕ berechnet durch:

$$h_i = \phi \left(\sum_j w_j \cdot x_j + b \right) \quad . \quad (3.2)$$

Durch die Verwendung einer nichtlinearen Aktivierungsfunktion ϕ kann das MLP auch nichtlineare Funktionen f approximieren [40, S. 187]. Die Aktivierungsfunktion *ReLU* (*Rectified Linear Unit*) gehört zu dieser Klasse der nichtlinearen Aktivierungsfunktionen [40, S. 191 f.], ist definiert durch

$$\phi_{\text{ReLU}}(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (3.3)$$

und nach Goodfellow eine der am häufigsten verwendeten Aktivierungsfunktionen [40, S. 192]. In der Definitionsgleichung von $\phi_{\text{ReLU}}(x)$ symbolisiert x die Klammer aus Gleichung 3.2. Analog zu Gleichung 3.2 werden die Werte der Neuronen in der *Output Layer* (auch: *Prediction*) \hat{y} aus den Neuronen der *Hidden Layer* berechnet. Somit sind die *Weights* w und die *Bias*-Werte b aller Neuronen und Verbindungen eines neuronalen Netzes ebenjene Parameter, die an die gegebenen Trainingsdaten angepasst werden müssen, sodass das MLP die Funktion f bestmöglich approximiert [40, S. 189]. Die Gewichte und die *Bias*-Werte werden daher auch *trainierbare Parameter* θ genannt und ihre Gesamtzahl

¹ Der Begriff steht in direktem Zusammenhang mit dem *Tensor* aus der Mathematik.

wird als *Kapazität* des neuronalen Netzes bezeichnet.

Das beispielhaft dargestellte MLP in Abbildung 3.1 besitzt eine Hidden Layer. Durch das Hinzufügen weiterer Hidden Layers wird das neuronale Netz *tiefer* und in seiner Gesamtheit ausdrucksstärker, da es eine größere Anzahl an Rechenoperationen der Form aus Gleichung 3.2 zur Darstellung der Funktion f nutzen kann [40, S. 185 f.]. Ab welcher Tiefe, beziehungsweise Kapazität, ein neuronales Netz zum Deep Learning zählt, ist in der Literatur nicht einheitlich definiert. Nach Goodfellow zählt ein neuronales Netz zu Deep Learning, sobald es komplexe Merkmale durch die Kombination abstrakter Teilmerkmale in den Hidden Layers ausdrücken kann [40, S. 6 f.]. So könnte ein tiefes neuronales Netz beispielhaft die Erfassung und Unterscheidung der Fellmuster von Giraffen (geflecktes Muster) und Zebras (gestreiftes Muster) als abstraktes Merkmal in einer der vielen Hidden Layers erlernen. In einer anderen Layer könnte es auch die relative Halslänge im Verhältnis zur Gesamtgröße des Tieres – ein weiteres nützliches Merkmal zur Unterscheidung zwischen Giraffen und Zebras – als abstraktes Merkmal erlernen.

3.2 TRAINING NEURONALER NETZE

Die im vorausgehenden Abschnitt erfolgte Vorstellung eines einfachen MLPs in Abbildung 3.1 lässt eine zentrale Frage offen: *Wie werden die trainierbaren Parameter θ an die gegebenen Trainingsdaten (x, y) angepasst?* Zur Beschreibung des sogenannten *Trainings*, ebendieser Anpassung der Parameter θ , werden drei Kernkomponenten benötigt, die im Folgenden aufgezählt werden:

- *Loss*: Im Allgemeinen berechnet ein neuronales Netz aus einem Input x (Input Layer) eine Prediction \hat{y} (Output Layer), die vom Ground Truth y mehr oder weniger stark abweicht. Die Abweichung, der sogenannte *Loss* \mathcal{L} , wird mithilfe sogenannter *Loss Functions* quantifiziert. Für verschiedene Problemarten, beispielsweise Regression und Klassifikation, gibt es verschiedene Loss Functions. Das im Rahmen dieser Dissertationsschrift formulierte Problem zählt formal zur Regression, da aus den Inputs x (CT-Bilder) kontinuierlich verteilte Predictions \hat{y} (Materialien) berechnet werden sollen. Somit wird im Folgenden eine Loss Function für Regressionsprobleme beschrieben. Eine häufig für Regressionsprobleme verwendete Loss Function ist der *Mean Squared Error (MSE-Loss)*, der den mittleren, quadratischen Fehler der Predictions \hat{y} gegenüber den Ground Truths y misst:

$$\text{MSE-Loss}(\hat{y}, y) = \frac{1}{L} \sum_i (\hat{y}_i - y_i)^2 . \quad (3.4)$$

Hierbei ist L die Länge der Vektoren y beziehungsweise \hat{y} , das heißt, konkret am Beispiel des MLPs aus Abbildung 3.1, die Anzahl der Neuronen in der Output Layer. Neben dem MSE-Loss gibt es eine Vielzahl alternativer Loss Functions, die zum Training von Regressionsproblemen verwendet werden und bei Ciampiconi et al. [41] eingesehen werden können.

- *Gradientenbestimmung:* Das vorgestellte MLP in Abbildung 3.1 verwendet größtenteils differenzierbare Rechenoperationen², sodass zur Minimierung des Losses ein Gradientenverfahren genutzt werden kann. Mithilfe der sogenannten *Backpropagation* kann der Gradient für einen Loss \mathcal{L} entlang der trainierbaren Parameter $\nabla_{\theta}(\mathcal{L})$ ermittelt werden [40, S. 225 f.]. Anschaulich wird hierbei der Einfluss jedes trainierbaren Parameters auf den Loss ermittelt.
- *Optimierer:* Mit dem Gradienten $\nabla_{\theta}(\mathcal{L})$ erfolgt die Anpassung der trainierbaren Parameter durch

$$\theta' = \theta - \epsilon \nabla_{\theta}(\mathcal{L})$$

mit der sogenannten *Lernrate* ϵ , sodass sich die aktualisierten trainierbaren Parameter θ' ergeben. Moderne Optimierer im Bereich des DLs verwenden adaptive Lernraten, sodass jeder Parameter aus θ zusätzlich eine multiplikative, individuelle Lernrate zugeordnet bekommt. Durch die Nutzung individueller Lernraten wird die Konvergenz beschleunigt, da Parameter, die einen eher geringen Einfluss auf den Loss haben, das heißt schwache Gradienten aufweisen, mit größeren individuellen Lernraten aktualisiert werden können, während Parameter mit steilen Gradienten niedrigere Lernraten und damit kleinere Optimierungsschritte erfahren. Ein häufig genutzter Optimierer heißt *AdamW* [42].

Anhand von Codeabschnitt 3.1 werden die soeben vorgestellten Kernkomponenten in den Ablauf des Trainings eines neuronalen Netzes eingebettet.

```

1   for batch in batches:
2       x, y = batch
3       y_hat = model(x)
4       loss = lossFunction(y_hat, y)
5       loss.backward()
6       optimizer.step(loss, model)

```

Codeabschnitt 3.1: Globalalgorithmus zum Training eines neuronalen Netzes `model` mit den Inputs `x`, den Ground Truths `y` und den Predictions `y_hat` anhand einer Loss Function `lossFunction` mithilfe eines Optimierers `optimizer`.

Zunächst werden die Trainingsdaten $(x, y)_0, \dots, (x, y)_N$ in sogenannte *Batches*, das heißt Teilstücke, der Größe B aufgeteilt. Jeder Batch wird in Inputs x und Ground Truths y zerlegt. Die Inputs dienen, gemäß obiger Beschreibung, dem neuronalen Netz zur Berechnung der Prediction \hat{y} , die mit den Ground Truths y verglichen werden und einen Loss ergeben. Mittels Backpropagation (`loss.backward()`) wird der Gradient des Losses bezüglich der trainierbaren Parameter $\nabla_{\theta}(\mathcal{L})$ berechnet [40, S. 225]. Der Optimierer kann anschließend die trainierbaren Parameter anpassen (`optimizer.step(..)`), sodass der Loss für den soeben prozessierten Batch reduziert wird. Dieses Verfahren wird für alle Batches der Trainingsdaten in mehreren Durchläufen, den sogenannten *Epochen*,

² Die einzige Ausnahme besteht in der Aktivierungsfunktion ReLU, die nur im Punkt $x = 0$ mathematisch nicht differenzierbar ist. An dieser Stelle beträgt ihr Gradient per Definition $\nabla_{\text{ReLU}} = 0$, um die Gradientenbestimmung weiterhin durchführen zu können. Moderne Alternativen, beispielsweise GELU, sind meist stetig differenzierbar.

durchgeführt.

Die Aktualisierung der trainierbaren Parameter θ auf Basis des Gradienten $\nabla_{\theta}(\mathcal{L})$ birgt in der bisher dargestellten Form eine Gefahr: Der Gradient, und damit die Aktualisierung, eines Gewichtes in einer bestimmten Hidden Layer wird basierend auf der Annahme berechnet, dass die anderen Hidden Layer konstant bleiben (vergleiche Goodfellow [40, S. 354 f.]). Durch die Aktualisierung der Gewichte innerhalb einer Hidden Layer h_i können ihre Ausgabewerte, die als Eingabewerte der nächsten Hidden Layer h_{i+1} agieren, deutlich geändert werden. Dieses Problem kann durch die Anwendung einer Normierung der Daten vor jeder Hidden Layer reduziert werden [40, S. 356 f.]. Mithilfe der sogenannten *Batch-Normalization-Layer* (kurz: *BNorm*) werden die Daten eines Batches bei Anwendung der Layer so normiert, dass sie den Erwartungswert 0 und die Standardabweichung 1 aufweisen [43]. Die *BNorm* reduziert das soeben beschriebene Problem deutlich und trägt insgesamt zu einer Stabilisierung des Trainings bei [40, S. 357 f.].

Rückblickend lassen sich zunächst zwei wichtige sogenannte *Hyperparameter*³ identifizieren: die Lernrate ϵ und die Größe der Batches, die sogenannte *Batch Size*, B . Die Lernrate wird zumeist anhand von Standardwerten aus der Literatur in der Größenordnung 1×10^{-3} für den AdamW-Optimierer gewählt und basierend auf der Historie durchgeführter Trainings angepasst. Gleichermaßen gilt für die Batch Size, die einerseits durch eine Vergrößerung die Trainingsdauer reduziert, jedoch andererseits durch eine Verkleinerung bessere Trainingsergebnisse liefern kann [44].

3.3 BEURTEILUNG DES TRAININGS

Der vorausgehende Abschnitt beschreibt, wie das Training eines neuronalen Netzes durchgeführt werden kann. Er lässt jedoch die Frage offen, wie der Trainingsverlauf beurteilt wird. Vor dem Training wird ein gegebener Trainingsdatensatz in zwei Teildatensätze, ohne gemeinsame Teilmenge, mit den Bezeichnungen *Training Dataset* und *Validation Dataset* zufällig aufgeteilt. Im Rahmen des in Codeabschnitt 3.1 durchgeföhrten Trainings werden nur die Daten aus dem *Training Dataset* verwendet. Der mittlere Loss aller Batches des *Training Datasets* wird als sogenannter *Training Loss* gespeichert. Nach einer Epoche werden die Daten des *Validation Datasets* durch das Modell, vorwärts (ohne Backpropagation und Optimierungsschritt), verarbeitet und der resultierende Loss als sogenannter *Validation Loss* gespeichert. Er ist somit der Loss, den das trainierte Modell auf Daten erreicht, die es nicht zum Training verwendet hat. Die graphische Auswertung des *Training* und *Validation Losses* für alle Epochen eines durchgeföhrten Trainings, die sogenannten *Loss Curves*, lassen Rückschlüsse auf den Verlauf, aber auch auf den Erfolg des Trainings zu. Abbildung 3.2 zeigt beispielhaft zwei verschiedene Trainingsverläufe. Das Modell im linken Plot von Abbildung 3.2 erfährt das sogenannte *Overfitting*. Während das Modell verhältnismäßig gute Vorhersagen auf den (bekannten) Daten des *Training Datasets* macht, verschlechtern sich die Vorhersagen auf den (unbekannten) Daten des *Validation Datasets*. Anschaulich kann dieser Prozess mit der

³ Neben den trainierbaren Parametern θ , die Bestandteil der Architektur sind und während des Trainings angepasst werden, gibt es die sogenannten *Hyperparameter*, die zur Steuerung des Trainings oder zur Erzeugung der Architektur verwendet werden [40, S. 133 f.].

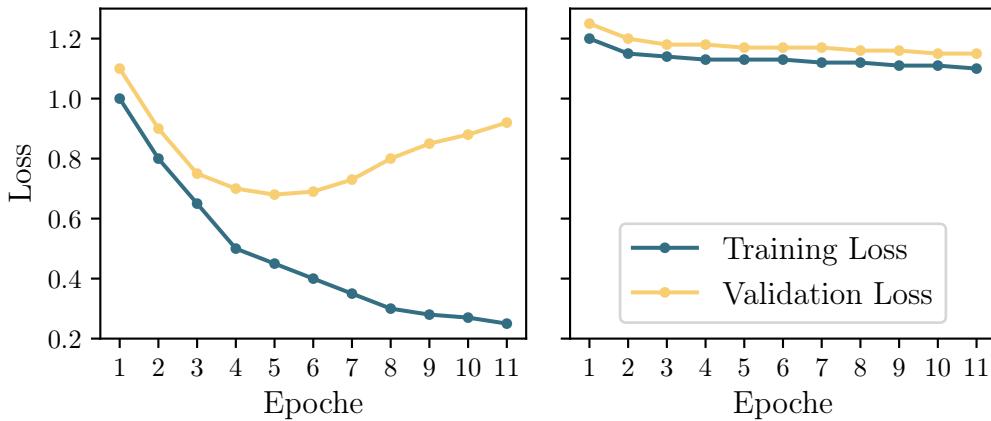


Abbildung 3.2: Training und Validation Loss eines Modells, das von Overfitting (links) und Underfitting (rechts) betroffen ist. Beide Plots teilen sich die Legende sowie die y-Achse.

folgenden Situation verglichen werden: Ein sehr fauler Mathematiklehrer stellt in seinen wöchentlichen Leistungsüberprüfungen stets dieselben Aufgaben. Es sollen immer die ersten Ableitungen derselben drei Funktionen berechnet werden (Training Dataset). Der effiziente Schüler (das neuronale Netz) erkennt nach kurzer Zeit dieses Muster und lernt nur die drei Ergebnisse auswendig und erzielt gute Noten (einen niedrigen Training Loss). Bei Ankunft eines neuen Lehrers, der die Schüler auffordert, andere Funktionen (Validation Dataset) zu differenzieren, scheitern die Schüler (Validation Loss), da sie in den letzten Wochen nicht das Konzept der Differenzierung, sondern nur dieselben drei Ergebnisse auswendig gelernt haben (*Overfitting*).

Anhand des Beispiels kann ebenfalls ein möglicher Ausweg aus dieser Situation identifiziert werden: Der faule Lehrer hätte nicht nur drei Funktionen in seinen Tests, sondern hunderte von Funktionen auflisten sollen. In diesem Fall ist es für den Schüler aufgrund seiner begrenzten *Gehirnkapazität* (vergleiche *Kapazität* als Gesamtheit der trainierbaren Parameter θ) nicht möglich, die Ergebnisse alle auswendig zu lernen. Seine Kapazität ist allerdings ausreichend, um das Konzept der Differenzierung zu erlernen und so alle Aufgaben dieser Art lösen zu können. Neuronale Netze können durch das Training basierend auf einem großen Datenbestand mit einer entsprechenden Varianz nach dem gleichen Schema gezwungen werden, Konzepte zu erlernen und nicht die Ergebnisse der Trainingsdaten auswendig zu lernen. Somit sind (ausreichend) große Datensätze ein zentraler Aspekt bei der Entwicklung neuronaler Netze.

Abbildung 3.2 (rechts) zeigt den anderen Extremfall, der beim Training eines neuronalen Netzes eintreten kann – das sogenannte *Underfitting*. In diesem Fall kann das Modell weder die Ergebnisse auswendig lernen, noch das zugrunde liegende Konzept erlernen. Dieses Verhalten kann in manchen Fällen auf eine zu niedrige Kapazität zurückgeführt werden. Hinzuzufügen ist, dass Underfitting, unabhängig von der Kapazität einer Architektur, auch bei einer ungeeigneten Wahl einer Architektur bezüglich einer Problemstellung beobachtet werden kann und somit das Lernen des zugrundeliegenden Zusammenhangs nicht möglich ist.

3.3.1 Weitere Verbesserungen des Trainings

Das häufig beobachtete Overfitting kann mithilfe unterschiedlicher Techniken reduziert werden. Neben dem oben genannten Vergrößern des Trainingsdatensatzes werden im Rahmen dieser Arbeit drei weitere Techniken angewandt, die sich im Laufe der Jahre als Standardverfahren durchgesetzt haben:

1. *Dropout*: Durch zufälliges Deaktivieren einzelner Neuronen innerhalb einer Layer durch eine sogenannte *Dropout-Layer* (kurz: *Dropout*) während des Trainings wird nach Srivastava et al. die sogenannte *Co-Adaption* reduziert [45]. Sie beschreibt die Ausbildung von Abhängigkeiten zwischen mehreren Neuronen oder Layern – vergleichbar mit dem Prozess des Auswendiglernens anstelle der Ausbildung robuster und unabhängiger Merkmale. Anschaulich kann sich das Netz bei Anwendung des Dropouts nicht auf einzelne, hochgradig spezialisierte Neuronen verlassen, da diese, bedingt durch die zufällige Deaktivierung, nicht immer zur Verfügung stehen. Der relative Anteil der deaktivierten Neuronen durch eine Dropout-Layer wird im Folgenden beispielhaft für den Wert 10 % durch die Notation $\text{Dropout}^{0.1}$ ausgedrückt.
2. *Weight Decay*: Der oben genannte Optimierer AdamW nutzt das sogenannte *Weight Decay*, einen Mechanismus, der in seiner Grundidee die Werte der trainierbaren Parameter θ abhängig von einem Zerfallsfaktor, dem namensgebenden Parameter *Weight Decay*, reduziert [42]. Der Zerfallsfaktor beträgt beispielsweise 0.99 und wird multiplikativ auf alle trainierbaren Parameter θ angewandt, sodass kleine Werte einem geringen Zerfall unterliegen, während große Werte deutlich angepasst werden. Somit wird das neuronale Netz indirekt gezwungen, große Werte in θ zu vermeiden, die häufig ein Symptom von Overfitting sind.
3. *Learning Rate Scheduling*: Bei der Verwendung des in Abschnitt 3.2 genannten AdamW-Optimierers werden intern, neben der globalen Lernrate ϵ , für jeden trainierbaren Parameter adaptive Lernraten verwendet. Trotz dieser Eigenschaft wurde von Loshchilov et al. gezeigt, dass die Variation der globalen Lernrate ϵ abhängig von der aktuellen Epoche, das sogenannte *Learning Rate Scheduling*, den Trainingserfolg mit dem AdamW-Optimierer begünstigen kann [42]. Konkret verwenden Loshchilov et al. das sogenannte *Cosine Annealing Warm Restarts*, das die Lernrate über die erste halbe Periode der Cosinus-Funktion zwischen einem Start- und Endwert abklingen lässt (*Cosine Annealing*) und anschließend die Lernrate auf den Startwert zurücksetzt, um den gleichen Zyklus zu wiederholen (*Warm Restarts*) [46, 42]. Abbildung 3.3 stellt einen möglichen Verlauf der Lernrate mit dem Cosine Annealing Warm Restarts mit insgesamt vier Lernzyklen dar. Weiterhin kann zu Beginn des Trainings die Lernrate einmalig, linear hochgefahren werden, das sogenannte *Warm-Up*, sodass das Training mit einer insgesamt höheren Lernrate ϵ beschleunigt und stabilisiert wird [47, 48]. Das Warm-Up wird ebenfalls in Abbildung 3.3 dargestellt.

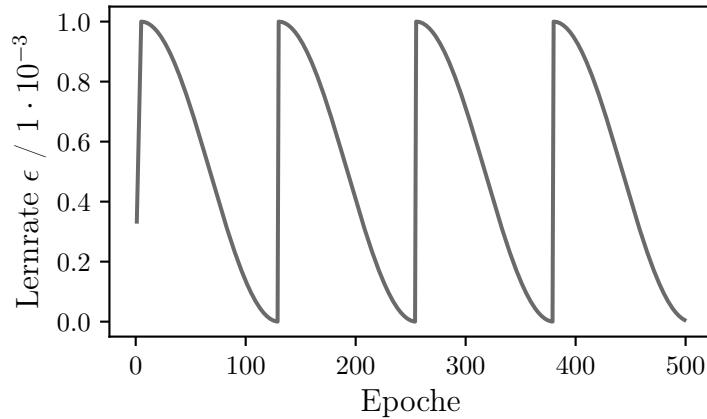


Abbildung 3.3: Beispielhafter Verlauf der Lernrate ϵ abhängig von der Epoche für das *Cosine Annealing Warm Restarts* mit *Warm-Up*.

3.4 DEEP LEARNING FÜR DIE BILDVERARBEITUNG

Für die Verarbeitung von (CT-Schicht-)Bildern ist das aus dem vorangehenden Abschnitt bekannte MLP nicht gut geeignet, wie die folgende Ausführung zeigt. Zunächst ändert sich die Datenmodalität: Bilder bestehen im Allgemeinen aus C (Farb-)Kanälen und weisen die Dimensionen H, W auf. Somit haben die Inputs die Form (C, H, W) .

Theoretisch könnte das MLP aus Abbildung 3.1 so angelegt werden, dass es $C \cdot H \cdot W$ Neuronen in der Input Layer besitzt. Für ein RGB-Bild der Form $(C, H, W) = (3, 256, 256)$ werden demnach fast 200000 Neuronen in der Input Layer benötigt. Selbst wenn die erste Hidden Layer nur aus zehn Neuronen besteht, gibt es in Summe bereits etwa zwei Millionen trainierbare Parameter θ . Ohne explizite Berechnung wird klar, dass bereits mit nur wenigen Hidden Layers und einer Output Layer die Kapazität eines derartigen neuronalen Netzes extrem groß wird und es für praktische Anwendungen unbrauchbar ist [49]. Einen Ausweg bietet eine Methode aus der klassischen Bildverarbeitung – die Faltung mit Bildfiltern. In der klassischen Bildverarbeitung kann beispielsweise die Kantenerkennung mit dem Sobel-Operator [50] explizit als Faltung eines Bildes mit einer Faltungsmatrix ausgeführt werden. Dieses Konzept lässt sich verallgemeinern, indem die Einträge der Faltungsmatrizen (auch: *Kernel*) trainierbare Parameter sind und somit die von der Faltung (englisch: *Convolution*) durchgeführte Operation für eine spezifische Aufgabe trainiert werden kann. LeCun et al. zeigen für die Handschrifterkennung einzelner Zeichen, dass neuronale Netze, basierend auf trainierbaren Convolutions, sogenannte Convolutional Neural Networks (CNNs), den MLPs überlegen sind [49]. In Kombination mit den Convolutions einer festen Kernelgröße 3×3 verwendet LeCun in seiner vorgestellten Architektur *LeNet-5* sogenannte *Pooling Layer*, die die räumlichen Dimensionen eines Bildes während der Verarbeitung schrittweise verkleinern. Abbildung 3.4 zeigt einen Tensor der Form $(1, 4, 4)$, der durch Anwendung einer Pooling-Layer, dem sogenannten *Maxpool*, auf die Form $(1, 2, 2)$ reduziert wird. Für eine feste Größe der Kernel wird durch das Pooling das Sichtfeld der Convolutions, das auch *rezeptives Feld* genannt wird, relativ zur Gesamtgröße des Bildes vergrößert.

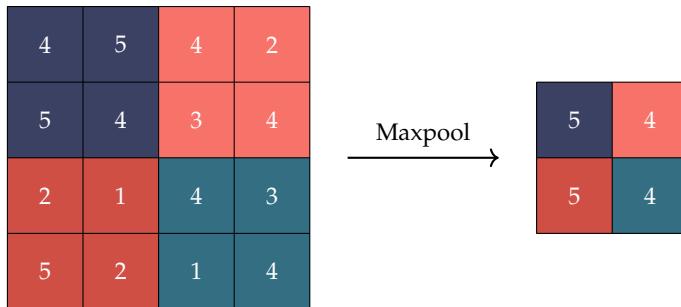


Abbildung 3.4: Funktionsskizze einer Maxpool Layer zur Halbierung der räumlichen Bilddimensionen durch Extraktion des Maximalwertes aus einem Fenster der Größe 2×2 .

Das Forschungsfeld der CNNs wurde in den folgenden Jahren um viele Architekturen erweitert. Eine der am häufigsten verwendeten Architekturen für die Bildverarbeitung ist das *ResNet* [51], das zusätzlich sogenannte *Residual Connections* oder *Skip Connections*, verwendet. Die Hintergründe dieser Skip Connections können in der Publikation von He et al. nachvollzogen werden [51].

Für die in Kapitel 1 motivierte Materialbestimmung sollen sowohl die Kernladungszahl als auch die Dichte pro Pixel, beziehungsweise im CT-Schichtbild pro Voxel, anhand von DECT-Bildern vorhergesagt werden. Somit müssen die Eingaben (Low-Energy und High-Energy) und Ausgaben (Kernladungszahl und Dichte) eines passenden neuronalen Netzes Bildern der Form $(C, H, W) = (2, H, W)$ entsprechen. Diese Form der Vorhersage wird auch *Dense Prediction* genannt, da für jedes Pixel der Eingabe ein Pixel in der Ausgabe berechnet wird. Eine der bekanntesten Architekturen für Dense Prediction ist das U-Net, welches in Abschnitt 3.4.1 vorgestellt wird. CNNs, die viele Jahre die Bildverarbeitung dominiert haben, werden zunehmend von sogenannten *Transformern* ersetzt. Die konzeptionelle Idee hinter der Transformer-Architektur, und wie diese für die Bildverarbeitung genutzt werden kann, wird in Abschnitt 3.4.2 beschrieben.

3.4.1 U-Net

Abbildung 3.5 stellt die Architektur des U-Nets anschaulich mit seiner namensgebenden, U-förmigen Struktur dar. Nach Ronneberger et al. besteht das U-Net aus einem sogenannten *Encoder* und einem sogenannten *Decoder*. Der Encoder verarbeitet eine Eingabe, ein Graustufenbild, der Form $(1, H, W)$ durch die Anwendung von Convolutions und Pooling in den sogenannten *Latent Space*, der wiederum die Form $(1024, \frac{H}{16}, \frac{W}{16})$ aufweist. Dieser Prozess verläuft auf mehreren, sogenannten *hierarchischen* Ebenen, sodass *hierarchische Feature-Maps*⁴ (C, H, W) mit den Dimensionen $(64, H, W), (128, \frac{H}{2}, \frac{W}{2}), (256, \frac{H}{4}, \frac{W}{4}), (512, \frac{H}{8}, \frac{W}{8})$ und $(1024, \frac{H}{16}, \frac{W}{16})$ berechnet werden (vergleiche Abbildung 3.5). Die Anzahl der Kanäle in der ersten Feature-Map $(64, H, W)$ wird im Folgenden *Feature-Map-Tiefe* genannt und ist der bestimmende Faktor für die Gesamtkapazität der Architektur. Anschließend dient der Latent Space als Startpunkt des Decoders, der, als Umkehrung

⁴ Der Begriff *hierarchische Feature-Map* bezeichnet hier konkret einen Tensor, der das letzte Element, das Zwischenergebnis, einer hierarchischen Ebene in einem CNN bildet. In der U-Net-Architektur (vergleiche Abbildung 3.5) werden die hierarchischen Feature-Maps des Encoders in den Decoder mittels Skip Connections übertragen.

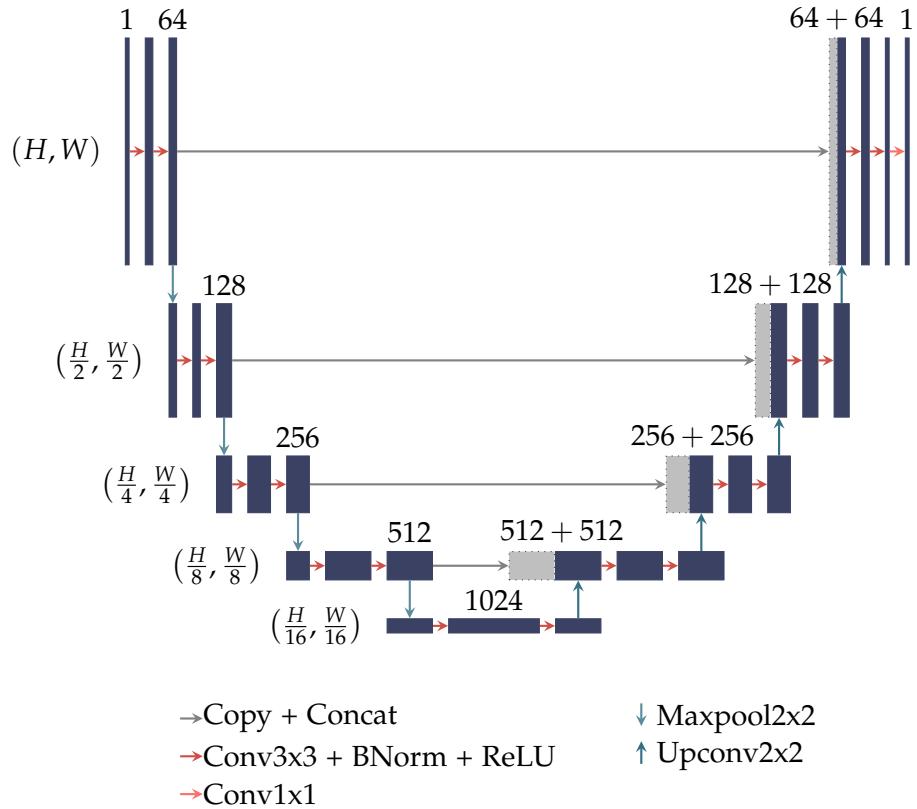


Abbildung 3.5: Architekturdiagramm des *U-Net* nach Ronneberger [36] mit den räumlichen Dimensionen (H, W) der Tensoren. Die Anzahl der Kanäle wird über den Tensoren angegeben.

zum Encoder, die räumlichen Dimensionen schrittweise erhöht und die Feature-Maps aus dem Encoder durch Skip Connections (vergleiche *Skip Connection* in Abschnitt 3.4) mitverarbeitet (Copy+Concat). Zur Erhöhung der räumlichen Dimensionen, als imperfekter Umkehrprozess des Poolings, werden sogenannte *Upconvolutions* oder *Transposed Convolutions* verwendet. Anschaulich ist die Transposed Convolution eine umgekehrte Convolution, das heißt sie führt die Erweiterung eines Pixels auf einen größeren Bereich aus. Die Skip Connections zwischen dem Encoder und dem Decoder helfen bei der Wiederherstellung der räumlichen, positionsabhängigen Informationen, die teilweise durch das Pooling im Encoder verloren gegangen sind.

Ronneberger entwickelte das U-Net ursprünglich zur Segmentierung medizinischer Bilddaten [36]. Zwischenzeitlich wurde das U-Net für unterschiedliche Anwendungsbereiche und Lernprobleme angewendet: zur Segmentierung von Kamerabildern aus fahrenden Autos [52], zur Erkennung von Gebäuden [53] in Satellitenbildern, zur Klassifizierung von Ackerland [54] in Satellitenbildern und zum Schätzen einer räumlichen Tiefe in Kamerabildern [55].

3.4.2 Swin Transformer

Ursprünglich entwickelt für das Natural Language Processing (NLP) [56], dringen sogenannte *Transformer* zunehmend in neue Anwendungsgebiete vor. Das Funktionsprinzip von Transformern wird im Folgenden vereinfacht erklärt, sodass ein Grundverständnis

erlangt werden kann. An entsprechenden Stellen wird auf die Literatur verwiesen, sodass bei Bedarf weitere Informationen eingesehen werden können.

Transformer basieren nicht auf Convolutions, sondern verwenden eine Kombination aus der sogenannten *Self-Attention* und MLPs [56]. Anhand eines einfachen Beispiels kann die Zielsetzung der Self-Attention illustriert werden: Gegeben sei der Satz

Moritz scannt sein Mittagessen

zusammen mit qualitativ bestimmten Self-Attention-Werten in Tabelle 3.1. Genaue Details zur Berechnung der Self-Attention sind für ein konzeptionelles Verständnis der Architektur nicht notwendig und können ergänzend bei Vaswani et al. eingesehen werden [56]. Für jede mögliche Permutation der Wortpaare wird ein Self-Attention-Wert berechnet, der bei der weiteren Verarbeitung ausgewählten Teilen des Satzes eine höhere Relevanz zuordnet. Anschaulich ist das Wortpaar *Moritz-sein* für die Gesamtbedeutung des Satzes wesentlich unwichtiger als das Wortpaar *Moritz-scannt*. Die Self-Attention ist somit ein Maß für die Verknüpfung relevanter Teile der Eingabe. Es fällt zudem auf,

Tabelle 3.1: Exemplarische Self-Attention-Werte für den Satz „Moritz scannt sein Mittagessen“. Die tabellierten Werte sind rein illustrativ und entstammen keiner realen Berechnung.

	Moritz	scannt	sein	Mittagessen
Moritz	0.10	0.30	0.20	0.40
scannt	0.20	0.10	0.40	0.30
sein	0.15	0.35	0.10	0.40
Mittagessen	0.25	0.25	0.25	0.25

dass die Matrix der Self-Attention-Werte nicht symmetrisch ist. Der phänomenologische Grund hierfür besteht in der Tatsache, dass die Reihenfolge der einzelnen Wörter relativ zueinander für die Gesamtaussage des Satzes wichtig ist. Jedes Wort (jede Zeile der Tabelle) berechnet die Self-Attention gepaart mit allen Wörtern des Satzes (Spalten) in der gegebenen Reihenfolge. Die Summe der Zeilen ist somit stets eins. Mathematisch kann dieser Zusammenhang in den Gleichungen zur Self-Attention nach Vaswani et al. nachvollzogen werden [56].

Das klassisch von CNNs dominierte Feld der Computer Vision wurde erstmalig durch Dosovitskiy et al. für die sogenannten Vision Transformer (ViT) eröffnet [57]. Die zentrale Idee ist, ein Bild in sogenannte *Patches* zu unterteilen, die als einzelne Wörter eines Satzes (dem gesamten Bild) interpretiert werden. Somit kann der oben genannte Mechanismus zur Berechnung der Self-Attention nach Vaswani et al. auf die Patches angewandt werden. Im Gegensatz zu CNNs, deren rezeptives Feld durch die Größe der Convolution-Kernel definiert ist, kann der ViT durch die Berechnung der Self-Attention auf allen Patches (Bildbereichen) Informationen aus weit verstreuten Bildbereichen miteinander kombinieren. Jedoch muss dieser offensichtliche Vorteil des ViTs durch das Training auf sehr großen Datensätzen zunächst erlernt werden [57]. CNNs nutzen durch die Anwendung der Convolution automatisch die Bildinformationen aus einem lokalen, meist semantisch zusammenhängenden Bereich, der häufig wertvolle (Bild-)Informationen, wie zum Beispiel Objektkanten, enthält. ViTs müssen diese Beziehung benachbarter Bildeigenschaften sehr

aufwendig erlernen, worin einer der Gründe für den großen Bedarf an Trainingsdaten besteht. Eine hybride Lösung bietet der sogenannte *Swin Transformer* nach Liu et al. [37, 38], der die Patches in größere Gruppen, die sogenannten *Windows*, zusammenfasst und die Self-Attention nur über die Patches innerhalb eines Windows berechnet. Abbildung 3.6 (links) zeigt die Aufteilung eines Bildes, bestehend aus den Patches, in Windows. Durch

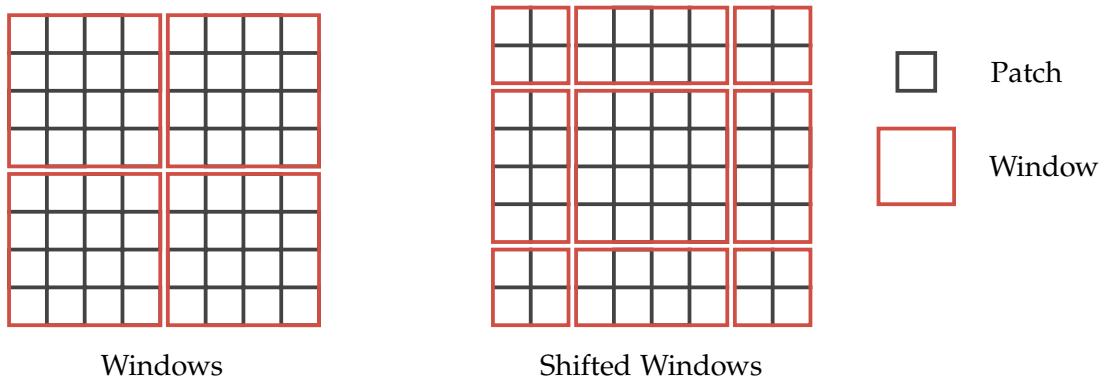


Abbildung 3.6: Positionierung der *Windows* und *Shifted Windows* zur Berechnung der Self-Attention über benachbarte Patches im Bild in der Swin-Transformer-Architektur. Die Abbildung entstammt in leicht veränderter Form der Publikation von Liu et al. [37].

diese Einschränkung des rezeptiven Feldes im Vergleich zum ViT, insbesondere wenn räumlich zusammenhängende Informationen an den Grenzen zwischen zwei Windows liegen, ist es sinnvoll, neben den Windows auch sogenannte *Shifted Windows* zu verwenden, die um eine halbe Window-Breite verschoben positioniert sind und in Abbildung 3.6 (rechts) dargestellt werden. Nach Liu et al. steigt der Berechnungsaufwand der Self-Attention beim Swin Transformer linear mit den Bilddimensionen [37], während dieser beim ViT quadratisch mit den Bilddimensionen anwächst [57]. Der Swin Transformer ist somit besser auf große Bilder skalierbar. Zudem wird ein Bild beim Swin Transformer in mehreren, hierarchischen Schichten verarbeitet, sodass hierarchische Feature-Maps, analog zu CNNs, entstehen. Anstelle der klassischen, parameterfreien Pooling Layer werden beim Swin Transformer trainierbare Layer wie beim MLP verwendet, um die räumlichen Dimensionen der Tensoren (Feature-Maps) schrittweise zu halbieren [37]. Durch Liu et al. wurden die Swin-Transformer-Modelle *Tiny* (Swin-T), *Small* (Swin-S), *Base* (Swin-B), *Large* (Swin-L), *Huge* (Swin-H) und *Giant* (Swin-G) vorgestellt, die sich hauptsächlich durch ihre, in der genannten Reihenfolge steigende, Kapazität unterscheiden [38]. Strukturell ist der Swin Transformer, aufgrund seiner hierarchischen Struktur, mit der Grundidee des U-Nets [36], speziell dem Encoder, kompatibel [37, 36, 58]. Weitere Details zur Umrüstung des U-Nets auf eine Transformer-basierte Architektur [58] folgen zusammen mit einer Gegenüberstellung beider Architekturen in Kapitel 6.4.

4

VERWANDTE FORSCHUNG

Dieses Kapitel basiert auf einer dieser Dissertationsschrift vorausgehenden Publikation [22], die den aktuellen Stand der Forschung sowie den Wandel der Materialbestimmung mit Deep Learning in CT-Daten in Form einer systematischen Literaturrecherche erfasst. Die Ergebnisse der Literaturrecherche stammen demnach aus dem Monat Juli im Jahr 2024. Eine Aktualisierung der Suchergebnisse im Februar des Jahres 2025 konnte neue Publikationen identifizieren, die zwar die bereits gesammelten Erkenntnisse untermauern, jedoch keine neuen Erkenntnisse hinzufügen. Somit sind die im Folgenden vorgestellten Ergebnisse weiterhin gültig und können als verwandte Forschung in dieser Dissertationsschrift genutzt werden.

Im Jahr 1976 legt Alvarez den Grundstein für die Materialzerlegung in CT-Schnittbildern [13]. Er beschreibt den linearen Abschwächungskoeffizienten μ_L als Linearkombination aus einem Term zur Beschreibung des Photoeffektes und einem Term zur Beschreibung der Compton-Streuung mit den materialspezifischen Konstanten a_1 und a_2

$$\mu_L^{(x,y,z)}(E) = \underbrace{a_1^{(x,y,z)} \cdot \frac{1}{E^3}}_{\text{Photoeffekt}} + \underbrace{a_2^{(x,y,z)} \cdot f_{\text{KN}}(E)}_{\text{Compton-Streuung}} \quad (4.1)$$

an einer Position (x, y, z) im Voxelgitter des rekonstruierten Volumens. Durch die Messung von μ_L für mindestens zwei verschiedene Energien E kann die Gleichung nach a_1 und a_2 aufgelöst werden. Alvarez tabelliert für die in seiner Publikation verwendeten CT-Parameter die charakteristischen Werte von a_1 und a_2 für Gehirngewebe und Fettgewebe. Etwa 25 Jahre später publiziert Heismann eine Erweiterung des Ansatzes von Alvarez, indem er eine Gleichung vorstellt, die jedem Voxel in einem DECT-Schichtbild eine Kernladungszahl und eine Dichte zuordnet – die sogenannte $Z\text{-}\rho$ -Zerlegung [14]. Ausgehend von Gewichtungsfunktionen $w_n(E)$, die die Eigenschaften des Röntgenspektrums und des Detektors abbilden, und tabellierten Massenschwächungskoeffizienten $\frac{\kappa}{\rho}$ aus der Literatur, gilt nach Heismann [14]

$$F(Z) = \frac{f_1(Z)}{f_2(Z)} = \frac{\mu_1}{\mu_2} \quad \text{mit} \quad f_n = \int w_n(E) \left(\frac{\kappa}{\rho} \right) (E, Z) dE \quad \text{und} \quad n \in \{0, 1\}. \quad (4.2)$$

Zur Ermittlung der Kernladungszahl Z aus dieser Gleichung genügt die numerische Bestimmung der Inversen F^{-1} . Nach Heismann wächst die Funktion $F(Z)$, in dem von ihm beschriebenen Energie- und Materialbereich, monoton, sodass die Bestimmung der Inversen F^{-1} durch Interpolation erfolgen kann [14].

Ein fundamentaler Unterschied zwischen den beiden vorgestellten Grundlagenarbeiten von Alvarez und Heismann besteht darin, dass die $Z\text{-}\rho$ -Zerlegung (nach Heismann) effektive Kernladungszahlen und Dichten berechnet, während die Basiszerlegung (nach Alvarez) diskrete Anteile an den gewählten (Material-)Basisvektoren ermittelt. In der

Basiszerlegung werden die Basisvektoren vorgegeben – zum Beispiel zwei Materialien wie Aluminium und Eisen. Das Ergebnis gibt dann die Anteile der Basisvektoren an einem bestimmten Voxel an – zum Beispiel 20% Aluminium und 80% Eisen. Mit dem nötigen Vorwissen ist die Basiszerlegung somit ein mächtiges Werkzeug, das auch auf drei oder mehr Basisvektoren ausgeweitet werden kann, wie im späteren Verlauf dieses Kapitels gezeigt wird.

Bei der Z - ρ -Zerlegung werden effektive Werte, das heißt im Fall einer Mischung (oder Legierung) Mittelwerte, ohne Vorwissen über die vorliegende Probe berechnet. Das soeben genannte Beispiel einer Zusammensetzung aus 20% Aluminium und 80% Eisen erscheint in der Z - ρ -Zerlegung nach der Gleichung von Heismann [14] als

$$Z_{\text{eff}}(\text{Al}_{0.2}\text{Fe}_{0.8}) = \left(\frac{\sum_i Z_i^3 \cdot \rho_i}{\sum_i \rho_i} \right)^{\frac{1}{3}} \approx 25.4$$

und

$$\rho_{\text{eff}} = 0.2 \cdot \rho(\text{Al}) + 0.8 \cdot \rho(\text{Fe}) = 6.8 \text{ g/cm}^3 .$$

Beide Methoden weisen somit Vor- und Nachteile auf und sind in der Literatur weit verbreitet, wie in den folgenden Abschnitten deutlich wird.

4.1 METHODIK DER LITERATURRECHERCHE

Nach der Definition von vom Brocke muss eine systematische Literaturrecherche reproduzierbar und umfassend sein sowie einen Erkenntnisgewinn liefern [59]. Hierzu nennt vom Brocke zwei verbreitete Suchmaschinen: *Scopus* und *Web of Science* [59]. Für diese Arbeit wird die Suchmaschine *Scopus* von Elsevier verwendet, da sie nach Pranckute eine breitere Abdeckung interdisziplinärer Themen, zu denen der Inhalt dieser Arbeit ebenfalls zählt, im Vergleich zu *Web of Science* bietet [60]. Konkret erfolgt die Literaturrecherche anhand sogenannter *Queries Q*. Ein Query kann durch einfache *Keywords* oder durch spezielle Operatoren erstellt werden. Für den unten genannten Query werden insgesamt zwei Operatoren verwendet:

1. TITLE-ABS-KEY(<str>) filtert nach dem String str in den Titeln, Abstracts und Keywords der Publikationen,
2. REF(<str>) filtert nach dem String str in den Quellenangaben der Publikationen.

Die Kombination mit einer Jahreszahl ist bei dem REF-Operator in der Form REF(Weiss 2024) ebenfalls möglich. Für die Suche nach materialauflösenden CT-Verfahren wird der Query aus drei Bestandteilen zusammengesetzt:

1. TITLE-ABS-KEY(computed tomography OR ct)
2. TITLE-ABS-KEY(material)
3. TITLE-ABS-KEY(dual-energy OR multi-energy OR photon-counting).

Die ersten beiden Bestandteile des Queries werden sofort durch das Thema dieser Dissertationsschrift motiviert. Verfahren, die nicht auf mehreren Energiekanälen basieren, werden durch den dritten Bestandteil des Queries bewusst ausgeschlossen, da, gemäß der Einführung der Materialbestimmung in Kapitel 2, mehrere Energiekanäle für die Materialerkennung benötigt werden [13, 14]. Die Anzahl der mit dem Query gefundenen Publikationen wird in Abbildung 4.1 aufgelöst nach dem Publikationsjahr graphisch dargestellt (blaue Linie). In der Gesamtheit kann ein steigender Trend der Publikationszahlen

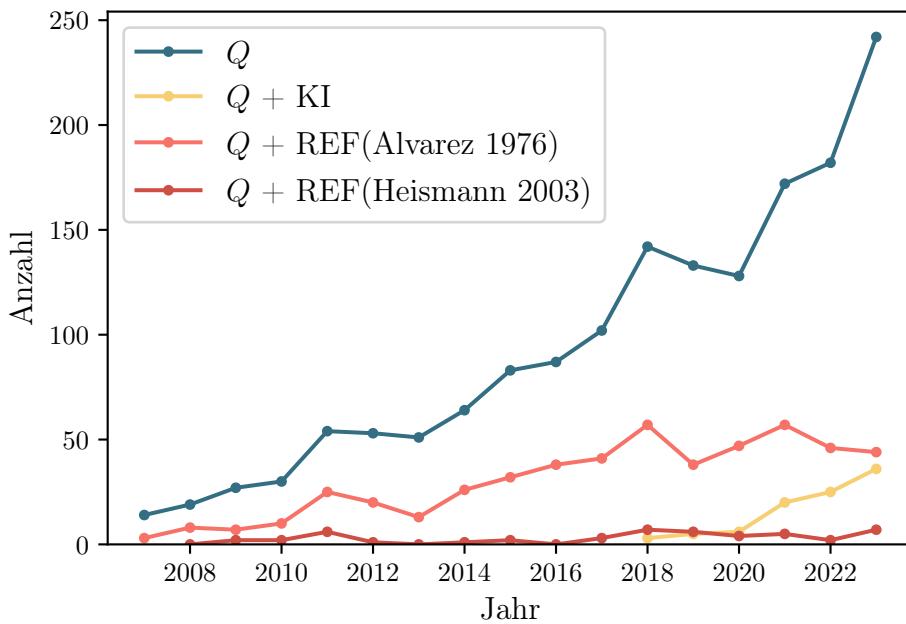


Abbildung 4.1: Anzahl der Publikationen unter Verwendung des genannten Queries bei Scopus. Mit den Zusatzfiltern REF(Alvarez 1976) und REF(Heismann 2003) werden Publikationen gefiltert, die ebenjene Autoren zitieren [13, 14].

mit dem Query beobachtet werden (blaue Linie). Durch weitere Verfeinerung des Queries nach KI-basierten Verfahren ($Q + \text{KI}$), explizit aufgebaut aus *artificial intelligence*, *machine learning* oder *deep learning*, ergibt sich die gelbe Linie in Abbildung 4.1. Mit der Vorstellung des U-Nets [36] wird im Jahr 2015, wie im weiteren Verlauf dieses Kapitels gezeigt wird, der Grundstein für die meisten DL-basierten Materialerkennungsverfahren nur wenige Jahre vor den ersten Publikationen zur Materialbestimmung mit Deep Learning gelegt. Alternativ werden mit den REF-Operatoren die Publikationen gefiltert, die die fundamentalen, klassischen Verfahren nach Alvarez [13] und Heismann [14] zur Materialerkennung in den Quellenangaben referenzieren (rote Linien). Während vor dem Jahr 2018 klassische Materialbestimmungsverfahren das Feld dominieren, steigt die Anzahl der Zitationen der klassischen Basismaterialzerlegung nach Alvarez an. Zu beobachten ist weiterhin, dass seit Einführung der KI-basierten Materialbestimmungsverfahren, die Anzahl der Zitationen von Alvarez stagniert. Es besteht, ohne weitere Begründung, die Vermutung, dass sich moderne, KI-basierte Verfahren weniger auf Wissen über die klassische, physikbasierte Materialzerlegung beziehen, während klassische Verfahren auf derartiges Domänenwissen angewiesen sind.

Zur Verbesserung der Vergleichbarkeit der verschiedenen Ansätze aus dieser Vielzahl an Publikationen müssen weitere Kriterien erfüllt werden. Zunächst muss der Zugriff

auf eine ausgewählte Publikation mindestens für akademische Einrichtungen möglich sein, ohne einzelne Publikationen gezielt erwerben zu müssen. Viele publizierte Ansätze basieren auf vorausgehenden Arbeiten, sodass versucht wird, Publikationen mit einem neuartigen Ansatz oder mindestens einer signifikanten Weiterentwicklung eines bestehenden Ansatzes zu selektieren. Hierbei muss zwischen klassischen und datengetriebenen Ansätzen differenziert werden. Klassische Ansätze basieren auf den Grundlagenarbeiten von Alvarez [13] und Heismann [14]. Somit grenzen sich substantielle Weiterentwicklungen im Feld der klassischen Algorithmen beispielsweise durch die Erweiterung des Ansatzes von Heismann auf den Energiebereich $> 1 \text{ MeV}$ unter zusätzlicher Berücksichtigung der Paarbildung (vergleiche Kapitel 2.1) ab. Datengetriebene Verfahren müssen entweder eine neue Modellarchitektur, Methodik der Trainingsdatenerzeugung oder generell neuartige Trainingsstrategien vorstellen. Die Architekturen müssen anhand der Beschreibung in der Publikation von versierten DL-Entwicklern nachvollziehbar und reproduzierbar sein. Zudem existieren nur wenige offene Benchmarks zum quantitativen Vergleich der Ansätze untereinander, sodass der jeweils verwendete Datensatz im Bestfall entweder mitpubliziert worden ist oder mithilfe einer Open-Source-Simulation und einem bekannten Parametersatz nachgestellt werden kann. Einige Daten, insbesondere aus dem klinischen Umfeld, unterliegen der Geheimhaltung, sodass in diesen Fällen mindestens eine ausführliche Beschreibung der Datenart und Datenzusammensetzung vorliegen muss. Verschiedene Röntgenspektren und Detektortechnologien, EID und PCD, stellen unterschiedliche Herausforderungen an die Materialbestimmung oder -zerlegung und müssen erfasst werden. Daher müssen die Konfiguration der CT-Anlage, in Bezug auf die Röntgenstrahlungsquelle und den Detektor, sowie die Aufnahmeparameter (Integrationszeit, Vergrößerung, Trajektorie) nachvollziehbar dokumentiert sein. Einige der gefundenen Publikationen verwenden nur einen Energiekanal eines PCDs, sodass abschließend manuell geprüft werden muss, ob die Ansätze auf mehreren Energiekanälen basieren. Zusammenfassend muss demnach eine Publikation die folgenden Filter erfüllen, um in den Korpus aufgenommen zu werden:

1. Der Zugriff erfolgt über einen der großen Verlage;
2. Der vorgestellte Ansatz muss entweder neuartig oder eine signifikante Weiterentwicklung eines bestehenden Ansatzes in mindestens einem der folgenden Aspekte sein:
 - 2.1. Physikalische Modellierung des CT-Messprozesses (insbesondere bei klassischen Algorithmen);
 - 2.2. Aufbau der DL-Architektur, Methodik der Datenerzeugung oder Trainingsstrategie.
3. Der CT-Scan muss in Bezug auf die Scanparameter und die Anlagenkonfiguration so beschrieben sein, dass er reproduziert werden kann.
4. Der vorgestellte Ansatz muss mindestens zwei Energiekanäle verwenden – eine Grundvoraussetzung, die durch Alvarez [13] und Heismann [14] für die Materialbestimmung vorgegeben wird.

Anhand der soeben genannten Filter resultiert ein Korpus, bestehend aus 24 Publikationen, wobei drei Publikationen industrielle CT-Anwendungsfälle beschreiben. Dieses starke Ungleichgewicht wird verbessert, indem die *iCT conference proceedings* seit dem Jahr 2016 zusätzlich durchsucht werden. Durch die ursprüngliche Suche wurden diese Publikationen nicht erfasst, da sie nicht bei *Scopus* indiziert werden. Die *iCT* ist eine der größten Konferenzen zur industriellen Computertomographie in Europa, sodass relevante Fortschritte mit hoher Wahrscheinlichkeit in diesem Rahmen publiziert werden. Aus den *iCT conference proceedings* konnten zwei Publikationen identifiziert werden, die die oben genannten Filter erfüllen, sodass der Korpus final 26 Publikationen umfasst. Der Korpus teilt sich in industrielle und klinische CT sowie in klassische Algorithmen und datengetriebene Verfahren (KI) auf. Eine Übersicht dieser Kategorien ist in Abbildung 4.2 dargestellt.

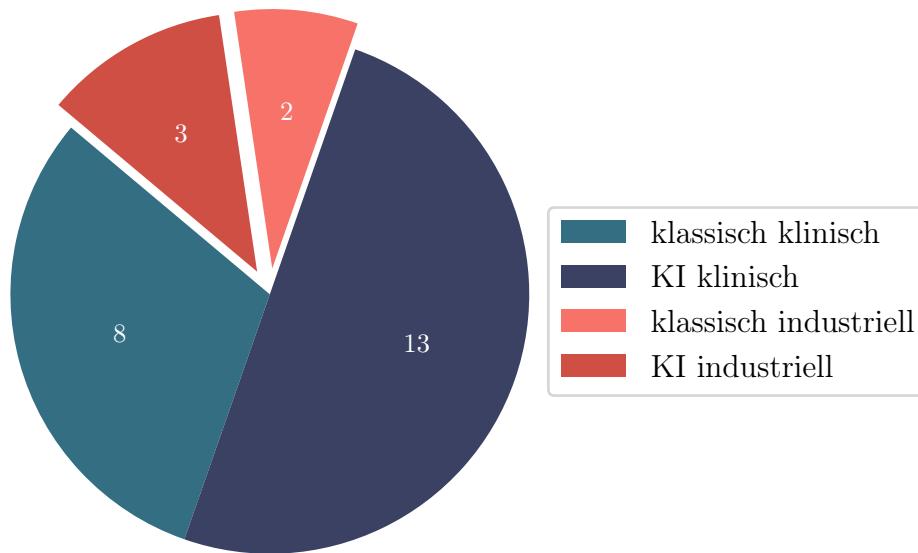


Abbildung 4.2: Kreisdiagramm zur anschaulichen Darstellung der Anteile klassischer und KI-basierter Publikationen für die klinische und industrielle Materialbestimmung und -zerlegung in CT-Bildern. Die Zahlenwerte in den Anteilen geben die absolute Anzahl der Publikationen in der jeweiligen Kategorie an.

4.2 ANALYSE DES KORPUS

Eine Übersicht der Publikationen im Korpus mit einer chronologischen Einordnung wird in Abbildung 4.3 dargestellt. Relevante Entwicklungen aus dem Umfeld der CT und des Deep Learnings sind chronologisch ergänzt. Bemerkenswert ist, dass Alvarez seinen Ansatz zur Basismaterialzerlegung [13] bereits fünf Jahre nach dem ersten CT-Scan, und damit auch zeitlich vor dem Nobelpreis von Hounsfield, publiziert hat. Zudem gibt es erst seit 2021 kommerzielle CT-Scanner, die im klinischen Umfeld PCDs verwenden, obwohl der Grundstein für die Technologie mit dem Medipix2 Detektor im Jahr 2007 gelegt worden ist. Die grundlegenden Arbeiten zu modernen CV-Architekturen wurden mit dem U-Net 2016 [36] und dem ViT 2019 [57] veröffentlicht.

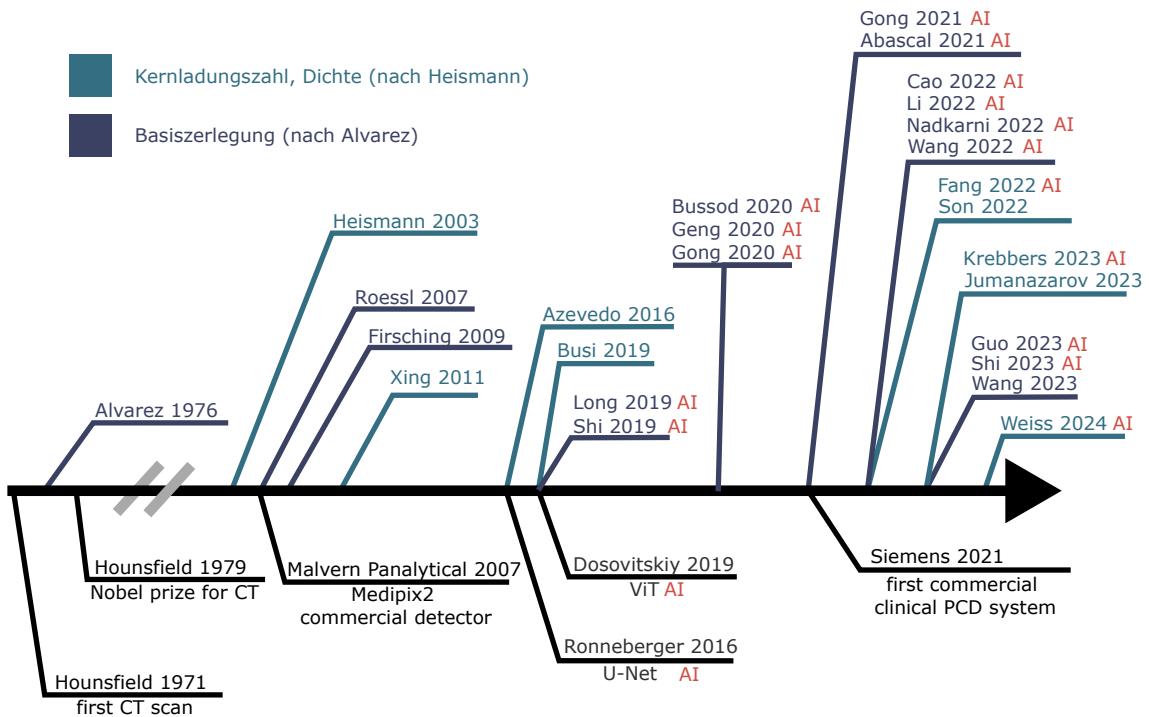


Abbildung 4.3: Zeitstrahl der Publikationen aus dem Korpus (oben) und relevanter Entwicklungen aus der CT- und DL-Forschung (unten). DL-basierte Einträge werden durch den roten Schriftzug *AI* markiert.

4.2.1 Beschleunigungsspannung und Detektortechnologie

Bei der Messung einer DECT können verschiedene Aufnahmestrategien verwendet werden, wie in Abschnitt 2.6 beschrieben worden ist. Im Rahmen dieser Dissertationsschrift wird zwischen integrierenden (EID) und zählenden (PCD) Detektoren unterschieden. Zählende Detektoren bieten den Vorteil, mehrere Energiekanäle gleichzeitig zu messen, während integrierende Detektoren in der Regel mehrere Einzelscans mit verschiedenen Quellparametern nacheinander aufzeichnen.¹ Abbildung 4.4 zeigt einen Überblick der verwendeten Quellspannungen und Detektortechnologien aus dem Korpus. Zunächst kann keine Einschränkung der klassischen oder datengetriebenen Materialbestimmung anhand der Quellspannung, der Detektortechnologie oder der Anwendungsdomäne identifiziert werden.

Es konnten fünf Publikationen identifiziert werden, die klassische Materialbestimmung auf PCD-Daten anwenden [61, 62, 16, 63, 64]. Roessl verwendet einen PCD mit acht äquidistanten Bins zwischen 10 keV und 80 keV mit einer Beschleunigungsspannung von 90 kV [61]. Son nutzt eine Quellspannung von 125 kV und teilt das Spektrum in zwei Bins $b_{1,2}$ mit $20 \text{ keV} \leq b_1 \leq 60 \text{ keV}$ und $60 \text{ keV} \leq b_2 \leq 125 \text{ keV}$ [62]. Die niedrigsten Quellspannungen aus dem Korpus werden bei Wang und Firsching mit 50 kV dokumentiert [16, 64]. Besonders hervorzuheben ist der Medipix2 Detektor (PCD) aus der Arbeit von Firsching, der 41 Energiekanäle zwischen 7.3 keV und 53.3 keV auflöst [64]. Jumanazarov erprobt die Verwendung von zwei, sechs und fünfzehn Energiekanälen, wobei sechs Energiekanäle die besten Ergebnisse liefern [63].

¹ Zur Vereinfachung wird hier auf eine weitere Aufschlüsselung der Messstrategie (Multi-Scan, Sandwich, Potential Switching, Multi-Tube) bei EIDs, wie in Abschnitt 2.6 angedeutet, verzichtet.

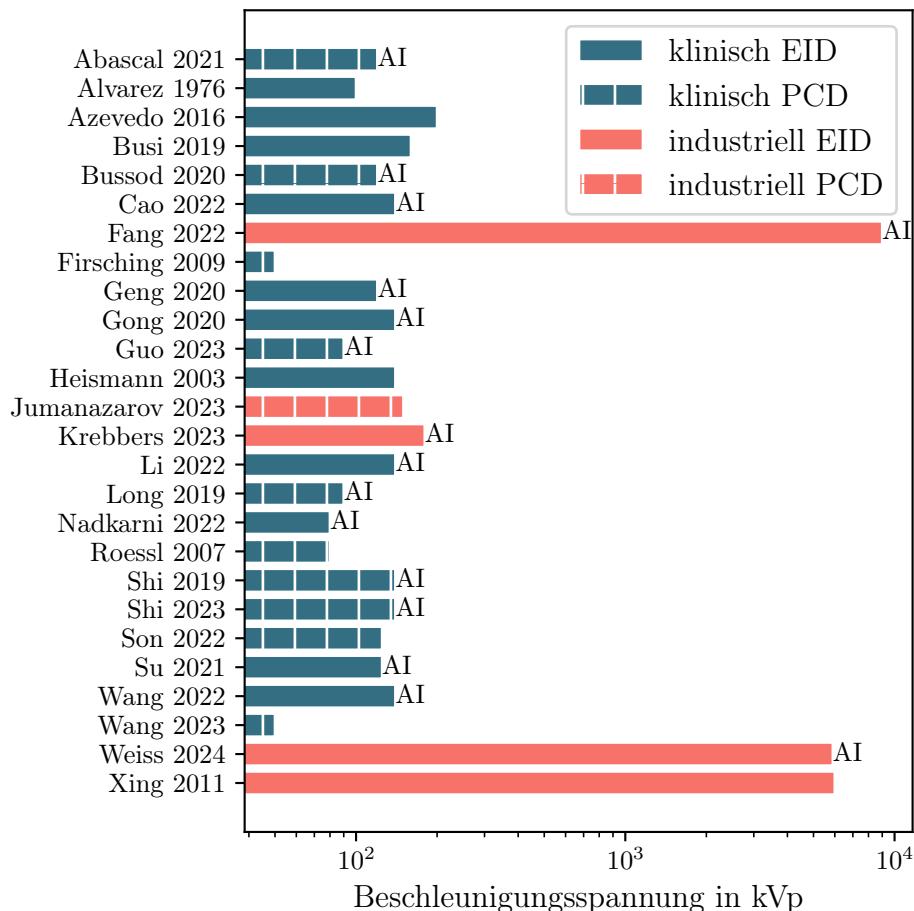


Abbildung 4.4: Verwendete Beschleunigungsspannungen der Publikationen aus dem Korpus. Die Farbe kennzeichnet die Anwendungsdomäne. Gestrichelte Balken stellen illustrativ die Verwendung eines PCDs dar. Die Position des Strichmusters gibt keine Auskunft über die Verteilung der Energiekanäle des PCDs. Methoden, die beide Detektortechnologien verwenden, sind mit durchgehenden Balken beschriftet. DL-basierte Ansätze sind durch den Schriftzug *AI* am rechten Ende des Balkens markiert.

Weiterhin wurden sechs Publikationen gefunden, die PCD-Daten mit datengetriebenen Ansätzen verarbeiten [65, 66, 67, 68, 69, 70]. Shi verwendet zwei Energiekanäle [70], Shi drei Energiekanäle [69], Bussod vier Energiekanäle [66], Abascal fünf Energiekanäle [65] sowie Long und Guo zwölf Energiekanäle [67, 68].

Heismann formuliert seinen klassischen Ansatz zur $Z\text{-}\rho$ -Zerlegung mit einem EID für einen DECT-Scan bestehend aus zwei Scans mit 80 kV und 140 kV Quellspannung. Xing erweitert Heismanns Ansatz auf den Energiebereich bis 6 MV ebenfalls für integrierende Detektoren [71].

Darüber hinaus konnten sieben Publikationen gefunden werden, die datengetriebene Methoden auf EID-basierte CT-Scans anwenden [11, 72, 73, 74, 75, 76, 77]. Drei der Publikationen nutzen hybride Ansätze, die sowohl EID- als auch PCD-Daten prozessieren [78, 79, 80]. Häufig führt die Nutzung eines EIDs zu einer besseren Bildqualität, während die Strahlenexposition des Patienten (der Probe) stark erhöht ist. In der klinischen Anwendung ist es daher wünschenswert, effiziente PCDs zu verwenden, um die Strahlenexposition zu senken, obwohl die Bildqualität vermindert ist. Ziel dieser hybriden Ansätze ist, die PCD-Daten als Inputs und die EID-Daten als Ground Truths zu nutzen, sodass die trainierten Modelle die Bildqualität der PCD-Daten verbessern [78, 80].

4.2.2 Materialien

Die Materialbestimmung oder -zerlegung ist für verschiedene Materialien aufgrund des jeweils charakteristischen Abschwächungskoeffizienten eine unterschiedlich komplizierte Aufgabe. Neben dem Auftreten von CT-Artefakten (vergleiche Kapitel 2.4.1) durch stark abschwächende Materialien spielt die Wahl einer geeigneten Basis für die Basismaterialzerlegung eine zentrale Rolle. Daher ist eine Differenzierung der verwendeten Materialien in den Publikationen des Korpus in Abhängigkeit der verwendeten Methode, klassisch oder datengetrieben, von großem Interesse. Zunächst werden die verwendeten Probenmaterialien aus den Publikationen des Korpus erfasst und nach der jeweils verwendeten Methode kategorisiert. Abbildung 4.5 bietet einen graphischen Überblick der Materialien. Aufgrund des Übergewichts der Publikationen aus dem klinischen Anwendungsgebiet, sind größtenteils klinisch-diagnostisch relevante Materialien vertreten. Hierzu zählen in erster Linie Knochen (*bone*) und Wasser (*water*). Calcium (*Ca*) ist zum Beispiel als Bestandteil von Hydroxylapatit in (menschlichen) Knochen und Zähnen zu finden. Aufgrund der geringen Dichte von Knochenmasse, ist sie auch für niedrige Photonenenergien ab 50 keV zur Bildgebung ausreichend transparent. Weiterhin werden Iod (*I*) und Gadolinium (*Gd*) aufgrund ihrer hohen Kernladungszahl und damit starken Abschwächung von Röntgenstrahlung als Kontrastmittel in der klinischen CT verwendet. Eine Beispielanwendung ist die Bildgebung von Blutgefäßen, die durch Injektion eines Kontrastmittels im Rahmen einer CT erfolgen kann. Die charakteristische K-Kante in den Absorptionsspektren von Iod bei $E_{k,I} = 33.17 \text{ keV}$ und Gadolinium bei $E_{k,Gd} = 50.24 \text{ keV}$ kann ebenfalls zur Bildgebung verwendet werden [81].

Technisch relevante, industrielle Materialien wie Aluminium (*Al*), Eisen (*Fe*), Kupfer

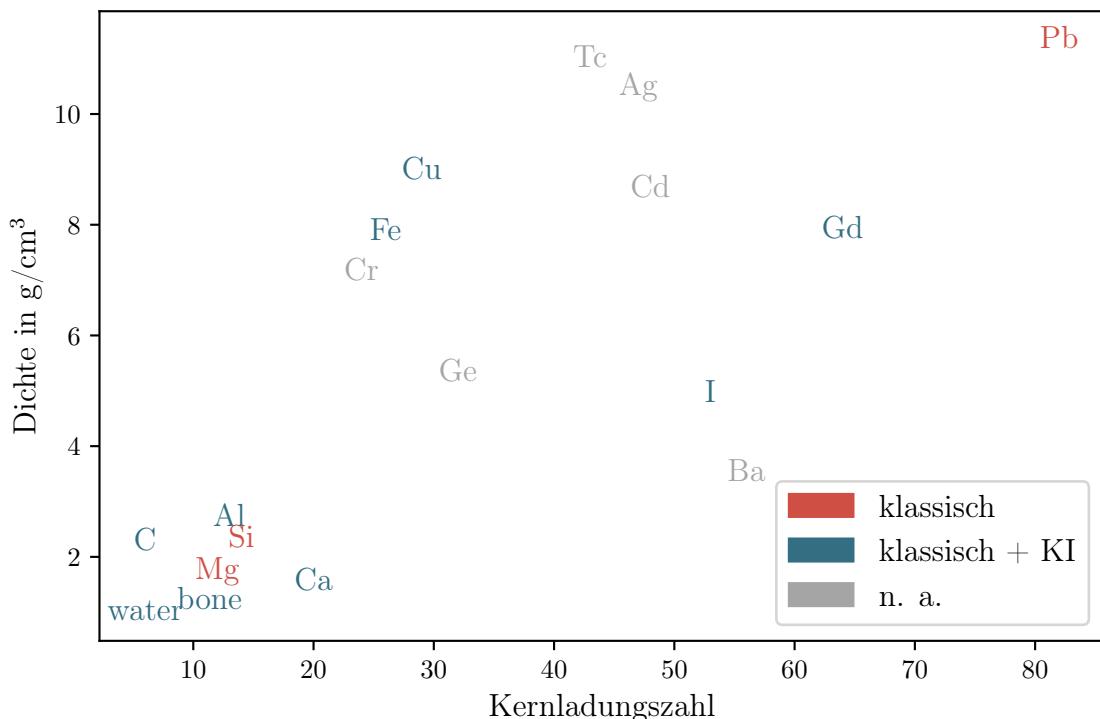


Abbildung 4.5: Graphische Darstellung der Materialien aus dem Korpus aufgelöst nach Kernladungszahl und Dichte. Im Periodensystem benachbarte Elemente wurden zur Ergänzung hinzugefügt und entsprechend gekennzeichnet (n.a.).

(Cu) und Blei (Pb) werden von Xing verwendet [71]. Aluminium kann aufgrund seiner niedrigen Abschwächung mit ebenfalls niedrigen Spannungen bis 140 kV in makroskopischen Proben identifiziert werden (vergleiche auch Heismann [14]). Zur Bildgebung der stärker abschwächenden Materialien verwendet Xing demnach Beschleunigungsspannungen bis 6 MV – die sogenannte LINAC-CT². Für Aluminium, Eisen und Kupfer konnte in einer dieser Thesis vorausgehenden Publikation gezeigt werden, dass die Z-ρ-Zerlegung von LINAC-CT-Daten mit DL möglich ist [76]. Drei Materialien wurden ausschließlich in Publikationen mit klassischen Ansätzen gefunden: Magnesium, Silicium und Blei [82, 71]. Aus technischer Sicht besteht kein Grund für den Ausschluss dieser oder anderer fehlender Materialien von datengetriebenen Ansätzen. Gase werden aufgrund ihrer grundsätzlich extrem niedrigen Abschwächung der Röntgenstrahlung in der ausgewählten Literatur nicht betrachtet.

4.2.3 Klassische Ansätze zur Materialbestimmung

Der Begriff *klassische Ansätze* bezeichnet in dieser Dissertationsschrift materialauflösende Algorithmen in CT-Bildern, die von Domänenexperten basierend auf physikalischen

2 Klassische Röntgenstrahlungsquellen funktionieren nach dem in Abbildung 2.5 dargestellten Prinzip: Freie Elektronen werden durch ein elektrisches Feld direkt zu einer Anode beschleunigt und wechselwirken mit ebendieser. Für extrem hohe Beschleunigungsspannungen ist dieses Prinzip nicht anwendbar. Freie Elektronen werden vorbeschleunigt und in Paketen durch Hohlraumresonatoren geleitet. Innerhalb der Hohlraumresonatoren können extrem hohe elektrische Feldstärken realisiert werden, sodass die Elektronen kinetische Energien im Bereich vieler Megaelektronenvolt erreichen können. Dieser Aufbau wird *Linearbeschleuniger* (englisch: linear accelerator; kurz: LINAC) genannt.

Erkenntnissen entwickelt werden. Konkret basieren diese Ansätze auf der Physik der Abschwächung von Röntgenstrahlung durch Materie – hauptsächlich durch den photoelektrischen Effekt und die Compton-Streuung (vergleiche Abbildung 2.4) wie Alvarez und Heismann zeigen [13, 14]. Erweiternde, klassische Methoden werden in diesem Abschnitt vorgestellt. Firsching zeigt im Jahr 2009, dass die Materialzerlegung nach Alvarez auf zählende Detektoren (PCD) ausgeweitet und mithilfe präziser Vermessung der Detektorantwort, des Quellspektrums und einer Monte-Carlo-basierten Simulation verbessert werden kann [64]. Zudem fasst er die Basismaterialzerlegung nach Alvarez anschaulich für mehrdeutige Materialkombinationen aus mathematischer Perspektive zusammen: Bei der Wahl der Basismaterialien wird idealerweise eine linear unabhängige Basis bereitgestellt [64]. Die lineare Unabhängigkeit ab dem dritten Basisvektor wird durch Unstetigkeiten der linearen Abschwächungskoeffizienten – den K-Kanten – erzeugt [64], während die ersten beiden Basisvektoren trivial unabhängig sind, weil sie auf dem Photoeffekt und der Compton-Streuung fußen. Da die K-Kanten der meisten klinisch relevanten Materialien sehr niedrig liegen und da am Detektor keine Photonen aus diesen Energiebereichen nachweisbar sind, können schwere Elemente mit höheren K-Kanten der Basis einen weiteren linear unabhängigen Basisvektor hinzufügen [64].

Roessl verwendet eine linear unabhängige Basis, bestehend aus dem Abschwächungsanteil des photoelektrischen Effektes und Gadolinium, um lokale Gadoliniumkonzentrationen zu messen [61]. Durch die Messung der Gadoliniumkonzentrationen können verstopfte Blutgefäße, zum Beispiel durch Kalkablagerungen, gefunden werden [61].

In der industriellen Anwendung nutzt Jumanazarov die Basismaterialzerlegung für leichte Elemente mit Kernladungszahlen $6 \leq Z \leq 15$.

Aufbauend auf der von Heismann publizierten $Z\text{-}\rho$ -Zerlegung [14] und der stöchiometrischen Kalibration von CT-Daten nach Schneider [83], entwickelt Son eine verbesserte Methode zur Materialbestimmung für verschiedene Gewebetypen basierend auf PCD-Daten [62]. Xing erweitert Heismanns Ansatz für die LINAC-CT durch die Berücksichtigung der Paarbildung bei hohen Photonenergien [71].

Zusammenfassend lässt sich festhalten, dass klassische Algorithmen stark von der Modellierung der zugrundeliegenden Physik abhängig sind. Bei der Basismaterialzerlegung ist die Wahl einer anwendungsspezifischen Basis entscheidend. Für ein System, bestehend aus drei Materialien, werden im Bestfall drei linear unabhängige Basisvektoren gewählt, sodass die Zerlegung eindeutig ist. Häufig ist diese Auswahl stark eingeschränkt oder nicht möglich.

4.2.4 Datengetriebene Ansätze

Im Kontext dieser Dissertationsschrift fungiert der folgende Abschnitt als Grundlage für die Ausrichtung der Forschung, die aufbauend auf der bestehenden Literatur erfolgt. Es werden die verwendeten Datensätze sowie die implementierten Architekturen erläutert.

Datensätze

Einige Autoren nutzen das XCAT-Phantom [84, 85] zur Erzeugung der Trainingsdaten. Hierbei handelt es sich um eine Simulationssoftware, die ein anatomisch detailliertes Abbild des menschlichen Körpers erstellt und in Form eines Voxelvolumens exportiert [85]. Die voxelbasierten Phantome können in gängige CT-Simulationsprogramme eingeladen und weiterverarbeitet werden.

Mit dieser Methode erzeugt Shi zwölf Phantome, von denen jeweils 140 Schichtbilder extrahiert werden [69]. Insgesamt nutzen drei Autoren das XCAT-Phantom bei der Simulation der Trainingsdaten für klinische Materialbestimmung [69, 70, 80]. Fang ersetzt die klinischen Materialien im XCAT-Phantom durch industriell genutzte Materialien (zum Beispiel Eisen, Magnesium und Teflon) und erzeugt so Trainingsdaten für einen industriellen Anwendungsfall [11].

Das *FORBILD*-Phantom [86] ähnelt konzeptionell dem XCAT-Phantom und wird von Cao zur Phantomerzeugung mit Gewebe, Wasser und Kontrastmittel verwendet [75]. Zu beachten ist, dass beide Programme (XCAT und FORBILD) nur zur Erzeugung eines Phantoms und nicht zur Simulation der CT dienen.

Eine alternative Strategie zur Sammlung der benötigten Trainingsdaten besteht in der Annotation bestehender, realer CT-Daten. Die Annotation ist ein zeitintensiver und teurer Prozess, der durch manuelle Annotationsfehler an Genauigkeit einbüßen kann. Der größte Vorteil dieser Methode besteht darin, dass die Trainingsdaten der realen Verteilung, den Messdaten aus einer realen CT-Anlage, entstammen und so keine Sim-To-Real-Gap auftritt. Vier Autoren verwenden manuell annotierte Trainingsdaten [72, 73, 67, 68]. Insbesondere bei der Annotation klinischer CT-Daten vereinfacht das Domänenwissen von Experten den Annotationsprozess und die damit verbundene Genauigkeit deutlich. Bussod verwendet die Synchrotron-CT³, um die Annotation der regulären, polychromatischen CT-Bilder vorzunehmen [66].

Sidky veröffentlicht einen Trainingsdatensatz [87], bestehend aus DECT-Bildern als Inputs und drei Materialkanälen als Ground Truths.⁴ Die simulierten DECT-Bilder zeigen unter anderem Verkalkungen, die gefunden werden sollen. Durch eine niedrige Winkelabdeckung der simulierten CT-Trajektorie sind die Bilder von starken Artefakten überlagert. Neben den Testdaten umfasst der Trainingsdatensatz effektiv 1000 DECT-Schichtbilder [87], sodass das Training parameterstarker Architekturen erschwert wird.

Krebbers nutzt energiedispersive Röntgenspektroskopie (auch bekannt als *EDX*) zur Annotation gegebener Graphitproben [77]. Grundsätzlich ist EDX eine Analysemethode, die besonders für Oberflächen geeignet ist, sodass die Methode von Krebbers nur auf sehr flachen Proben anwendbar ist.⁵

³ Zur Erzeugung der Röntgenstrahlung wird ein Synchrotron verwendet. Aufgrund des extrem hohen Photonenfluxes kann die polychromatische Röntgenstrahlung mithilfe eines Monochromators auf ein schmales Energieband gefiltert werden, sodass nahezu monoenergetische Photonen zur CT verwendet werden können. In diesem Fall kann die in Abschnitt 2.6, insbesondere Abbildung 2.14, beschriebene Methode trivial zur Materialbestimmung angewandt werden.

⁴ Methodisch ist für diese Problemstellung die Basismaterialzerlegung zu wählen.

⁵ Bei dem Beschuss einer Probe mit einem Elektronenstrahl werden aufgrund der Eindringtiefe der Elektronen vorzugsweise oberflächennahe, aber auch tiefer liegende Atome angeregt. Bei der Relaxation geben die Atome Photonen einer charakteristischen Energie ab, sodass anhand dieser Photonen ein Element identifiziert werden kann. Emittierte Photonen aus tieferen Lagen der Probe verlassen jedoch nur selten die Probe, da

In einer dieser Arbeit vorausgehenden Publikation wird ein Verfahren zur Erzeugung zufälliger Phantome, bestehend aus einer Vielzahl einzelner Regionen mit variablen Formen, veröffentlicht [76]. Durch die hohe Varianz der so erzeugten Phantome, wird die Generalisierungsfähigkeit eines trainierten Modells gefördert, da keine festen Muster existieren, die auswendig gelernt werden können.

Architekturen

Abgesehen von leichten Modifikationen, zum Beispiel der notwendigen Anpassung der Featurezahl der Input und Output Layer zur Verarbeitung von DECT-Bildern, nutzen viele Autoren das Vanilla U-Net⁶ nach Ronneberger [36]. Ein Überblick über die verwendeten Architekturen und die Trainingsdatensätze ist in Tabelle 4.1 dargestellt. Cao nutzt das Vanilla U-Net zur Vervollständigung fehlender Projektionsdaten in einem sogenannten *Sparse-Angle-CT* [75] – einem CT-Scan mit niedriger Winkelabdeckung, das heißt mit zu wenigen Projektionen. Darüber hinaus werden drei weitere Vanilla U-Nets verwendet, um die Trainingsdaten mit verschiedenen Kontrastmittelkonzentrationen zu verarbeiten [75]. Li nutzt ein Vanilla U-Net zur Basismaterialzerlegung von Wasser und Iod anhand von CT-Schichtbildern [73]. Die gleiche Zerlegung wird von Shi anhand von Projektionen mit einem Vanilla U-Net durchgeführt [70]. Beide Autoren nutzen mehrere MLPs, um die Predictions der U-Nets im Kontext der Inputs auf physikalische Konsistenz zu überprüfen [73, 70].

Fang, Abascal und Bussod nutzen flachere Versionen des U-Nets mit einer reduzierten Tiefe (Anzahl der Pooling beziehungsweise Upconv Layer) und Breite (Anzahl der Features nach Convolutions) [11, 65, 66]. Nadkarni verwendet ebenfalls eine flachere Version des U-Nets, um dreidimensionale Volumen verarbeiten zu können [78]. Aufgrund der zusätzlichen räumlichen Dimension steigt die benötigte Rechenzeit signifikant an.

Neben diesen Vereinfachungen des ursprünglichen U-Nets versuchen einige Autoren, Verbesserungen an der Architektur zu implementieren. Shi verwendet variable Kernelgrößen (3×3 , 5×5 , 7×7) im Encoder zusammen mit einer sogenannten *Local and Non-Local Feature Aggregation*, die zur verbesserten Erfassung lokaler Bildeigenschaften über die verschiedenen Größenskalen im Encoder beitragen soll [69]. Es wird quantitativ gezeigt, dass die entwickelte Architektur im Vergleich zum Vanilla U-Net auf simulierten XCAT-Phantom-Daten die Materialzerlegung verbessert [69]. In einer dieser Thesis vorausgehenden Publikation wird gezeigt, dass der Encoder des U-Nets durch einen Swin Transformer [38] ersetzt werden kann [76].

Die bisher vorgestellten Ansätze führen die Berechnungen in einer festen Datendomäne, Projektionen oder Volumen, durch. Su stellt eine Architektur namens *DIRECT-NET* vor, die Projektionen direkt in rekonstruierte Volumen in Form einer Basismaterialzerlegung umrechnet [74]. Der vorgestellte Ansatz ist zweistufig [74]. Zunächst werden aus DECT-Sinogrammen acht neue Sinogramme mit einem leicht modifizierten U-Net berechnet. Diese acht Sinogramme werden mit einem CT-Rekonstruktionsoperator in

sie aufgrund ihrer niedrigen Energie sofort wieder innerhalb der Probe wechselwirken. Somit stammt der Großteil des Signals, das beim EDX gemessen wird, aus oberflächennahen Schichten.

⁶ Der Begriff *Vanilla* bezieht sich im Kontext von Deep Learning auf die Originalversion einer Architektur ohne Modifikationen.

acht Schichtbilder umgerechnet. Ein weiteres U-Net fasst die acht Schichtbilder zu zwei Basismaterialbildern zusammen. Das Alleinstellungsmerkmal dieses Ansatzes ist, dass der Rekonstruktionsoperator Teil der Architektur ist [74]. Er muss differenzierbar sein, um das sogenannte *End-to-End* Training der gesamten Architektur zu ermöglichen. Mit dieser Methode können CT-Artefakte, die während der Rekonstruktion entstehen, effektiv unterdrückt werden. Nach Maier wird die Integration bekannter Operatoren, insbesondere der CT-Rekonstruktion, in größere Architekturen als *Precision Learning* bezeichnet [88]. Gong verwendet ein CNN mit sogenannten *Inception Blocks*, um weiches Gewebe, Knochen und Iod zu erkennen [79]. Das trainierte Modell wird mit dem Vanilla U-Net verglichen und liefert bessere Ergebnisse [79].

Drei Autoren nutzen sogenannte *Generative Adversarial Networks (GANs)*, die ursprünglich für das Lernen von Verteilungen vorgestellt worden sind [89]. GANs werden aus einem *Generator* und einem *Discriminator* aufgebaut. Der Generator erzeugt Outputs, die sogenannten *Candidates*, die zur Verteilung der Trainingsdaten passen sollen. Anhand des Outputs entscheidet der Discriminator, ob der gegebene Candidate *real* oder *artificial* ist, also ob er direkt aus dem Trainingsdatensatz stammt oder ob er vom Generator künstlich innerhalb der gelernten Verteilung erzeugt worden ist [89]. Generator und Discriminator agieren als Gegenspieler (*adversarial*) und werden gemeinsam trainiert, sodass der Generator zunehmend bessere, das heißt der Verteilung der Trainingsdaten ähnlichere, künstliche Outputs erschafft, während der Discriminator zunehmend besser zwischen den realen Trainingsdaten und den künstlichen Outputs des Generators unterscheiden kann. Das Funktionsprinzip von GANs wird auch für Transfer Learning verwendet [90]. Wang verwendet ein GAN mit einem modifizierten U-Net als Generator [72]. Nach dem Encoder im U-Net wird ein Transformer-Modul, bestehend aus einer Self-Attention-Layer und einem MLP, auf den Latent Space angewandt [72]. Zwei Discriminatoren entscheiden anhand der generierten Wasser- und Iod-Bilder (Basismaterialzerlegung), ob die Bilder *real* oder *artificial* sind. Der Loss der gesamten Architektur wird durch

$$\mathbb{L} = \mathbb{L}_1 + \mathbb{L}_{\text{adversarial}} + \mathbb{L}_{\text{VGG}}$$

definiert, wobei \mathbb{L}_1 der Generator-Loss, $\mathbb{L}_{\text{adversarial}}$ der Discriminator-Loss und \mathbb{L}_{VGG} der sogenannte *Perceptual Loss* nach Johnson [91] ist. Guo verwendet ein Vanilla U-Net als Generator und ein CNN als Discriminator, der Inputs (DECT-Bilder) und Predictions (Materialkarten) als Entscheidungsgrundlage verwendet, sodass die Predictions zusätzlich im Kontext der Inputs auf Konsistenz geprüft werden [68]. Geng nutzt eine ähnliche Architektur zur Segmentierung einer metallischen Nadel in Röntgenprojektionsbildern [80].

Im vorherigen Abschnitt zu den verwendeten Datensätzen wurde ein von Sidky publizierter Datensatz [87] genannt, welcher im Rahmen der AAPM *deep-learning spectral CT Grand Challenge* – einem Wettbewerb zur Materialzerlegung – erstellt worden ist. Sidky fasst die besten zehn Modelle dieser Challenge in seiner Publikation zusammen [87]. Im Rahmen dieser Literaturrecherche werden die Platzierungen 1, 3 und 4 erfasst. Der zweitplatzierte Ansatz basiert auf einem klassischen Verfahren, das eine akkurate

CT-Simulation voraussetzt und iterativ eine optimale Lösung erprobt.⁷ Die Gewinner der Challenge, *Team GM_CNU*, verwenden einen zweistufigen Ansatz. Mithilfe einer iterativen Rekonstruktion⁸ werden Volumen rekonstruiert, die anschließend durch ein CNN (genauer: RED-CNN [92]; eine Variation des U-Nets) in Basismaterialbilder zerlegt werden [87]. Diese Basismaterialbilder werden mit einer CT-Simulation rückprojiziert und ergeben zusammen mit den ursprünglichen Inputs sogenannte *Residual Images* – Differenzbilder, die zusammen mit den Basismaterialbildern in ein weiteres CNN gegeben werden, um die Basismaterialbilder weiter zu verbessern. Insgesamt ist dieser Ansatz sehr vergleichbar mit der bereits vorgestellten Methoden nach Liu [73].

Den dritten Platz belegt *Team MIR* mit vier U-Nets, die jeweils Teilaufgaben der Zerlegung erfüllen [87]. Das erste U-Net berechnet die Summe der Basismaterialien aus den DECT-Bildern. Das zweite U-Net erstellt eine binäre Segmentierungsmaske zur Detektion der Verkalkungen. Das dritte U-Net erstellt die gesuchten Calcium- und Fettmaterialkarten. Das vierte U-Net verfeinert die Materialkarten.

Den vierten Platz belegt *Team WashUDEAM* mit einer Basismaterialzerlegung in alle möglichen Kombinationen der drei Basismaterialien paarweise, das heißt *AB*, *AC*, *BC* sowie die drei zugehörigen Permutationen, in insgesamt sechs verschiedene Materialbasen [87]. Ein U-Net verrechnet die sechs Materialbasen in die gesuchten drei Materialkarten.

Zusammenfassend nutzt die absolute Mehrheit der Autoren das U-Net mit teilweise kleinen Anpassungen sowohl im klinischen als auch im industriellen Umfeld (vergleiche Tabelle 4.1). Es konnten einige Hybrid-Architekturen identifiziert werden, die dennoch in ihrer Kernidee ein U-Net benutzen und durch weitere Mechanismen, beispielsweise die Konsistenzprüfung zwischen Inputs und Outputs mit einem Discriminator, ergänzen. Weiterhin spielen Transformer-basierte Architekturen derzeit eine zu vernachlässigende Rolle, obwohl ihr Beitrag zum Fortschritt anderer Bereiche der Computer Vision unbestreitbar ist. Diese Beobachtung konnte durch die Aktualisierung der Literaturrecherche im Februar 2025 erneut bestätigt werden.

Beobachtungen zur Rechenzeit und Hardware

DL-basierte Modelle benötigen einen gewissen Rechenzeitaufwand, der im Rahmen des Trainings aufgebracht werden muss. Abhängig von der Komplexität der Architektur und des zu lernenden Problems, variieren die Anforderungen an die Hardware. Die meisten Autoren aus dem Korpus machen keine Angaben zur verwendeten Hardware. Es ist anzunehmen, dass die meisten Architekturen, die auf U-Nets basieren, mit einer hinreichend effizienten Implementierung auf einzelnen, modernen Graphikkarten trainiert werden können. Somit sind derartige Architekturen für die allermeisten Entwickler verfügbar und ebenfalls effizient in praktischen Anwendungen nutzbar. Abascal führt das Training auf einer Nvidia GTX 1080Ti [65] durch und Shi nutzt eine Nvidia RTX 2080 [70] – beide Graphikkarten entstammen dem Gaming-Portfolio von Nvidia, benötigen

⁷ In grober Auslegung kann das Verfahren zu Machine Learning gezählt werden. Große Teile der Methode basieren auf physikalischer Modellierung und unterscheiden sich in ihrer Komplexität deutlich von den anderen Verfahren.

⁸ Durch die Verwendung einer iterativen Rekonstruktion werden die CT-Artefakte, teils bedingt durch die Sparse-Angle CT des gegebenen Datensatzes, reduziert.

keine speziellen Systemanforderungen und sind für Konsumenten frei erhältlich. In einer dieser Thesis vorausgehenden Publikation wird ein Transformer-basiertes U-Net mithilfe eines Nvidia DGX-A100-Systems trainiert [76].

Derzeit ist unklar, ob die Mehrheit der Autoren das U-Net verwendet, weil es die besten Ergebnisse liefert oder weil parameterstärkere Architekturen die verfügbaren Rechenressourcen überlasten würden. Hinzu kommt, dass die Datenbeschaffung, insbesondere realer CT-Daten, ein zeitaufwändiger und daher teurer Prozess ist, sodass dateneffiziente Architekturen weiterhin im Vorteil sind. Die zunehmende Leistungsfähigkeit, und die damit verbundene breite Verfügbarkeit der Hardware wird in den kommenden Jahren zeigen, ob sich das U-Net gegen neuartige Architekturen, zum Beispiel gegen den Transformer, langfristig durchsetzen kann.

Tabelle 4.1: Überblick über die Architekturen und Datensätze der datengetriebenen Ansätze aus dem Korpus. Die Publikationen von Nadkarni [78] und Krebbers [77] machen keine Angaben zur Größe des Datensatzes.

Autor	Architektur	Datensatzart	Datensatzgröße	Datendomäne
Long 2019 [67]	FC-PRNet	Scan	≈200	Volumenschnitt
Shi 2019 [70]	U-Net	Simulation	140	Projektion
Bussod 2021 [66]	U-Net	Scan (Synchrotron)	450 K	Projektion
Gong 2020 [79]	U-Net + InceptNet	Scan	110 K	Volumenschnitt
Geng 2021 [80]	PMS-GAN	Simulation + Scan	124 + 124	Projektion
Abascal 2021 [65]	U-Net	Simulation	5400	Volumenschnitt + Projektion
Su 2022 [74]	U-Net	Simulation	10 K	Volumenschnitt + Projektion
Fang 2022 [11]	U-Net	Simulation	300	Volumenschnitt
Nadkarni 2022 [78]	U-Net	Scan	-	Volumenschnitt
Wang 2022 [72]	GAN	Scan	8159	Volumenschnitt
Li 2023 [73]	U-Net + MLP	Scan	7218	Volumenschnitt
Cao 2022 [75]	CNN	Simulation	≈12 K	Volumenschnitt
Guo 2023 [68]	GAN + U-Net	Scan	1 K	Volumenschnitt
Shi 2024 [69]	U-Net	Simulation	35 K	Volumenschnitt
Krebbers 2023 [77]	sensor3D	Scan + XRD	-	Volumenschnitt
Weiss 2024 [76]	U-Net	Simulation	64 K	Volumenschnitt

4.3 ERKENNTNISGEWINN AUS DER LITERATUR

Die Ergebnisse der Literaturrecherche werden im Folgenden zusammengefasst. Als Basis für die aufbauende Forschung in dieser Dissertationsschrift wird auf die Erzeugung, oder Sammlung, der Trainingsdaten sowie die DL-Architekturen ein besonderer Fokus gelegt.

4.3.1 *Erzeugung der Trainingsdaten*

Datengetriebene Ansätze benötigen einen Datensatz, um die zugrundeliegenden Zusammenhänge selbstständig aus den Daten zu erlernen. Die Herkunft dieser Datensätze kann entweder eine Simulation oder eine Serie realer CT-Scans sein. Es wurden vier Publikationen gefunden, die das XCAT-Phantom [84, 85] verwenden, um (digitale) Phantome für eine anschließende CT-Simulation zu erzeugen [69, 70, 80, 11]. Fang ersetzt die Materialien des XCAT-Phantoms, um das Phantom für einen industriellen Anwendungsfall mit den Materialien Eisen, Magnesium und Teflon zu nutzen [11]. Einige Autoren nutzen reale Scandaten und annotieren die CT-Bilder händisch. Diese Methode bietet den Vorteil, dass die Verteilung der Trainingsdaten sehr genau ebenjener Verteilung entspricht, für die das trainierte Modell später ausgeführt werden soll. Teilweise werden andere Messmethoden, beispielsweise Synchrotron-CT [66], verwendet, um die Annotation vorzunehmen. Aus dem klinischen Umfeld stellt Sidky einen offenen, simulierten Datensatz zur Materialzerlegung und gezielten Suche von Calcium in Brust-DECT-Bildern bereit. Grundsätzlich sind weder offene Datensätze zur Materialbestimmung noch performante Simulationsprogramme aus dem industriellen Umfeld im Korpus gefunden worden, die zur Erzeugung von Trainingsdaten lückenlos geeignet sind. Einer der Forschungsziele dieser Arbeit ist daher, eine schnelle und modulare Methode zur Trainingsdatenerzeugung für die Materialbestimmung anhand von industriellen CT-Schichtbildern zu finden, oder bei Bedarf zu entwickeln, sowie die verwendeten Trainingsdaten zu veröffentlichen.

4.3.2 *Architekturen*

Eine Zusammenfassung der verwendeten Architekturen ist in Tabelle 4.1 angegeben. Die absolute Mehrheit der Autoren verwendet entweder das Vanilla U-Net oder eine stark an das U-Net angelehnte Architektur mit kleinen, teils äußerst effektiven Anpassungen [69, 70, 66, 65, 74, 11, 78]. In einer dieser Dissertationsschrift vorausgehenden Publikation wurde gezeigt, dass, bei der Materialbestimmung in LINAC-CT-Bildern, der Encoder des U-Nets durch einen Swin Transformer [38] ersetzt werden kann. Einige Autoren verwenden GANs, um mithilfe des Discriminators die Konsistenz zwischen den Inputs (DECT-Bilder) und den Predictions (Materialvorhersage) zu überprüfen [80, 68, 72]. Die in Abschnitt 4.1 geäußerte Vermutung, datengetriebene Ansätze basierten zunehmend weniger auf der Verwendung des Domänenwissens klassischer Materialzerlegung, kann somit nicht bestätigt werden. Vielmehr machen die Autoren des Korpus deutlich, dass sie Domänenexperten der Computertomographie sind und dass sie Deep Learning als praktisches Werkzeug zur weiteren Verbesserung ihrer Methodik nutzen.

Aufgrund mangelnder Angaben in den Publikationen des Korpus kann nicht abschließend ermittelt werden, ob die Verwendung des U-Nets erfolgt, da es die besten Ergebnisse liefert oder weil parameterstärkere Architekturen mit den verfügbaren Ressourcen nur eingeschränkt verfügbar sind. Somit besteht ein Forschungsbeitrag dieser Arbeit in der Implementierung und dem quantitativen Vergleich Convolution- und Transformer-basierter Architekturen für die Materialbestimmung in CT-Bildern.

5

SIMULATION DER COMPUTERTOMOGRAPHIE

Ziel dieses Kapitels ist die Definition und Umsetzung einer simulationsbasierten Methode, um Trainingsdaten für die DL-basierte Materialbestimmung zu erzeugen. Die Vorteile einer Simulation bei der Erzeugung großer Datenmengen gegenüber einer realen Messserie, bezogen auf die DL-basierte Materialbestimmung, wurden bereits in Kapitel 1.1 angeführt und dienten zur Formulierung der ersten Forschungsfrage, die im Rahmen dieses Kapitels beantwortet werden soll. Zunächst erfolgt eine Anforderungsanalyse in Abschnitt 5.1, die sowohl die CT-spezifischen Aspekte der benötigten Simulation als auch die durch die Wahl eines DL-basierten Ansatzes induzierten Randbedingungen beleuchtet. Einige existierende Simulationsprogramme werden in Abschnitt 5.2 knapp vorgestellt und im Kontext der Anforderungsanalyse eingeordnet. Aufgrund einer unzureichenden Abdeckung der Anforderungen aus Abschnitt 5.1 durch die existierenden Simulationen aus Abschnitt 5.2 muss eine neue Simulation entwickelt werden, die in Abschnitt 5.3 beschrieben wird. Abschließend erfolgt die Evaluation der spezifizierten Anforderungen aus Abschnitt 5.1 an die implementierte Simulation in Abschnitt 5.4.

5.1 ANFORDERUNGSANALYSE: CT-SIMULATION

Die funktionalen Anforderungen an eine CT-Simulation, die zur Erzeugung der Trainingsdaten für die DL-basierte Materialbestimmung eingesetzt werden kann, lassen sich in zwei Kategorien unterteilen: die CT-spezifischen Anforderungen und die DL-spezifischen Anforderungen. Für den erfolgreichen Transfer eines trainierten DL-Modells auf eine konkrete Anwendung an einer realen CT-Anlage muss die Verteilung der Trainingsdaten möglichst genau mit der Verteilung der Messdaten der realen CT-Anlage übereinstimmen. Somit müssen die relevanten Einflussfaktoren, die die Messdaten an der realen CT-Anlage charakterisieren, in der Simulation physikalisch korrekt abgebildet werden. Die Identifikation dieser CT-spezifischen Einflussfaktoren erfolgt in Abschnitt 5.1.1.

Wie in Kapitel 3 beschrieben, ist Deep Learning ein flexibles und damit an viele Fragestellungen anpassbares Werkzeug, sofern (a) die Qualität der Trainingsdaten hoch ist und (b) diese in ausreichender Menge zur Verfügung stehen. Die Qualität der Trainingsdaten muss in zweierlei Hinsicht, bezogen auf die Fragestellung dieser Arbeit, gewährleistet sein: Erstens muss die Simulation den physikalischen Messprozess der CT gut modellieren. Dies wurde im Rahmen der CT-spezifischen Anforderungen bereits erfasst. Zweitens muss der Trainingsdatensatz eine große Vielfalt unterschiedlicher CT-Bilder umfassen, sodass das Modell im Training möglichst viele unterschiedliche Formen der virtuellen Prüfteile (auch: *Phantome*) und Materialien erlernen kann (Abschnitt 5.1.2). Die absolut benötigte Menge der Trainingsdaten (b) hängt stark von der Komplexität, oder auch Ka-

pazität, der gewählten Architektur ab, welche sich wiederum aus der Komplexität des zu lernenden Problems ableitet. Im Rahmen dieser Anforderungsanalyse soll eine sinnvolle Größenordnung für die Menge der benötigten Trainingsdaten begründet abgeschätzt werden (vergleiche DL-spezifische Anforderung in Abschnitt 5.1.2).

Abschließend werden einige nicht-funktionale Anforderungen in Abschnitt 5.1.3 diskutiert.

5.1.1 CT-spezifische Anforderungen

Wie in Kapitel 2 beschrieben, fungiert die CT als ortsaufgelöstes Messinstrument für lineare Abschwächungskoeffizienten μ_L der Röntgenstrahlung. Demnach wird zunächst ein Simulationsmodul zur Berechnung des erzeugten Röntgenstrahlungsspektrums benötigt. Die in Abschnitt 2.2 eingeführten Röntgenquellen, beziehungsweise die von ihnen emittierten Spektren, werden typischerweise durch eine Beschleunigungsspannung sowie das Material und durch die Maße einer Anode parametrisiert. Häufig werden (Vor-)Filter (vergleiche Kapitel 2.4.1; Abbildung 2.11), auch in Form von Strahlfenstern (vergleiche Abbildung 2.5), ebenfalls bei der Simulation der Quelle berücksichtigt. Mithilfe dieser Angaben kann das emittierte Spektrum, das heißt der Photonenzahl pro Energie- und Winkelbereich, berechnet werden.

Im nächsten Schritt wird der durch die Probe transmittierte Anteil des Quellspektrums ermittelt. Hierzu werden die Schnittlängen x durch jedes in der Probe vorliegende Material ermittelt und durch Anwendung des LAMBERT-BEER-Gesetzes aus Gleichung 1.1 mit dem gegebenen Quellspektrum in Intensitäten umgerechnet. Der lineare Abschwächungskoeffizient $\mu_{L,m}(E)$ für das gegebene Material m kann der Literatur entnommen werden [93].

Die durch die Probe transmittierte Röntgenstrahlung trifft auf den Detektor und wird, analog zur Probe, erneut abgeschwächt. Durch erneute Anwendung des LAMBERT-BEER-Gesetzes aus Gleichung 1.1 folgt die Anzahl der Photonen, die im Szintillator des Detektors eine Wechselwirkung verursachen und damit ein Signal erzeugen (vergleiche Abschnitt 2.3). Wie in Kapitel 2.3 beschrieben, führt die zeitliche Integration über das Spektrum der wechselwirkenden Photonen im Szintillator zu einer deponierten Energie, die im Idealfall proportional zum Detektorsignal, den Counts eines EIDs, ist. Momentan stehen PCDs nicht im Fokus der Anwendung, sodass sie nicht berücksichtigt werden. Der Messprozess ist somit in dieser abstrahierten Darstellung in drei Schritte aufteilbar: die Simulation des Quellspektrums, die Abschwächung durch die Probe sowie die Abschwächung und Integration (Signalentstehung) im Detektor.

Anhand der Definition des linearen Abschwächungskoeffizienten in Gleichung 2.3 kann gefolgert werden, dass die Dichte für ein bestimmtes Material proportional zum linearen Abschwächungskoeffizienten ist. Somit kann der Toleranzbereich des linearen Abschwächungskoeffizienten direkt aus der geforderten Präzision für die Dichtevorhersage bestimmt werden. Aus Gesprächen mit CT-Anwendern ergibt sich der Toleranzbereich $\pm 1\%$ für die Berechnung der Dichte, der proportional zu $\frac{1}{\mu_m}$ auch für den linearen Abschwächungskoeffizienten angenommen werden kann (vergleiche Fehlerfortpflanzung

erster Ordnung von Gleichung 2.3). Eine weitere Anforderung besteht in der Berücksichtigung poröser Materialien, die für viele (chemische) Anwendungen von hoher Relevanz sind. Durch die Porosität wird die Kontaktoberfläche vergrößert, sodass gewisse chemische Prozesse effizienter ablaufen. Daher ist nicht nur die Bestimmung des eigentlichen Materials, sondern ebenso die Bestimmung der absoluten Dichte (in g/cm³) für poröse Materialien wichtig. Die simulierten Proben müssen daher durch variable, relative Dichten η , beispielsweise $\rho(\text{Al}^{80\%}) = 0.8 \cdot 2.7 \text{ g/cm}^3$, charakterisiert werden können.

5.1.2 Deep-Learning-spezifische Anforderungen

Die erste Anforderung besteht in der Wahl eines Datenformats, das gleichermaßen für die Problemstellung der Materialbestimmung sowie für die Verarbeitung durch DL geeignet ist. Wie in Kapitel 2.4 beschrieben, liegen die Ausgabedaten eines CT-Scans in Form von Volumen, beziehungsweise einer Serie von zweidimensionalen Schnittbildern durch das rekonstruierte Volumen, vor. Aufgrund ihrer Entwicklungsgeschichte sind die meisten DL-Architekturen aus dem CV-Bereich für die Verarbeitung zweidimensionaler Bilder optimiert. Ebenso ergibt die Literaturrecherche in Kapitel 4, dass lediglich ein Autor dreidimensionale Volumen als direkte Eingabe für ein neuronales Netz verwendet. Der Hauptgrund hierfür liegt vermutlich in dem extremen Anstieg der benötigten Rechenzeit beim Training der Modelle auf dreidimensionalen Daten. Somit wird im Folgenden der zweidimensionale Grenzfall der CT, die sogenannte *Fächerstrahl-CT* (vergleiche Abschnitt 2.4.1; *Fächerstrahl-Geometrie*), modelliert. Sie entspricht der Mittelschicht eines rekonstruierten Volumens. Die Algorithmik wird dem Ablauf eines CT-Scans nachempfunden: Aus den zweidimensionalen Phantomen werden Sätze eindimensionaler Projektionen für jede Energie des DECT berechnet, die wiederum in Volumen mit zwei Kanälen (Low-Energy und High-Energy) und zwei räumlichen Dimensionen (DECT-Schichtbilder), rekonstruiert werden. Die Phantome werden als Materialkarten, räumlich aufgelöst nach Kernladungszahl und Dichte, ausgedrückt, sodass sie als Ground Truths der durchgeführten Materialbestimmung anhand der DECT-Schichtbilder (Inputs) eingesetzt werden können. Sowohl die DECT-Schichtbilder als auch die Materialkarten haben demnach zwei Kanäle C und zwei räumliche Dimensionen (durch die Verwendung der Fächerstrahl-CT) H, W . Das Ergebnis eines Durchlaufs dieser Datenerzeugung ist ein sogenanntes *Tupel* bestehend aus den DECT-Schichtbildern (Inputs) und den Materialkarten (Ground Truths). Die Zusammensetzung der Trainingsdaten sollte idealerweise keinen sogenannten *Bias*¹ während des Trainings induzieren. Beispielsweise besteht die Möglichkeit, dass ein Modell, das exklusiv auf runden Phantomen trainiert worden ist, bei der Verarbeitung eckiger Phantome falsche Vorhersagen produziert. Gleches kann für die Materialzusammensetzung der Phantome gelten: Ein Modell, das nur auf leichten Elementen wie Aluminium oder Magnesium trainiert worden ist, wird Kupfer höchstwahrscheinlich nicht korrekt bestimmen können. Somit muss die Zusammensetzung der Trainingsdaten anwendungsspezifisch steuerbar sein. Da in dieser Arbeit vorrangig moderne,

¹ Gemeint ist nicht der Bias b , der bei der Einführung neuronaler Netze in Kapitel 3 genannt wurde, sondern eine grundsätzliche Voreingenommenheit des Modells aufgrund der Zusammensetzung der Trainingsdaten.

parameterstarke Architekturen untersucht werden sollen, besteht das Risiko, dass die gewählte Größe des Trainingsdatensatzes zu klein ausfällt, insbesondere wenn diese Wahl auf Basis der Literaturrecherche erfolgt in der die absolute Mehrheit der Autoren verhältnismäßig dateneffiziente Architekturen verwendet. Moderne Architekturen für CV-Anwendungen werden in der Regel auf Datensätzen mit mehreren Millionen Bildern vortrainiert [37, 38, 94], sodass die Möglichkeit bestehen muss, einen simulierten Trainingsdatensatz mit einigen Millionen Tupeln mindestens exemplarisch zu berechnen. Idealerweise ist die Implementierung so effizient, dass auch Rechner mit einzelnen Graphikkarten in weniger als 24 Stunden Datensätze der benötigten Größe erzeugen können. Konkret wird im Rahmen dieser Dissertationsschrift die Partition *WestAI* am Forschungszentrum Jülich verwendet, die einer Rechenzeitbeschränkung von maximal 24 Stunden pro Job unterliegt. Wie in der vorausgehenden Literaturrecherche in Kapitel 4 aufgezeigt wurde, schwanken die von den Autoren verwendeten Trainingsdatenmengen zwischen wenigen hundert Tupeln [70, 67] bis hin zu 450.000 Tupeln [66]. Es ist zu beachten, dass der auffällig große Datensatz mit 450.000 Tupeln auf Projektionen basiert, während in dieser Arbeit ein Verfahren auf Basis der CT-Schichtbilder entwickelt wird. Aus Tabelle 4.1 kann entnommen werden, dass Gong et al. den größten Datensatz im Korpus mit 110.000 Tupeln verwenden, der die Materialbestimmung basierend auf CT-Schnittbildern durchführt. Für das Training vergleichbarer Architekturen wird demnach abgeschätzt, dass eine Datensatzgröße von 250.000 (250K) CT-Schnittbildern ausreichend ist. Hieraus ergibt sich für einen Eintrag, bestehend aus einem DECT-Schichtbild mit zwei Kanälen und zwei Materialphantomen (Kernladungszahl und Dichte), die Gesamtgröße des Datensatzes

$$V = 250000 \cdot 2 \cdot 2 \cdot H \cdot W \cdot 4 \text{ Byte} = \begin{cases} \approx 66 \text{ GB} & \text{mit } H = W = 128 \\ \approx 262 \text{ GB} & \text{mit } H = W = 256 \\ \approx 1049 \text{ GB} & \text{mit } H = W = 512 \end{cases}$$

für die Bilddimensionen H und W sowie das Datenformat float32 mit 4 Bytes. Die Wahl der Bilddimensionen anhand von Zweierpotenzen bietet den Vorteil, dass die Speicherzugriffe bei den Rechenoperationen effizienter sind, da die Speicherblöcke ebenfalls anhand von Zweierpotenzen angeordnet sind. Zusätzlich werden die räumlichen Bilddimensionen H, W durch das in Kapitel 3.4 eingeführte Pooling bei der Verarbeitung mehrfach halbiert, sodass auch hier H und W möglichst oft durch zwei teilbar sein sollten. In Anbetracht des benötigten Speicherplatzes, insbesondere bei Datensätzen mit bis zu einer Million Tupeln, wird die Bildgröße für die Ausgabe der Simulation im Folgenden auf 256×256 festgelegt. Die Folgen dieser Wahl werden an den entsprechenden Stellen in den weiteren Kapiteln anwendungsbezogen diskutiert. Theoretisch können die Materialkarten der Kernladungszahlen im uint8-Format abgespeichert werden, um Speicherplatz auf Kosten der Verarbeitungsgeschwindigkeit zu sparen.²

² Durch das Speichern der Daten in einem anderen Datenformat müssen die Daten bei jedem Ladevorgang, das heißt in jeder Epoche des Trainings, konvertiert werden. Die Reduktion des benötigten Speicherplatzes durch das Speichern der Kernladungszahlen im uint8-Format beträgt etwa 19 %.

5.1.3 Nicht-funktionale Anforderungen

Neben den funktionalen Anforderungen aus den vorangehenden Abschnitten können einzelne, nicht-funktionale Anforderungen identifiziert werden. Zunächst muss die Simulation zur Vereinfachung der Reproduzierbarkeit auf quelloffener Software basieren. Sie soll zudem mithilfe einzelner, lesbbarer Konfigurationsdateien, bevorzugt im JSON-Format, parametrisierbar sein, um eine einfache Reproduzierbarkeit der Ergebnisse zu erreichen. Gekoppelt an die Anforderung bezüglich der Größe des zu simulierenden Datensatzes und der daraus folgenden Anforderung an den Berechnungsaufwand soll die Simulation Cluster-fähig sein, das heißt in erster Instanz mit dem Betriebssystem *Linux* kompatibel und einfach skalierbar sein.

5.1.4 Überblick über die Anforderungen

Angelehnt an die Durchführung eines realen CT-Scans, bestehend aus der Auswahl einer zu scannenden Probe, der Projektion der Probe mithilfe von Röntgenstrahlung sowie der abschließenden Rekonstruktion, kann der Ablauf der benötigten Simulation zur Erzeugung der Trainingsdaten im Rahmen dieser Arbeit abgeleitet werden. Abbildung 5.1 stellt diesen Ablauf dar, wobei die Auswahl der (virtuellen) Probe durch den sogenannten *Phantomgenerator* randomisiert erfolgt.

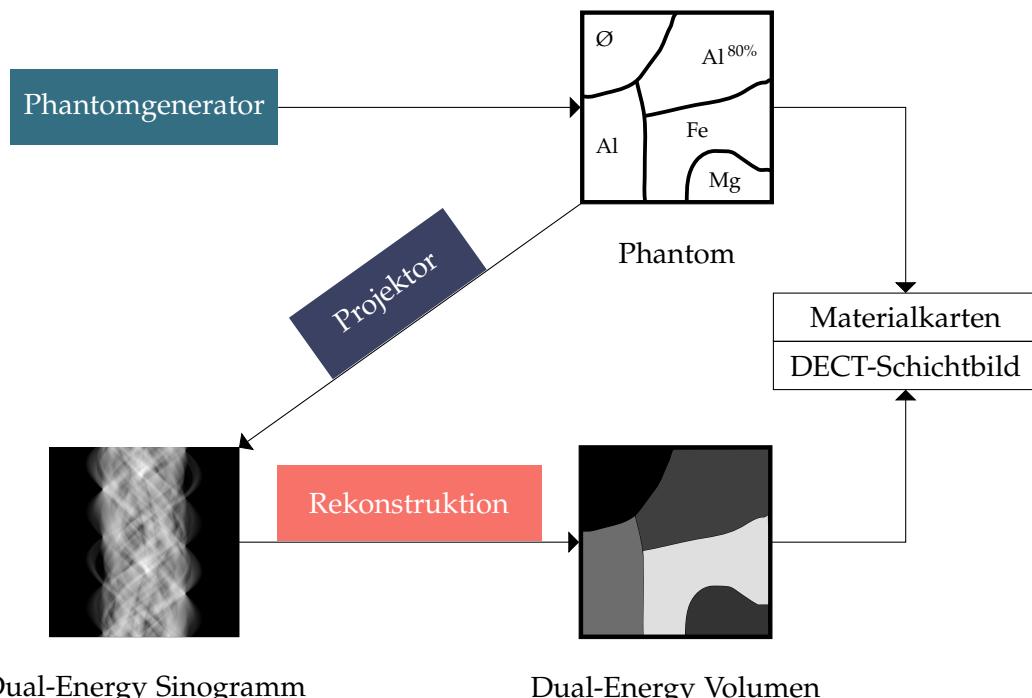


Abbildung 5.1: Ablaufdiagramm der vorgestellten Methode zur Erzeugung der Trainingsdaten. Der Phantomgenerator erzeugt zufällige Phantome, die durch den Projektor und die Rekonstruktion in DECT-Schichtbilder umgerechnet werden. Die Phantome werden in Form von Materialkarten, das heißt räumlich aufgelöst nach Kernladungszahl und Dichte, dargestellt. Im oben rechts gezeigten Phantom ist eine leere Region mit \emptyset beschriftet. Poröse Materialien werden durch den hochgestellten, relativen Volumenanteil angezeigt: $\text{Al}^{80\%}$ stellt poröses Aluminium mit einer mittleren Dichte von $\rho = 0.8 \cdot 2.7 \text{ g/cm}^3 = 2.16 \text{ g/cm}^3$ dar.

5.2 EXISTIERENDE CT-SIMULATIONEN

Anhand der soeben beschriebenen Anforderungen werden einige bestehende Simulationsprogramme hinsichtlich ihrer Eignung für die Trainingsdatenerzeugung im Rahmen dieser Arbeit bewertet. Grundsätzlich lassen sich CT-Simulationen in zwei Kategorien einteilen: Monte-Carlo-basierte Simulationen und Raytracing-basierte Simulationen [95]. Monte-Carlo-basierte CT-Simulationen modellieren den physikalischen Prozess auf Teilchenebene – die Wechselwirkung der einzelnen Photonen mit Materie. Sie erreichen somit die präziseste Modellierung einer realen CT-Anlage und sind den anderen Verfahren in diesem Aspekt weit überlegen. Anhand des Photonenflusses durch den Detektor einer realen CT-Anlage (typischerweise $\gg 1 \times 10^8 \text{ s}^{-1}$) kann beispielhaft überschlagen werden, dass Monte-Carlo-basierte CT-Simulationen für die Berechnung einer einzelnen Projektion viele Minuten bis hin zu mehreren Tagen benötigen [96, 97]. Die Monte-Carlo-basierte Simulation einer großen Sammlung unterschiedlicher CT-Scans, jeweils bestehend aus tausenden von Projektionen, ist praktisch nicht durchführbar und demnach zur Trainingsdatenerzeugung im Rahmen dieser Arbeit nicht geeignet.

Raytracing-basierte-Verfahren gehen nach einer anderen Methodik vor: Die emittierte Röntgenstrahlung wird in Form einzelner Strahlen zwischen der Röntgenquelle und jedem Pixel des Detektors modelliert. Schematisch, jedoch in einem anderen Kontext, ist diese Modellierung in Abbildung 2.1 bereits dargestellt worden. Jeder Strahl wird durch ein Photonenspektrum, das heißt durch Photonenzahlen abhängig von ihrer Energie $I(E)$, charakterisiert. Aus der Schnittlänge mit der Probe kann durch Anwendung des LAMBERT-BEER-Gesetzes aus Gleichung 1.1 das Photonenspektrum nach Durchgang durch die Probe ermittelt werden. Somit wird pro Strahl (pro Detektorpixel) einmal die Schnittlänge und einmal das LAMBERT-BEER-Gesetz berechnet, um das transmittierte Photonenspektrum zu erhalten. Die konkreten Verfahren zur Berechnung der Schnittlängen werden im Rahmen dieser Arbeit nicht weiter ausgeführt. Eine beispielhafte Methode kann bei Vidal et al. eingesehen werden [95].

Die Raytracing-basierten Verfahren, speziell für die industrielle CT-Modellierung, unterliegen einer erheblichen Einschränkung: die modellierten Strahlen sind *gerade*. Wie in Kapitel 2.1 beschrieben, steigt der Anteil der Compton-Streuung an der totalen Wechselwirkungswahrscheinlichkeit mit der Energie im CT-relevanten Energiebereich an (vergleiche Abbildung 2.4). Die räumliche (Zer-)Streuung des Photonenstrahls³ wird mit einem Raytracing-basierten Ansatz nicht korrekt abgebildet, obschon die absolute Abschwächung durch die Compton-Streuung mit der Auswertung des LAMBERT-BEER-Gesetzes korrekt erfasst wird. Im Folgenden werden ausgewählte, Raytracing-basierte Simulationsprogramme kurz vorgestellt und hinsichtlich ihrer Eignung für die Trainingsdatenerzeugung anhand der in Abschnitt 5.1 aufgestellten Anforderungen bewertet. Die Gründe für die Auswahl der vorgestellten Programme werden in den jeweiligen Abschnit-

³ Die Winkelverteilung des Photons nach dem Compton-Streuprozess folgt der Gesetzmäßigkeit nach Klein-Nishina [3, S. 50 f.] und ist daher für steigende Photonenergien zunehmend nach vorne, das heißt in die ursprüngliche Ausbreitungsrichtung des Photons, gerichtet. Bei steigender Photonenergie steigt somit auch das Detektorsignal, da es zu einer Summation mit den gestreuten Photonen kommt.

ten dargelegt. Ein grundsätzlicher Orientierungspunkt besteht stets durch Gespräche mit anderen CT-Entwicklern.

5.2.1 *aRTist*

Im Rahmen des öffentlich geförderten Forschungsprojektes *CTSimU*, das sich mit der Messunsicherheitsbestimmung realer CT-Anlagen anhand von Simulationen befasst, wird die Simulation *aRTist* häufig referenziert [98, 99]. Das Konsortium des Forschungsprojektes besteht sowohl aus Universitäten als auch aus Vertretern der Industrie, sodass der genannten Simulation *aRTist* ein gewisser Stellenwert zuzuschreiben ist. *aRTist* wird von der *Bundesanstalt für Materialforschung und -prüfung* entwickelt und liegt aktuell in der Version 2.12.6 vor [24]. Eine der großen Stärken von *aRTist* besteht in der umfassenden Sammlung an ergänzenden Paketen, die zum Beispiel die Simulation verschiedener Röntgenquellen oder verschiedener Detektormodelle ermöglicht. Im Kern verwendet *aRTist* einen Raytracing-basierten STL-Projektor, der die Schnittlängen der virtuellen Proben basierend auf STL-Modellen⁴ berechnet. Im zweiten Schritt werden die Schnittlängen unter Berücksichtigung des Röntgenspektrums, des Materials des Prüfkörpers und des Detektormodells in Intensitäten, typischerweise in J/m^2 , umgerechnet. Hauptsächlich wird hierzu das LAMBERT-BEER-Gesetz aus Gleichung 1.1 ausgeführt. Zusätzlich besteht die Möglichkeit, eine Monte-Carlo-basierte Berechnung der Projektionen mit reduzierter Auflösung durchzuführen, sodass realistische Wechselwirkungs- und Streuprozesse, beispielsweise die angesprochene Compton-Streuung, in der Simulation abgeschätzt werden können.

Trotz dieser vielversprechenden Funktionen hat *aRTist* zwei signifikante Nachteile. Erstens wird für den Betrieb der Simulation eine graphische Benutzeroberfläche benötigt, die zur Einrichtung und qualitativen Bewertung der Simulation zunächst praktisch erscheint, allerdings die Parallelisierung mehrerer Instanzen ineffizient gestaltet. Zur Berechnung großer Trainingsdatensätze entsteht durch die graphische Benutzeroberfläche ein Nachteil, da die simulierten Proben nicht einzeln betrachtet werden müssen und so die Rechenleistung ineffizient verbraucht wird. *aRTist* benötigt etwa 1.9 min zur Berechnung von 2000 Projektionen eines Zeilendetektors mit 2000 Pixeln und demnach bereits fast 660 Tage, um diese Berechnung für 250.000 unterschiedliche Phantome und zwei Energiekanäle durchzuführen. Wie in Abschnitt 5.1.2 beschrieben, muss die Simulation innerhalb eines sinnvollen Zeitrahmens einen angemessen dimensionierten Trainingsdatensatz erzeugen können, sodass *aRTist* für diesen Zweck ausscheidet. In den folgenden Kapiteln wird *aRTist* teilweise verwendet, um simulierte Phantome unter Einbezug einer realistischen Röntgenstrahlungsquelle, der realistischen Interaktion der Strahlung mit der Probe und einer realistischen Nachbildung des Detektors zu erzeugen, um quantitative Aussagen zu treffen. Somit wird die Stärke von *aRTist*, die realistische Simulation der CT, zur Validierung der im Folgenden erzielten Ergebnisse genutzt. Der zweite, jedoch nicht entscheidende, Nachteil von *aRTist* besteht in der Begrenzung auf

⁴ Ein Dateiformat zur Darstellung von Oberflächen im dreidimensionalen Raum basierend auf Dreiecken [100].

das Betriebssystem *Windows*, sodass der Betrieb auf Großrechnern, die typischerweise ein Betriebssystem auf der Basis von *Linux* verwenden, erschwert wird. Zudem muss *aRTist* erworben werden und ist nicht quelloffen, sodass die eigenständige Erweiterung der Funktionen nicht möglich ist.

5.2.2 *gVXR*

Eine vielversprechende Alternative zu dem soeben vorgestellten Programm *aRTist* besteht in *gVXR*, das quelloffen ist und kostenlos für die Forschung eingesetzt werden kann [101]. *gVXR* vereint in Summe die gleichen Funktionalitäten, mit Ausnahme der Monte-Carlo-basierten Berechnungen, und nutzt hierfür teilweise andere, ebenfalls quelloffene Pakete. So stammen die simulierten Quellspektren von *SpekPy*[102] und die tabellierten Abschwächungskoeffizienten von *xraylib* [103]. Analog zu *aRTist* werden die Phantome in *gVXR* als STL-Modelle eingeladen, die Schnittlängen bestimmt sowie das LAMBERT-BEER-Gesetz aus Gleichung 1.1 angewandt. Im Gegensatz zu *aRTist* werden die Schnittlängenberechnung sowie die Anwendung des LAMBERT-BEER-Gesetzes auf der Graphikkarte ausgeführt, sodass die Berechnungsgeschwindigkeit deutlich zunimmt. *gVXR* ist plattformunabhängig und kann mit der Programmiersprache Python angesteuert werden. Die Verwendung einer graphischen Benutzeroberfläche ist optional.

5.2.3 *Diskussion der Anforderungen*

Anhand der in Abschnitt 5.1 formulierten Anforderungen, werden die beiden vorgestellten Simulationen hinsichtlich ihrer Eignung zur Trainingsdatenerzeugung im Rahmen dieser Arbeit bewertet. Im Kontext von Abbildung 5.1 werden die einzelnen Kernfunktionalitäten abgearbeitet. Keins der vorgestellten Programme bietet die Möglichkeit der zufälligen Erzeugung von Phantomen (*Phantomgenerator*). In der Literaturrecherche (Kapitel 4.2.4) wurde eine Publikation identifiziert, die das XCAT-Phantom verwendet, um Trainingsdaten für die Materialbestimmung in der industriellen CT zu generieren [11]. Da das XCAT-Phantom für klinische Anwendungen entwickelt wurde [84], ist seine Morphologie an den menschlichen Körper angelehnt, sodass der Transfer auf industrielle Prüfteile potentiell aufgrund der abweichenden Morphologie erschwert wird.

Beide Programme aus Abschnitt 5.2 ermöglichen die Berechnung der Röntgenprojektionen (*Projektor*), sofern das Phantom als STL-Modell vorliegt. Die Materialkarten müssen als Voxelgitter vorliegen, damit diese mit den rekonstruierten CT-Volumen räumlich registriert sind (vergleiche Anforderungsanalyse Abschnitt 5.1). Somit bietet es sich an, die Erzeugung der Phantome ebenfalls in einem Voxelgitter auszuführen, sodass bei der Umrechnung in ein STL-Modell keine Verschiebungen und damit Registrierungsfehler auftreten können. Demnach scheiden sowohl *aRTist* als auch *gVXR* für den *Projektor* aus. Es wird ein voxelbasierter Projektor gesucht, sodass nicht zwischen Voxelgittern und STL-Modellen konvertiert werden muss. Rekonstruktionsalgorithmen verwenden voxelbasierte Projektoren, um die Rückprojektion durchzuführen (vergleiche Kapitel 2.4). Die *ASTRA Toolbox* (kurz: *ASTRA*) [104] ist eine Sammlung verschiedener Module, die

überwiegend auf die Rekonstruktion von CT-Daten ausgerichtet sind. Sie bietet zudem einen sogenannten *Vorwärtsprojektor*, der die Schnittlängen durch voxelbasierte Volumen berechnet. Durch Ergänzen der Abschwächungsberechnung nach dem LAMBERT-BEER-Gesetz aus Gleichung 1.1, einer Möglichkeit zur Simulation realistischer Quellspektren sowie dem Abruf der tabellierten Abschwächungskoeffizienten kann somit die Projektion des voxelbasierten Phantoms realisiert werden. Die Quellspektren können mit *SpekPy*[102] oder *aRTist* [24] berechnet werden. Die energieabhängigen, linearen Abschwächungskoeffizienten μ (vergleiche beispielhaft Abbildung 2.4) können der Literatur [26] entnommen oder mithilfe von *xraylib* [103] oder *aRTist* [24] berechnet werden.

Weiterhin wird gemäß Abbildung 5.1 eine Rekonstruktion benötigt, die die gesammelten Projektionen, zusammengefasst dargestellt als Sinogramm, in die DECT-Volumen rekonstruiert. Hierzu ist die bereits genannte *ASTRA Toolbox* [104] geeignet. Sie bietet neben dem aus Kapitel 2.4 bekannten FDK-Algorithmus ebenfalls iterative Rekonstruktionsalgorithmen an, die zu einer Verbesserung der Bildqualität führen können, jedoch aufgrund ihres immensen Rechenzeitbedarfs, insbesondere auf großen CT-Volumen, von niedriger Relevanz in der industriellen Computertomographie sind. Somit wird die gefilterte Rückprojektion, ein Spezialfall des FDK-Algorithmus, zur Rekonstruktion verwendet. Darüber hinaus ist die *ASTRA Toolbox* quelloffen und plattformunabhängig [104], sodass die nicht-funktionalen Anforderungen ebenfalls erfüllt werden.

Zusammenfassend bleiben die folgenden Module übrig, die implementiert werden müssen: ein Phantomgenerator und die Anwendung des LAMBERT-BEER-Gesetzes (Gleichung 1.1).

5.3 EIGENE SIMULATION

Basierend auf der Anforderungsanalyse aus Abschnitt 5.1 wird eine eigene Simulation implementiert. Gemäß Abschnitt 5.2.3 werden einige der gesuchten (Teil-)Funktionalitäten, die in anderen Programmen verfügbar sind, übernommen. Die im Folgenden vorgestellte Simulation besteht aus drei Modulen, die bereits in Abbildung 5.1 im Kontext der Anforderungsanalyse eingeführt worden sind: dem *Phantomgenerator*, dem *Projektor* und der *Rekonstruktion*.

Gemeinsam werden diese drei Module in eine übergeordnete Funktion eingebettet, die den gesamten Ablauf, beginnend mit einer Zufallszahl zur Initialisierung des Phantomgenerators bis hin zum Speichern eines Datentupels, bestehend aus einem DECT-Volumen und den Materialkarten, zusammenfasst. Durch die Verwendung von *Dask*, einer Python-Bibliothek für verteiltes Rechnen [105], wird die parallele Ausführung dieser Funktion zur Erzeugung eines einzelnen Datentupels in voneinander unabhängigen Instanzen stark vereinfacht. Kurz gefasst kann Dask verwendet werden, um eine große Anzahl kleiner Aufgaben (hier: die Erzeugung eines Datentupels) in einem Netzwerk aus sogenannten *Workern* (auch: *Dask-Worker*) zu verteilen und die Ergebnisse zu sammeln. Mithilfe von Dask ist somit die im Folgenden beschriebene Simulation zwischen einem Laptop und

einem HPC-Cluster beliebig skalierbar⁵. In den folgenden Abschnitten werden die Implementierungen des *Phantomgenerators* und des *Projektors* aus dem Ablaufdiagramm in Abbildung 5.1 beschrieben. Die *Rekonstruktion* wird direkt aus der *ASTRA Toolbox* übernommen. Abgesehen von einigen Paketen in den verwendeten Bibliotheken ist die Simulation in Python implementiert.

Phantomgenerator: Erzeugung eines zufälligen Phantoms

Eine der DL-spezifischen Anforderungen nach Abschnitt 5.1 besteht in der Erzeugung zufällig geformter Phantome, die aus verschiedenen, ebenfalls zufällig ausgewählten Materialien zusammengesetzt sind. Im Folgenden werden die von einzelnen Materialien besetzten Bereiche in einem Phantom als *Regionen* oder *Materialregionen* bezeichnet. Der im Folgenden beschriebene Algorithmus ist gemäß den Anforderungen aus Abschnitt 5.1 für die Erzeugung zweidimensionaler Phantome konzipiert. Für dreidimensionale Phantome ist der Algorithmus direkt erweiterbar, wobei die benötigte Rechenzeit sowie der Speicherbedarf linear mit der Gesamtzahl der Voxel ansteigen.

Das im Rahmen dieser Arbeit gewählte Prinzip zur Erzeugung der zufälligen Regionen entstammt der Biologie: dem Wachstum von unterschiedlichen Bakterienkolonien in einer Petrischale. Zu Beginn wird die leere Petrischale mit einer Anzahl unterschiedlicher Bakterien an zufälligen Stellen punktuell besetzt. Dies sind die Startpunkte des Wachstums jeder Kolonie (Materialregion). Pro Zeitschritt wächst jede Kolonie zufällig in angrenzende und bisher unbesetzte Gitterpunkte (Voxel). Sobald ein festgelegter Anteil der Gesamtfläche der Petrischale von den Bakterienkolonien besetzt worden ist, wird das Wachstum eingefroren und das Experiment (die Erzeugung der Materialregionen) beendet. Codeabschnitt 5.1 fasst den soeben beschriebenen Algorithmus zur Erzeugung zufälliger Regionen innerhalb eines zweidimensionalen Bereiches zusammen.

```

1  volume = Array2D(yDim, xDim)
2  seed_points = Random(number_of_seed_points, volume)
3  while volume.infill < target_infill:
4      for seed in random_choice(seed_points):
5          seed.grow()

```

Codeabschnitt 5.1: Globalgorithmus zur Erzeugung der zufälligen Regionen eines Phantoms.

Nach der Ausbreitungsphase der Bakterienkolonien wird jede Kolonie, identifiziert durch eine eindeutige *ID*, zufällig in ein Material (Kernladungszahl Z und absolute Dichte ρ) umgewandelt:

$$\text{ID} \mapsto (Z_i, \eta\rho_i) \quad \text{mit} \quad \eta \in [0.8, 1.0].$$

Der Faktor η wird zufällig gewählt und stellt die Porosität eines Materials dar. Die Auswahl des Materials erfolgt ebenfalls zufällig aus einem gegebenen Materialpool, der im Rahmen dieser Thesis in Tabelle 5.1 zusammengefasst ist. Insbesondere die Materialien Aluminium, Chrom, Eisen, Kupfer, Magnesium, Nickel, Titan und Zink sind von hoher

⁵ Die Mindestanforderungen für das Ausführen einer einzelnen Simulation umfassen etwa 1 GB verfügbaren Haupt- und Graphikspeicher. Mehr Speicher ermöglicht eine erhöhte Parallelisierung und Auslastung.

Symbol	Al	Be	Ca	Co	Cr	Cu	Fe	Ga	Ge	K	Li
Z	13	4	20	27	24	29	26	31	32	19	3
ρ in g/cm ³	2.7	1.8	1.6	8.9	7.2	9.0	7.9	5.9	5.3	0.9	0.5
Symbol	Mg	Mn	Ni	Sc	Ti	V	Zn				
Z	12	25	28	21	22	23	30				
ρ in g/cm ³	1.7	7.3	8.9	3.0	4.5	6.1	7.1				

Tabelle 5.1: Symbole, Kernladungszahlen Z und Dichten ρ der selektierten Materialien für die Phantome der CT-Simulation.

technischner Relevanz, sodass sie in den Materialpool aufgenommen werden. Die auf dem Periodensystem benachbarten Elemente werden zur Verbesserung der Generalisierung auf ähnliche Materialien ergänzend hinzugenommen. Für eine spezifische Anwendung der Materialbestimmung mit Vorwissen, beispielsweise zur gezielten Identifikation von Kupferpartikeln, kann es zielführend sein, den Materialpool deutlich einzuschränken. Das Ergebnis des Phantomgenerators ist insgesamt eine zweidimensionale Materialkarte (vergleiche Abbildung 5.1), die für jeden Gitterpunkt (Voxel) eine Kernladungszahl Z und eine Dichte ρ angibt.

Projektor: Projektion des Phantoms

Jedes nach obigem Prinzip erzeugte Phantom wird durch eine Materialkarte beschrieben. Ziel dieses Moduls ist die Berechnung der energie- und materialabhängigen Projektion des Phantoms in Anlehnung an die reale CT-Bildaufnahme. Ein Überblick über diese Berechnung wird in Abbildung 5.2 dargestellt.

Zunächst werden aus den materialaufgelösten Phantomen unter Angabe der CT-Geometrie Schnittlängen mithilfe eines Schnittlängenprojektors \mathcal{P} aus der *ASTRA Toolbox* berechnet. Die resultierenden Schnittlängen hängen somit zusätzlich vom Projektionswinkel φ aus der CT-Geometrie und einem Pixelindex u des eindimensionalen Detektors (Fächerstrahl-CT) ab. Im nächsten Schritt wird aus dem einfallenden Röntgenspektrum $I_0(E)$ sowie den tabellierten Abschwächungskoeffizienten μ_m für jedes Material m das **LAMBERT-BEER**-Gesetz aus Gleichung 1.1 angewandt, um das Spektrum $I_P(E)$ nach Abschwächung durch das Phantom zu berechnen. Aus dem Abschwächungskoeffizienten des Detektormaterials μ_D und der Schnittlänge durch ein Pixel $x_D^{(u)}$ des Detektors wird die Intensität $I_D(E)$ nach zusätzlicher Abschwächung durch den Detektor bestimmt. Die Differenz $I_P(E) - I_D(E)$ entspricht der Intensität, beziehungsweise der Photonenzahl, aufgelöst nach Energie, die das Signal im Detektor $I_{\text{Signal}(E)}$ erzeugt. Nach spektraler Integration $\int dE$ für einen energieintegrierenden Detektor nach Kapitel 2.3 folgt das gemessene Signal I_{Signal} – das sogenannte *Sinogramm*. Die gesamte in Abbildung 5.2 beschriebene Berechnung ist hochgradig parallelisierbar. Nach der Projektion zur Berechnung der Schnittlängen wird jede folgende Operation pro Pixel ausgeführt und ist unabhängig von den anderen Pixeln, sodass die Parallelisierung auf einer Graphikkarte effizient implementiert werden kann.

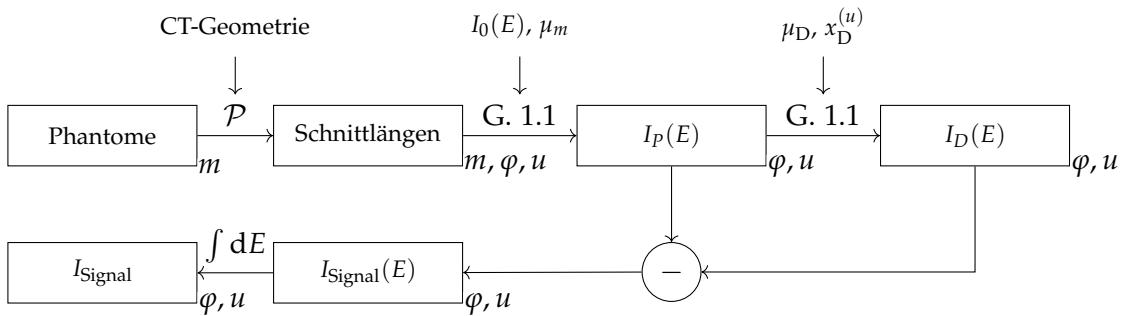


Abbildung 5.2: Ablaufdiagramm der CT-Projektion ausgehend von m materialauflösten Phantomen mit dem Schnittlängenoperator \mathcal{P} unter einer gegebenen CT-Geometrie mit dem Projektionswinkel φ und für jedes Detektorpixel u eines eindimensionalen Detektors. Das erzeugte Quellspektrum $I_0(E)$ sowie die durch Phantom und Detektor abgeschwächten Spektren $I_P(E)$ und $I_D(E)$ werden zusammen mit den Abschwächungskoeffizienten μ zur Anwendung des LAMBERT-BEER-Gesetzes benötigt. Nach Subtraktion und spektraler Integration folgt das gemessene Detektorsignal I_{Signal} .

5.4 EVALUATION DER IMPLEMENTIERTEN SIMULATION

Die implementierte Simulation muss hinsichtlich der in Abschnitt 5.1 aufgeführten Anforderungen untersucht werden. Einige Anforderungen, beispielsweise das Datenformat, wurden bereits in der Implementierung entsprechend erfüllt und bedürfen keiner weiteren Prüfung. Exemplarisch wird die Ausgabe der Simulation sowie die statistische Verteilung eines Trainingsdatensatzes in Abschnitt 5.4.1 vorgestellt. Anschließend muss die Präzision des modellierten CT-Bildgebungsprozesses im Vergleich zu einer verbreiteten Simulation (*aRTist*) sowie einem realen CT-Scan verglichen werden. Die DL-spezifischen Anforderungen in Bezug auf die Datensatzgröße in Abhängigkeit zur Berechnungsdauer werden in Abschnitt 5.4.4 überprüft.

5.4.1 Exemplarischer Datensatz & Datenverteilung

Anhand eines exemplarisch simulierten Datensatzes können einige der funktionalen Anforderungen aus Abschnitt 5.1 an die implementierte Simulation evaluiert werden. Die verwendeten CT-Parameter werden in Tabelle 5.2 zusammengefasst und sind dem Modell einer realen CT-Anlage nachempfunden (vergleiche CT-Anlage in Kapitel 2.7). Im Vergleich zum realen Detektor werden in diesem CT-Parametersatz vier Pixel des Originaldetektors mit dem Pixel Pitch 139 µm zu einem Pixel mit dem Pixel Pitch 278 µm zusammengefasst, um Rechenzeit auf Kosten der räumlichen Auflösung zu sparen. Durch die Verdopplung des Pixel Pitch, kann die Anzahl der Pixel pro Zeile (im dreidimensionalen Fall auch pro Spalte) halbiert werden, sodass die CT-Abbildungsgeometrie in allen anderen Parametern konstant bleibt und somit die simulierten Abschwächungskoeffizienten quantitativ mit den an der realen CT-Anlage gemessenen Werten vergleichbar sind. Zusätzlich wird, begründet durch die DL-spezifischen Anforderungen aus Abschnitt 5.1, kein zweidimensionaler Detektor, sondern nur eine einzelne Zeile des Detektors simuliert, da die so gesammelten eindimensionalen Projektionen ausreichen, um eine

einzelne CT-Mittelschicht zu rekonstruieren (Fächerstrahl-CT). Zunächst wird die Ausga-

Geometrie	
Anzahl Projektionen	1500 (Vollkreis)
SDD	500 mm
SOD	35 mm
Detektor	
Detektor-Typ	Flachdetektor, 200 μm Caesiumiodid
Pixel Pitch	278 μm
Pixelmatrix Dimensionen	(1500 \times 1) Pixel
Pixeldatentyp	16 Bit unsigned-int
Integrationszeit	> 100 ms
Strahlungsquelle	
Typ	Wolfram-Transmissionsquelle
U_B + Vorfilter (Low-Energy)	140 kV + 1.0 mm Al + 0.5 mm Cu
U_B + Vorfilter (High-Energy)	240 kV + 1.0 mm Cu + 1.0 mm Sn
Phantome	
Anzahl Datentupel	250000
Seitenlänge des Volumens	1.0 cm
Auflösung	(256 \times 256) Voxel
Anzahl Regionen N	$0 < N \leq 1000$ (gleichverteilt)
Ausfüllfaktor (vgl. <code>target_infill</code>)	90%
relative Dichte η	$0.8 \leq \eta \leq 1.0$

Tabelle 5.2: CT-Simulationsparameter für einen DECT-Scan mit dem Abstand zwischen Quelle und (Objekt-)Drehachse SOD , dem Abstand zwischen Quelle und Detektor SDD und der Beschleunigungsspannung U_B . Die Einstellungen der Strahlungsquelle werden getrennt für den Low- und High-Energy-Scan angegeben, wobei die Integrationszeit am Detektor jeweils so gewählt wird, dass der Dynamikumfang im Mittel 70% des Datenraumes des Detektors 0 – 65535 abdeckt.

be der implementierten Simulation visualisiert. Abbildung 5.3 zeigt ein mit den genannten Parametern simuliertes Tupel, bestehend aus den DECT-Schnitbildern und den Materialkarten, aufgelöst nach Kernladungszahl und Dichte. Der Phantomgenerator erzeugt die Materialkarten in der rechten Spalte in Abbildung 5.3 zufällig. Mithilfe des Projektors und der Rekonstruktion werden die Phantome (Materialkarten) in die DECT-Schichtbilder in der linken Spalte umgerechnet. Diese etwas kontraintuitive Anordnung – von rechts nach links – wird anhand des zu lernenden Problems, der Materialzerlegung, motiviert: Die DECT-Schichtbilder (links) sollen in die Materialkarten (rechts) umgerechnet werden. Aus den DL-spezifischen Anforderungen geht weiterhin hervor, dass die generierten Phantome von zufälliger Materialzusammensetzung und Form sein müssen, um erstens

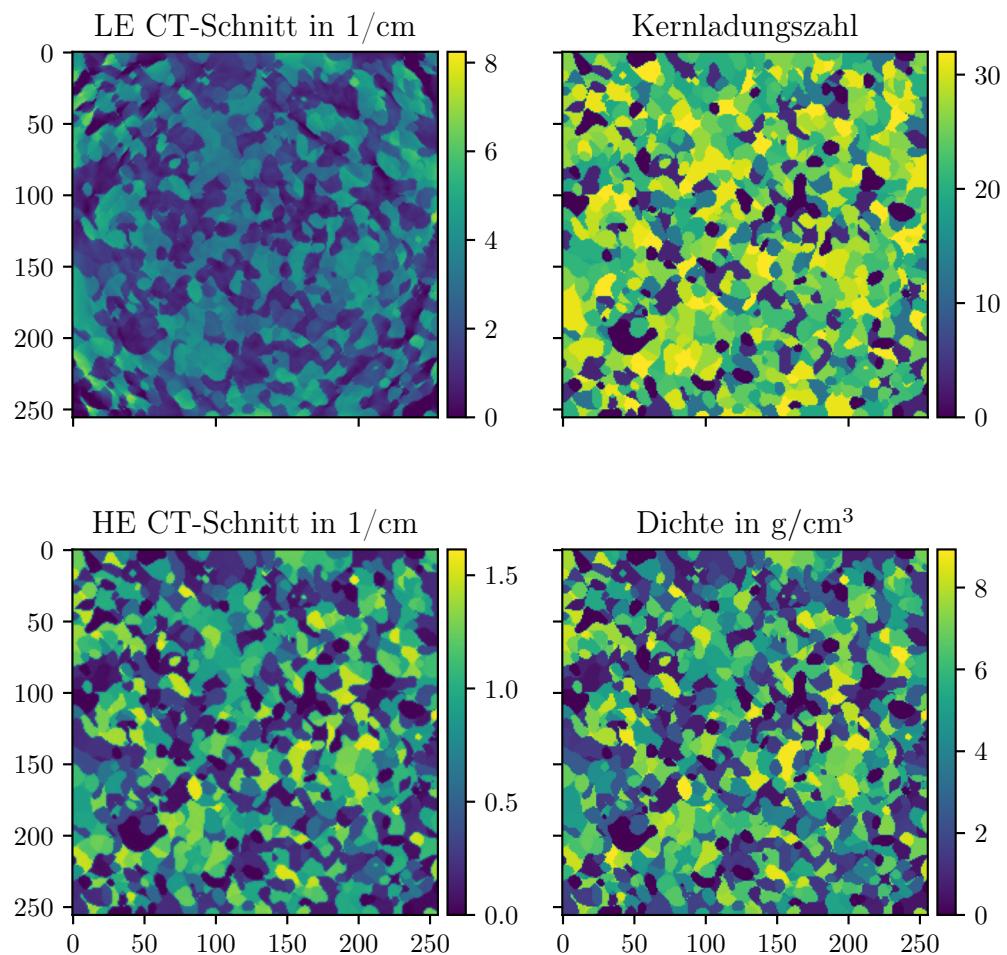


Abbildung 5.3: Simuliertes Datentupel bestehend aus Low-Energy (LE) und High-Energy (HE) CT-Schnitten (linke Spalte) durch ein zufällig erzeugtes Phantom. Das Phantom wird durch die Angabe der Kernladungszahlen und Dichten räumlich charakterisiert (rechte Spalte). Alle x- und y-Achsen geben räumliche Positionen im Voxelgitter an. Die Einheiten der Farbskalen werden in den Plotüberschriften angegeben.

eine große Auswahl unterschiedlicher Materialien bestimmen zu können und zweitens die morphologische Varianz innerhalb des Datensatzes zu erhöhen. Anhand des exemplarisch simulierten Datensatzes kann mittels einer Stichprobe die Zusammensetzung des gesamten Datensatzes statistisch untersucht werden. Die Verteilungen in Abbildung 5.4 wurden aus 8000 zufällig ausgewählten Tupeln berechnet und geben einen Hinweis auf die Verteilung der Materialien innerhalb der Stichprobe. Zu sehen sind die quantisierten, nahezu gleichverteilten Kernladungszahlen (oben rechts) der für die Simulation ausgewählten Elemente. Die relativen Dichten η werden aus dem Bereich $0.8 \leq \eta \leq 1.0$ zufällig und gleichverteilt gewählt, sodass im Histogramm (unten rechts) der absoluten Dichten eine Streuung beobachtet werden kann. In den Histogrammen der linken Spalte von Abbildung 5.4, den effektiv gemessenen Abschwächungskoeffizienten, wird deutlich, dass die Strahlaufhärtung einen signifikanten Einfluss auf die Datenzusammensetzung ausübt. Wie bereits in Kapitel 2.4.1 beschrieben, hängt der rekonstruierte Abschwächungskoeffizient μ_L stark von der lokalen Umgebung innerhalb der Probe ab und ist keineswegs konstant für ein festes Material. Zur Beurteilung der morphologi-

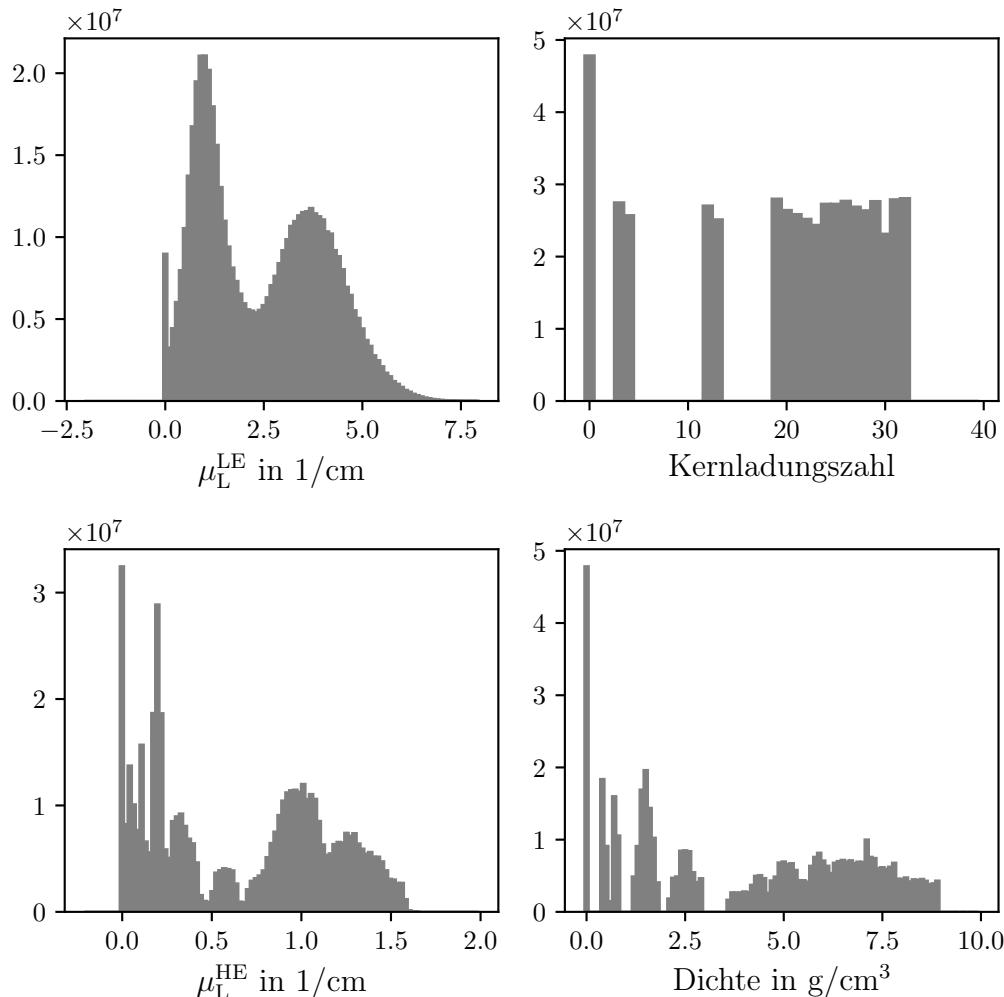


Abbildung 5.4: Histogramme der auftretenden Abschwächungskoeffizienten μ_L für den Low-Energy LE und High-Energy HE Kanal sowie die Kernladungszahlen und Dichten aus der in Tabelle 5.2 beschriebenen Simulation anhand einer Stichprobe.

schen Varianz der einzelnen Phantome ist Abbildung 5.4 ungeeignet. Eine detaillierte Analyse der vorkommenden Formen kann sehr aufwändig gestaltet werden, sodass im Rahmen dieser Arbeit eine vereinfachte Methode angewandt wird: Die Phantome werden mit dem Canny-Edge-Operator [106] in Kantenbilder umgerechnet und anschließend normiert. Bei geeigneter Parameterwahl des Canny-Edge-Operators erscheinen so die Kanten zwischen den unterschiedlichen Regionen innerhalb eines Phantoms gleich und sind nicht abhängig von ihrem Gradienten. Anhand der oben genannten Stichprobe kann die Summe über alle Kantenbilder berechnet werden und so sollten feste Muster, die immer an der gleichen räumlichen Position in den Phantomen auftreten, sichtbar werden. Abbildung 5.5 stellt das Summenbild über alle Kantenbilder der Stichprobe dar. Der dunkel abgesetzte Randbereich ist eine Folge der räumlichen Begrenzungen an den

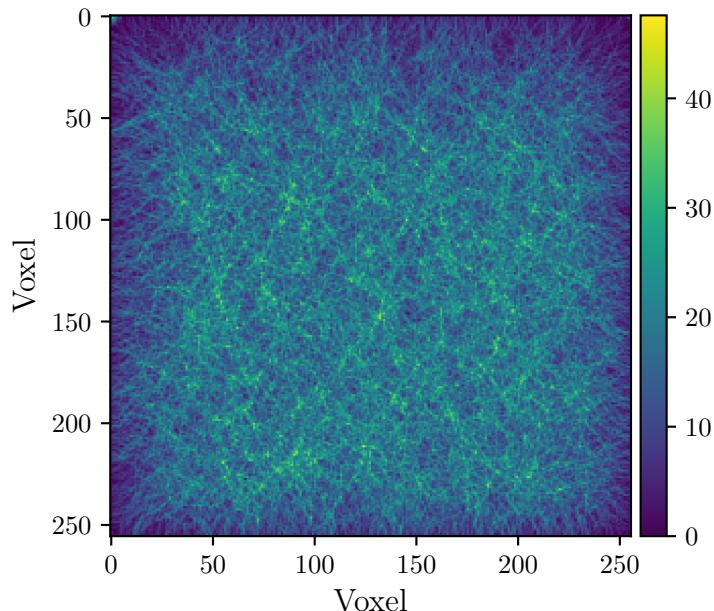


Abbildung 5.5: Summe aus den Kantenbildern ermittelt durch den Canny-Edge-Operator mit anschließender Normierung aus 8000 zufällig ausgewählten Phantomen. Die Farbskala hat keine Einheit und ist als eine *Häufigkeit* zu interpretieren.

Rändern der Phantome. Statistisch ist es wahrscheinlicher, dass mittlere Bildbereiche von verschiedenen Regionen besetzt werden als Randbereiche. Ansonsten kann kein eindeutiges Muster identifiziert werden, sodass insgesamt geschlussfolgert werden kann, dass die Erzeugung der zufälligen Phantome, in Bezug auf Form und Materialzusammensetzung, mit der implementierten Simulation erfolgreich ist.

5.4.2 Vergleich mit *aRTist* und mit einem realen Scan

Weiterhin sind die CT-spezifischen Anforderungen, insbesondere in Bezug auf die Präzision des CT-Bildgebungsprozesses, zu prüfen. Dies erfolgt anhand eines Vergleiches mit der in Abschnitt 5.2.1 vorgestellten Simulation *aRTist* sowie einem realen CT-Scan. Ein

direkter Vergleich zwischen der implementierten Simulation und *aRTist* ist in Abbildung 5.6 dargestellt. Das simulierte Phantom ist ein Aluminiumwürfel, der verschiedene Por-

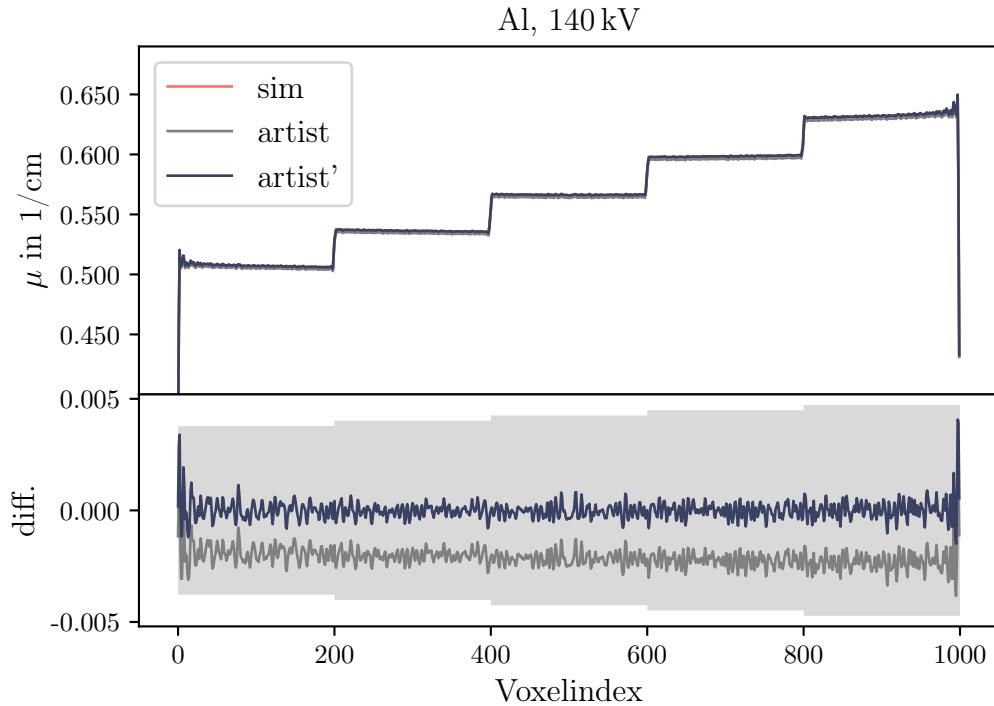


Abbildung 5.6: Vergleichender Schnitt durch einen Aluminiumwürfel mit verschiedenen relativen Dichten $0.8 \leq \eta \leq 1.0$ simuliert mit *aRTist*, der implementierten Simulation *sim* und der korrigierten *aRTist*-Simulation *aRTist'* aus Gleichung 5.1. Durch die starke Übereinstimmung der Kurven ist die rote Linie nur schlecht erkennbar. Zur besseren Vergleichbarkeit werden im unteren Teil der Abbildung die absoluten Differenzen *diff.* zwischen *aRTist* beziehungsweise *aRTist'* mit *sim* visualisiert. Der grau schattierte Bereich markiert die Zielgenauigkeit ($\pm 1\%$ -Genauigkeit der Dichte nach Gleichung 2.3).

sitäten aufweist, sodass im CT-Schnitbild ein Stufenprofil sichtbar wird. Ein konstanter Offset ist zwischen der implementierten Simulation *sim* und *aRTist* in der unteren Darstellung von Abbildung 5.6 qualitativ zu beobachten (graue Linie). Dieser Offset ist für verschiedene Materialien, beziehungsweise verschiedene Abschwächungskoeffizienten, nicht konstant. Daher wird ein lineares Modell der Form

$$y = m \cdot x + b \quad \text{mit} \quad \begin{cases} y : \text{aRTist-Daten} \\ x : \text{sim-Daten} \end{cases} \quad (5.1)$$

angesetzt, um die Abweichung zu korrigieren. Es ergeben sich die Koeffizienten $m = 0.9912$ und $b = 0.0113$ für den in Abbildung 5.6 dargestellten Fall und unter zusätzlicher Berücksichtigung der Materialien Magnesium, Eisen und Kupfer. Durch Anwendung der Offset-Korrektur aus Gleichung 5.1 ergibt sich das Profil *aRTist'* in Abbildung 5.6 sowie die dargestellte Abweichung unter dem Plot (blaue Kurve). Es erscheint zunächst unintuitiv, die mit *aRTist* simulierten Werte an die implementierte Simulation anzupassen, da *aRTist* im Rahmen dieser Dissertationsschrift als verbreitete und etablierte CT-Simulation vorgestellt wurde. Durch die Verwendung realer CT-Daten kommt in späteren Kapiteln

eine dritte Datenquelle hinzu, für die erneut ein Korrekturmödell nach Gleichung 5.1 angepasst werden muss. Zur besseren Vergleichbarkeit der Materialbestimmung wird daher die implementierte Simulation als gemeinsame Basis gewählt, sodass die mit aRTist simulierten Daten sowie die Daten von der CT-Anlage in die Basis der implementierten Simulation übertragen werden müssen. Die korrigierten CT-Schnitbilder nach aRTist und nach der implementierten Simulation werden für die genannten vier Materialien in Abbildung 5.7 zusammengefasst. Die Plots der oberen Zeile in Abbildung 5.7 zeigen die leicht dämpfenden Materialien Aluminium und Magnesium, die keine starken CT-Artefakte aufweisen. In den Plots der unteren Zeile ist deutlich der Effekt der Strahlauflaufhärtung zu erkennen, der auf die verhältnismäßig niedrige Beschleunigungsspannung $U_B = 140 \text{ kV}$ zurückzuführen ist und korrekt von der implementierten Simulation dargestellt wird. Der grau schattierte Bereich markiert die geforderte Genauigkeit ($\pm 1\%$ -Genauigkeit der Dichte) und wird durch die Fehlerfortpflanzung von Gleichung 2.3 individuell für jedes Material berechnet. Die korrigierten Abschwächungskoeffizienten liegen nach Anwendung der Korrektur innerhalb der geforderten Genauigkeiten. Basierend auf diesen Ergebnissen und im Kontext der geforderten Genauigkeit der linearen Abschwächungskoeffizienten (vergleiche Abschnitt 5.1), wird Gleichung 5.1 als ausreichend gewertet und fortan zum Angleich der mit aRTist simulierten Daten an die Daten der implementierten Simulation verwendet. Die Koeffizienten ebendieser Gleichung sind stets für einen Satz gegebener CT-Parameter, insbesondere Quellspannung und Vorfilterung, individuell zu ermitteln und demnach nicht für andere CT-Parameter gültig.

Zum Abgleich der simulierten CT-Daten mit dem Scanergebnis einer realen CT-Anlage (im Folgenden zur Vereinfachung auch: *Anlage*) müssen mehrere Dinge beachtet werden. Zunächst müssen die CT-Geometrie der Anlage, das emittierte Quellspektrum sowie der Aufbau des Detektors möglichst genau erfasst werden. Während die CT-Geometrie durch Kalibrationsmessungen mit metrologisch eingemessenen Prüfkörpern sehr exakt bestimmt werden kann, ist über den inneren Aufbau der Röntgenquelle und des Detektors nur wenig bekannt. Somit basieren die berechneten Quellspektren und Detektorantworten teilweise auf Angaben der Datenblätter, gut begründeten Annahmen sowie interner Kommunikation mit den Herstellern. Die genauen Angaben zur Berechnung der verwendeten Quellspektren und der Detektorantwort dürfen daher in dieser Dissertationsschrift nicht publiziert werden. Das rekonstruierte Volumen des Scans wird manuell segmentiert und anhand der Segmentierung in ein voxelbasiertes Phantom umgerechnet, welches durch den in Abbildung 5.1 vorgestellten Ablauf beginnend ab dem *Projektor* verarbeitet werden kann. Mithilfe dieser Methodik wird das real gescannte Phantom als Eingabeparameter der implementierten Simulation genutzt, sodass ein quantitativer Abgleich der rekonstruierten Volumen möglich ist.

Der an der Anlage durchgeführte Scan wird mit einem Aluminiumquader, einem Edelstahlzylinder und einem Kupferquader gemeinsam mit den CT-Parametern aus Tabelle 5.2 durchgeführt. Die rekonstruierten Volumina des Scans und der Simulation sind für beide Energien (140 kV und 240 kV gemäß Tabelle 5.2) in Abbildung 5.8 dargestellt. Quantitativ können bereits einige Unterschiede in Bezug auf die Abschwächungskoeffizienten innerhalb und außerhalb der verschiedenen Objekten identifiziert werden. Besonders

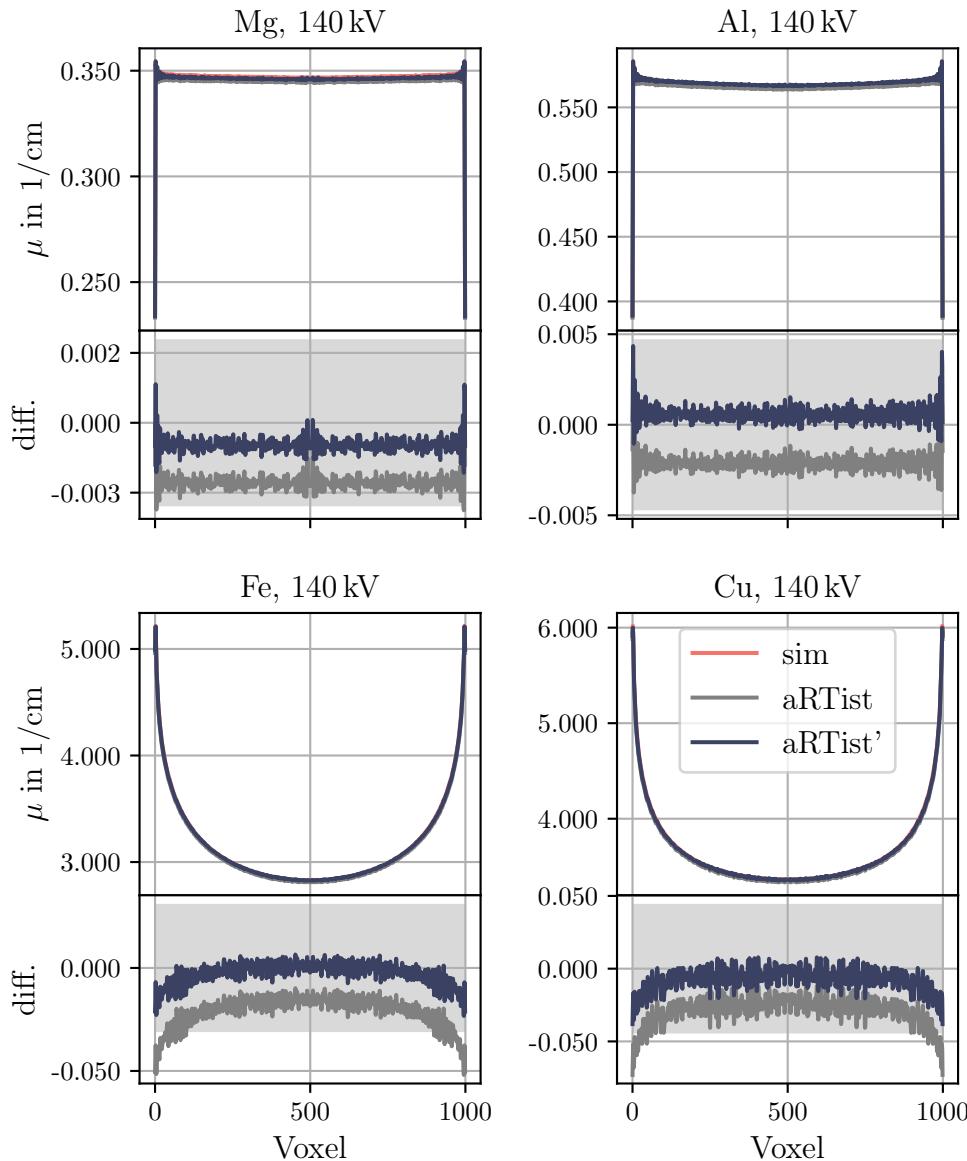


Abbildung 5.7: Schnitte durch die rekonstruierten Volumina eines Phantoms für verschiedene Materialien mit der entwickelten Simulation *sim* und *aRTist* unter Berücksichtigung der linearen Korrektur (*aRTist'*) aus Gleichung 5.1. Unter den Plots werden die absoluten Differenzen *diff.* zusammen mit der grau schattierten Zielgenauigkeit ($\pm 1\%$ -Genauigkeit der Dichte nach Gleichung 2.3) dargestellt. Neben dem Element wird ebenfalls die Quellspannung angegeben. Alle Plots teilen die Legende unten rechts. Die rote Linie wird aufgrund der Übereinstimmung mit den anderen Linien teilweise durch diese verdeckt.

auffällig ist die Abweichung der CT-Artefakte außerhalb der Objekte, die hauptsächlich durch die fehlende Modellierung der Streustahlung (vergleiche Compton-Streuung; Kapitel 2.1) in der implementierten Simulation entsteht. Da die Strahlaufl härtung korrekt modelliert wird und in Abbildung 5.8 (oben rechts) deutlich zu erkennen ist, wird deutlich, dass die verbleibende Abweichung innerhalb der Materialien zur realen Messung,

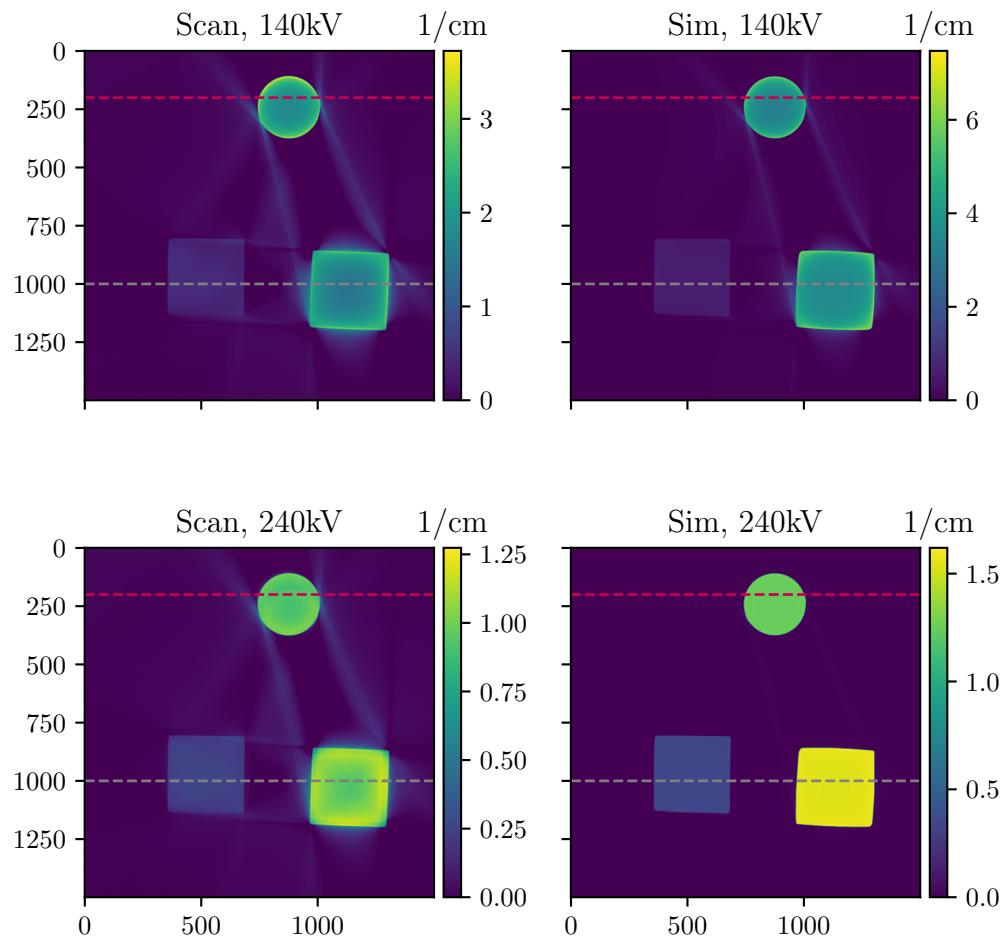


Abbildung 5.8: Rekonstruierte Volumina für zwei Beschleunigungsspannungen aus der CT-Anlage beschriftet mit *Scan* (links) und der implementierten Simulation *Sim* (rechts). Der Vollzylinder besteht aus dem Edelstahl V2A. Der linke Quader besteht aus Aluminium und der rechte Quader besteht aus Kupfer. Die graue Linie zeigt das Referenzprofil an, um das in Gleichung 5.1 vorgestellte lineare Modell anzupassen. Die rote Linie markiert das Testprofil, um die Interpolationsfähigkeit des linearen Modells nachzuweisen. Alle x- und y-Achsen geben räumliche Positionen im Voxelgitter an.

insbesondere bei höheren Photonenenergien in Abbildung 5.8 (unten), auf die Streustrahlung zurückzuführen ist. Entlang der grauen Linie werden die Werte des Scans mithilfe des eben beschriebenen linearen Modells aus Gleichung 5.1 an die simulierten Werte angeglichen. Somit werden die Materialien Aluminium und Kupfer zur Anpassung des Modells verwendet. Zu prüfen ist die Interpolationsfähigkeit des linearen Modells entlang der roten Linie (Edelstahl V2A) – dem Testprofil, das aus Sicht der Kernladungszahl und Dichte zwischen den Stützpunkten des linearen Modells liegt. Abbildung 5.9 zeigt den Schnitt durch die Volumina entlang des Testprofils durch den Edelstahlzyylinder. Das Modell verschiebt das ursprünglich stark abweichende Scanprofil (gelb) in die Nähe des simulierten Profils (rot). Wie bereits beschrieben, gibt es bei der Modellierung einer realen CT-Anlage Unsicherheiten, vorwiegend induziert durch die Röntgenquelle und den Detektor, die nicht vollständig korrigiert werden können. Trotz der weiterhin

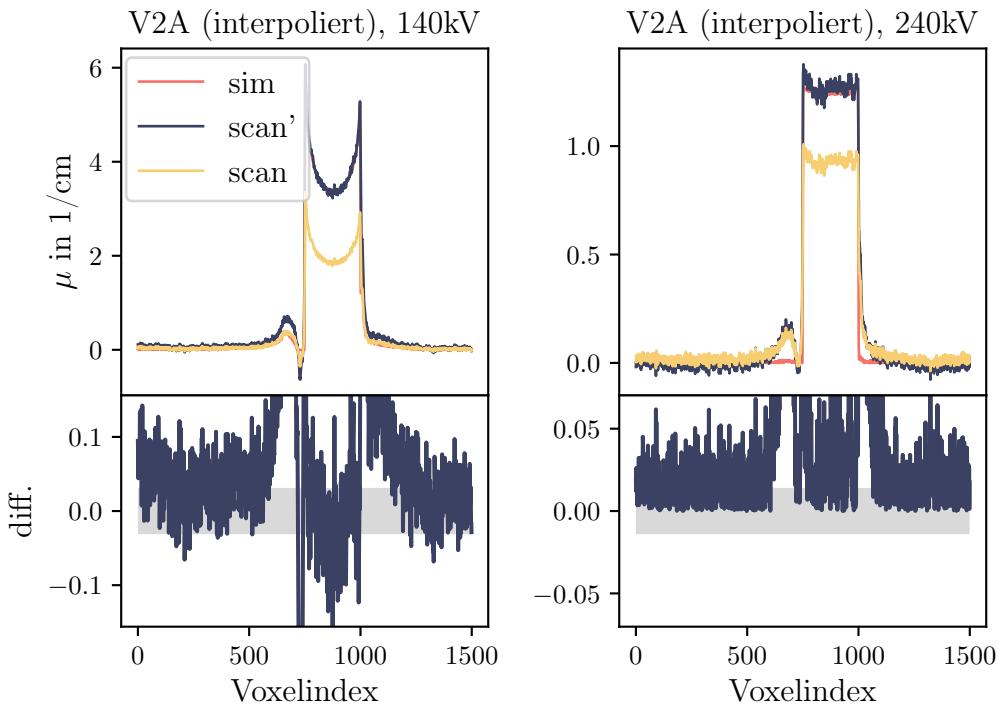


Abbildung 5.9: Schnitt entlang der roten Linie in Abbildung 5.8 zur Evaluation der Transferfähigkeit des linearen Modells auf Edelstahl V2A. Das auf der grauen Linie in Abbildung 5.8 angepasste Modell wird verwendet. Unter den Plots werden die absoluten Differenzen $diff.$ zwischen $scan'$ und sim zusammen mit der grau schattierten Zielgenauigkeit ($\pm 1\%$ -Genauigkeit der Dichte nach Gleichung 2.3) dargestellt. Die absoluten Differenzen weichen deutlich von der geforderten Zielgenauigkeit ab.

bestehenden Abweichung, der effektiven Sim-To-Real-Gap, zwischen dem korrigierten Profil $scan'$ und der implementierten Simulation sim , wurde mithilfe der Korrektur dieser ursprünglich noch viel stärkere Unterschied erheblich reduziert. Im Folgenden wird das angepasste lineare Modell zur Anpassung der realen Messdaten an die Simulation verwendet, um die Sim-To-Real-Gap zu verkleinern und somit die Materialbestimmung mithilfe der trainierten DL-Modelle zu verbessern.

5.4.3 Beobachtungen mit (simulierter) Dual-Energy CT

Wie bereits im Rahmen der Einführung der Strahlauflärtung (vergleiche Kapitel 2.4.1; insbesondere Abbildung 2.11) beschrieben, werden bei der Verwendung polychromatischer Röntgenstrahlung homogene Objekte nicht mit konstanten Abschwächungskoeffizienten rekonstruiert. Somit verändert sich der in Abbildung 2.14 dargestellte Plot, in dem jedes Material durch eine separate, scharfe Linie beschrieben wird, mit zunehmendem Einfluss der Strahlauflärtung und weiterer CT-Artefakte deutlich. Dieser Effekt soll, in Anlehnung an die Darstellung aus Abbildung 2.14, anhand simulierter Abschwächungskoeffizienten einzelner Materialien beobachtet werden, um erstens die Funktionalität der polychromatischen CT-Simulation zu überprüfen und um zweitens die Problemstellung der Materialbestimmung weiter zu verdeutlichen. Abbildung 5.10 zeigt die DECT-Linien

für die in Tabelle 5.2 beschriebene Simulation. Im linken Plot von Abbildung 5.10 ist

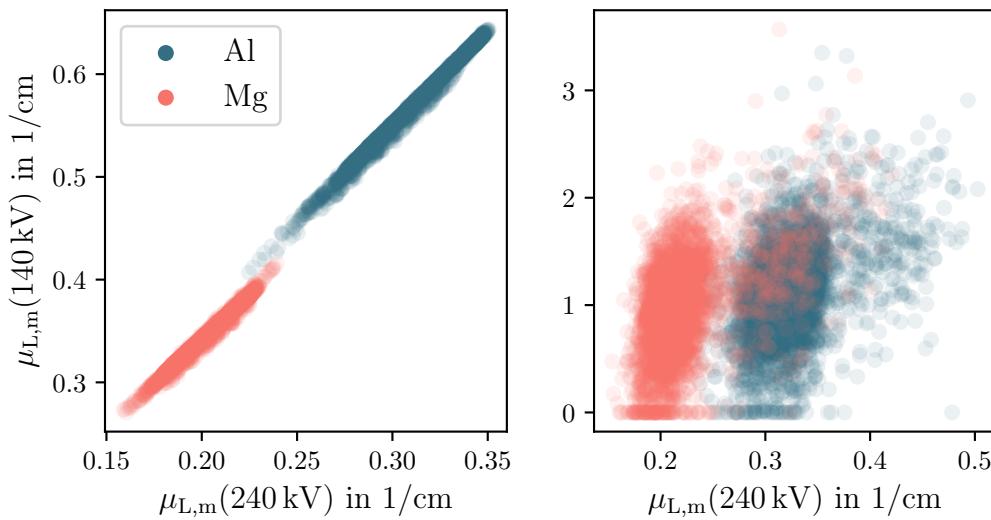


Abbildung 5.10: Aus einer Simulation extrahierte, materialabhängige Abschwächungskoeffizienten $\mu_{L,m}$ für zwei Quellspannungen 140 kV und 240 kV. Die Werte im linken Plot entstammen einer Simulation nur mit den Materialien Aluminium und Magnesium, während die Werte im rechten Plot aus einer Simulation mit den zusätzlichen Materialien Eisen und Kupfer entstammen. Es wurden jeweils 3000 Voxel zufällig ausgewählt. Die einzelnen Datenpunkte sind leicht transparent eingezeichnet, um die Dichte übereinander liegender Punkte zu visualisieren.

die Aufweichung der scharfen DECT-Linien sichtbar – die Linien erscheinen nunmehr als Punktewolken, die für die gezeigten Materialien Aluminium und Magnesium noch eindeutig trennbar sind. Der rechte Plot zeigt den Einfluss der Materialien Eisen und Kupfer (nicht im Plot) neben Aluminium und Magnesium in den Phantomen. Durch die starke Abschwächung der sehr viel schwereren Elemente Eisen und Kupfer werden die DECT-Punktewolken von Aluminium und Magnesium weiter zerstreut. An der y-Achse ist ebenfalls Clipping am Nullpunkt zu erkennen, da rekonstruierte Voxel mit negativen Werten per Definition in der Simulation geclipped werden.⁶ Analog zeigt Abbildung 5.11 die DECT-Plots für die Materialien Eisen und Kupfer. Auch hier ist eine deutliche Zerstreuung der Punktewolke sowie eine teilweise Überlagerung zu beobachten.

Die Schwierigkeit für einen Algorithmus zur Materialbestimmung wird anhand der Abbildungen 5.10 und 5.11 deutlich. Wird lediglich ein Voxel in einem DECT-Schichtbild betrachtet, sind zwei Zahlenwerte gegeben – der $\mu_{L,m}(140 \text{ kV})$ und der $\mu_{L,m}(240 \text{ kV})$. Die DECT-Punktewolken in den Abbildungen 5.10 und 5.11 sind weitestgehend getrennt, jedoch gibt es einzelne Ausreißer, die zu einer leichten Überlagerung der Punktewolken beitragen. Mit einem klassischen, (einzel-)voxelbasierten Algorithmus, nach der Definition klassischer Algorithmen zur Materialbestimmung aus Kapitel 4, kann in diesen Fällen nicht entschieden werden, welches Material vorliegt. Nach der Einführung des Deep Learnings (Kapitel 3) und den Ergebnissen der Literaturrecherche zur Material-CT (Kapitel 4) besteht die These, dass DL-basierte Verfahren mit ihrer Fähigkeit, einen lokalen (Bild-)Kontext in eine Berechnung einzubeziehen, diese Limitierung auflösen können.

⁶ Negative Abschwächungskoeffizienten sind physikalisch nicht sinnvoll, entstehen immer als Folge von CT-Artefakten und werden in der Regel genutzt.

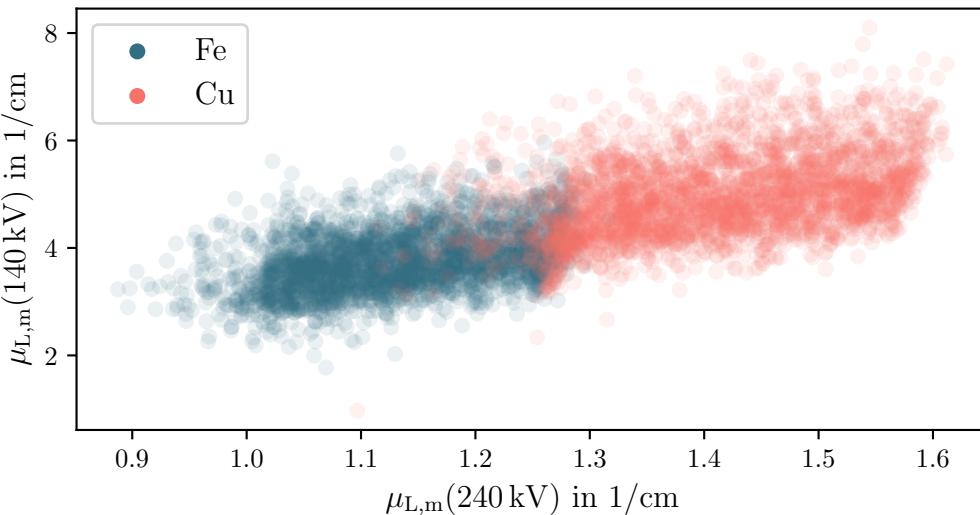


Abbildung 5.11: Aus einer Simulation extrahierte, materialabhängige Abschwächungskoeffizienten $\mu_{L,m}$ für zwei Quellspannungen 140 kV und 240 kV. Es wurden jeweils 3000 Voxel zufällig ausgewählt. Die einzelnen Datenpunkte sind leicht transparent eingezeichnet, um die Dichte übereinander liegender Punkte zu visualisieren.

5.4.4 Geschwindigkeit und Skalierbarkeit

Die am Anfang des Kapitels in Abschnitt 5.1 formulierten Anforderungen bezüglich der Geschwindigkeit der Trainingsdatenerzeugung werden nun für die implementierte Methode evaluiert. Die implementierte Simulation berechnet einen Trainingsdatensatz gemäß Tabelle 5.2 mit insgesamt 250.000 Tupeln in weniger als zwei Stunden. Für die Bestimmung der angegebenen Zeit wurde ein GPU-Server der Partition *WestAI* am Forschungszentrum Jülich verwendet. Jeder GPU-Server dieser Partition besitzt 64 CPU-Kerne, 512 GB Speicher und vier Nvidia H100 NVL mit jeweils 94 GB Graphikspeicher. Die einzelnen Simulationen werden in Batches mit jeweils 64 Datentupeln innerhalb der oben angesprochenen Dask-Worker gestartet. Abbildung 5.12 zeigt den Einfluss der Worker-Anzahl pro GPU auf die Geschwindigkeit der Simulation. Es gilt, genügend

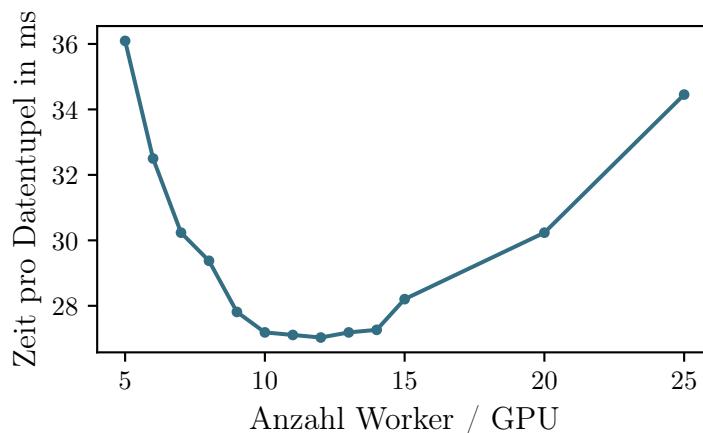


Abbildung 5.12: Mittlere benötigte Zeit pro Simulation eines Datentupels für verschiedene Worker-Anzahlen pro GPU.

Worker zu starten, sodass die GPUs voll ausgelastet werden. Bei einer zu großen Worker-Anzahl blockieren sich die Prozesse gegenseitig und bremsen das gesamte System aus. Die minimale Zeit pro Datensatz wird bei 12 Workern beobachtet, sodass im Folgenden 12 Worker pro GPU verwendet werden.

Weiterhin ist zu klären, wie die benötigte Rechenzeit von der Anzahl der verwendeten Graphikkarten innerhalb eines Rechenknotens abhängt. Abbildung 5.13 zeigt die benötigte Zeit zur Berechnung eines verkleinerten Testdatensatzes bestehend aus 12800 Datentupeln. Die theoretische, ideale Skalierung bezüglich der GPU-Anzahl (graue Linie)

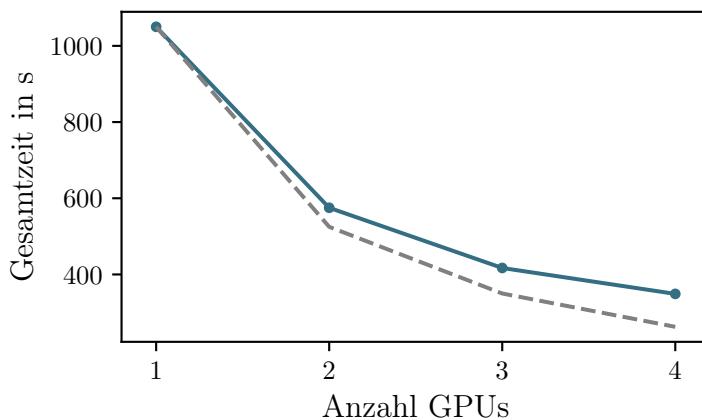


Abbildung 5.13: Mittlere benötigte Zeit für die Simulation von 12800 Datentupeln mit einer oder mehreren GPUs. Die graue Linie markiert den Idealfall: die Verdopplung der GPU-Anzahl führt zu einer Halbierung der Rechenzeit.

wird erwartungsgemäß nicht erreicht. Dennoch ist ein Leistungszuwachs, das heißt eine Reduktion der Berechnungsdauer, für eine zunehmende Anzahl an GPUs zu beobachten. Aufbauend kann untersucht werden, wie sich die Skalierung mit einer größeren Anzahl an GPUs, verteilt auf mehrere Rechenknoten, in Bezug zur Rechenzeit verhält. Es ist anzunehmen, dass die in Abbildung 5.13 dokumentierte Leistungsabnahme durch Erhöhung der GPU-Anzahl weiter zunimmt. Im Rahmen dieser Thesis wird auf diese Untersuchung bewusst verzichtet, da die Simulation auf einem Rechenknoten mit der oben genannten Spezifikation innerhalb des Zeitlimits von 24 h etwa drei Millionen Datentupel erzeugen und speichern kann.

5.5 ZUSAMMENFASSUNG

In diesem Kapitel wurden zunächst die CT-spezifischen sowie die DL-spezifischen Anforderungen für die Trainingsdatenerzeugung mithilfe einer Simulation für die Materialbestimmung in der Computertomographie identifiziert. Motiviert anhand eines realen CT-Messablaufes wurde ein Ablaufdiagramm für die Trainingsdatenerzeugung in Abbildung 5.1 entwickelt. Einige der Funktionalitäten, der Schnittlängenprojektor sowie die Rekonstruktion, konnten in bestehenden Simulationsprogrammen identifiziert und isoliert für die eigene Implementierung genutzt werden. Die fehlenden Funktionen, das Generieren zufälliger Phantome sowie die Abschwächungsberechnung nach Gleichung

1.1, wurden implementiert. Insgesamt wurde die entwickelte Methode zur Datenerzeugung abschließend anhand der zu Beginn des Kapitels aufgestellten Anforderungen evaluiert. Es konnte gezeigt werden, dass die implementierte Simulation eine ausreichende Übereinstimmung mit der vielerseits genutzten Simulation *aRTist* aufweist. Der Abgleich mit realen CT-Scandaten zeigt jedoch deutliche Abweichungen, die vermutlich aufgrund systematischer Modellierungsungenauigkeiten auftreten, sodass die geforderte Genauigkeit nicht erreicht wird. Diese Abweichungen können vermutlich mit der Verwendung einer Monte-Carlo-basierten Simulation signifikant reduziert werden, wozu jedoch die verfügbare Rechenzeit um mehrere Größenordnungen überschritten wird. Die implementierte Simulation wird im weiteren Verlauf dieser Arbeit zur Trainingsdatenerzeugung verwendet, wobei der Fehler aufgrund der angesprochenen Sim-To-Real-Gap jeweils im Kontext der Auswertung verschiedener Aspekte diskutiert wird. Es wird festgehalten, dass mithilfe der implementierten Methode zur Trainingsdatenerzeugung die im Rahmen der Literaturrecherche identifizierte Forschungslücke mindestens teilweise geschlossen werden kann. Durch die Geschwindigkeit und Skalierbarkeit der implementierten Methode können erstmals sehr große Trainingsdatenmengen mit mehreren Millionen Tupeln innerhalb einer wirtschaftlichen Zeitspanne erzeugt werden. Mithilfe ebendieser Datensätze können parameterstarke DL-Modelle, die auf große Datensätze angewiesen sind, trainiert und für die Material-CT erprobt werden. Auch alternative Fragestellungen der CT abseits der Materialbestimmung können mithilfe der implementierten Methodik durch die Simulation entsprechender Trainingsdaten erschlossen werden. Offen bleibt eine Lösung der beobachteten Sim-To-Real-Gap, die als Folge der fehlenden Modellierung der Streustrahlung auftritt und weiterhin nicht trivial mithilfe einer Raytracing-basierten Simulation erschlossen werden kann.

6

DEEP-LEARNING-ARCHITEKTUR FÜR DIE MATERIALBESTIMMUNG

Nach der Einführung des Deep Learnings als flexibles Werkzeug zur Darstellung komplexer Funktionen in Kapitel 3 wird in diesem Kapitel zur Beantwortung der zweiten Forschungsfrage eine DL-Architektur gesucht, die die Materialbestimmung basierend auf CT-Daten effektiv lösen kann. Zunächst werden in Abschnitt 6.1 die industriellen Rahmenbedingungen beschrieben, die für den späteren Transfer der DL-Modelle in die praktische CT-Anwendung berücksichtigt werden müssen. Es werden verschiedene Szenarien von CT-Scans vorgestellt, die individuelle Anforderungen an die Inferenzgeschwindigkeit der DL-Modelle stellen. Weiterhin ist die zu erreichende, absolute Präzision der DL-basierten Materialbestimmung festzulegen, die ebenfalls im Rahmen der industriellen Rahmenbedingungen anwendungsbezogen motiviert und berechnet wird. Die in Kapitel 5 beschriebene Simulation wird zur Erzeugung der Trainingsdaten für die DL-Modelle in diesem Kapitel verwendet. Somit unterliegt die Beurteilung der Präzision bei der Materialbestimmung anhand von realen CT-Scandaten weiterhin der Sim-To-Real-Gap, die nach Kapitel 5.4.2 nicht vernachlässigbar ist. Daher werden zunächst simulierte Phantome verwendet, um die Präzision der Materialbestimmung in diesem Kapitel, insbesondere beim Vergleich verschiedener DL-Architekturen, zu quantifizieren. Das hierzu simulierte Phantom wird in Abschnitt 6.2 anhand der Rahmenbedingungen motiviert und beschrieben. CT-Schichtbilder weisen in praktischen Anwendungen meist große räumliche (Bild-)Dimensionen H, W auf (vergleiche Kapitel 6.1). Ein Standardansatz, um sehr große Bilder mit neuronalen Netzen effizient zu verarbeiten, ist das sogenannte *Cropping* – das Zerlegen eines Bildes in kleinere Teilbilder, sodass die Teilbilder einzeln von einem neuronalen Netz verarbeitet werden können [107, 108]. In Abschnitt 6.3 wird untersucht, ob diese Strategie für die Materialbestimmung mit CT-Schichtbildern angewandt werden kann. Basierend auf der Literaturrecherche in Kapitel 4.2.4 werden in Abschnitt 6.4 zwei DL-Architekturen für die Materialbestimmung trainiert und quantitativ verglichen. Abschließend wird die ausgewählte Architektur in Abschnitt 6.5 anhand der zuvor beschriebenen Rahmenbedingungen bewertet.

6.1 INDUSTRIELLE RAHMENBEDINGUNGEN FÜR DL IN DER CT

Zunächst ist die Laufzeit, die sogenannte *Inferenzzeit*, der entwickelten DL-Architektur im Kontext der Mess- und Rekonstruktionszeit eines CT-Scans zu beurteilen. In der industriellen CT lassen sich drei Scanmodi anhand der zur Verfügung stehenden Messzeit identifizieren: die *Labor-CT*, die *At-line-CT* und die *In-line-CT*. Die Labor-CT ist auf eine bestmögliche Bildqualität ausgelegt, sodass die Messzeit praktisch nicht limitiert ist. Im Gegensatz dazu steht die In-line-CT, die zur lückenlosen Prüfung in einem Produkti-

onsprozess eingesetzt wird und deren Messzeit direkt an den Produktionstakt T der zu prüfenden Teile (sogenannte *Prüfteile*) gekoppelt ist. Durch die Einschränkung der Messzeit kann die Bildqualität negativ beeinträchtigt werden. Zwischen diesen beiden Prüfkonzepten gibt es die At-line-CT, die produktionsbegleitend nicht jedes produzierte Teil, sondern nur stichprobenhaft die CT-Prüfung durchführt, sodass ein Kompromiss aus Bildqualität und aufgebrachter Messzeit erreicht wird. Grundsätzlich lässt sich für die Labor-CT keine eindeutige obere Schranke hinsichtlich der Inferenzzeit eines DL-Modells bestimmen, da die Messzeiten anwendungsabhängig sind und ebenso die Dimensionen der rekonstruierten Volumina stark variieren. Hingegen wird durch den Produktionstakt bei der At-line- und In-line-CT die maximale Verarbeitungszeit eingeschränkt. Anhand eines Beispiels aus der industriellen Fließbandproduktion lassen sich einige Faktoren aus der Anwendungsumgebung anschaulich erfassen. Abbildung 6.1 zeigt den Ausschnitt einer industriellen Fertigungsline mit einem CT-Scanner zur Qualitätsüberwachung. Während das Prüfteil nach der CT-Bildaufnahme im Fertigungsprozess weiterverarbeitet

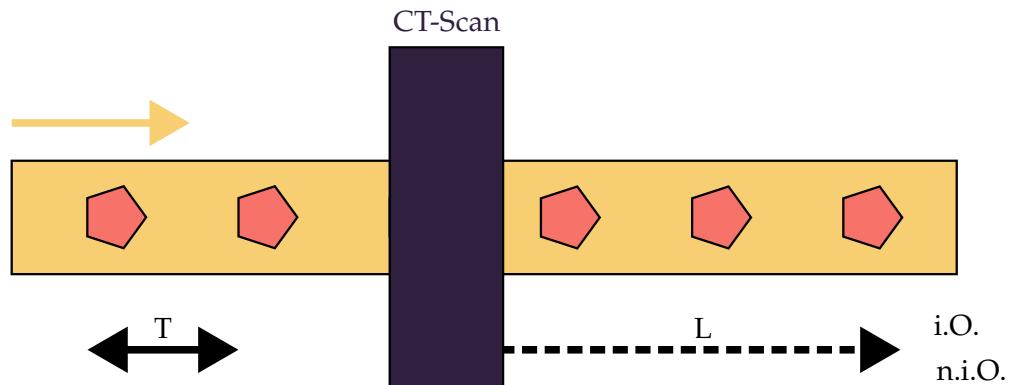


Abbildung 6.1: Schematische Darstellung einer In-line-CT gekoppelt an ein Fließband mit einem zeitlichen Abstand T , der *Taktzeit*, zwischen zwei Prüfteilen (oder: *Proben*). Nach der Bildaufnahme im CT-Scanner werden die Prüfteile auf dem Fließband weitertransportiert und die Zeit, beziehungsweise die zurückgelegte Strecke, bis zur Prüfentscheidung *in Ordnung* (i.O.) oder *nicht in Ordnung* (n.i.O.) durch den CT-Scan wird als *Latenz* L bezeichnet.

wird, laufen im Hintergrund die Rekonstruktion und die (DL-basierte) Auswertung. Die Laufzeit der DL-Modelle hat somit, additiv zur Laufzeit der Rekonstruktion, einen direkten Einfluss auf die Latenz der Prüfentscheidung. Im Idealfall erfolgen die Rekonstruktion und die Auswertung innerhalb der Taktzeit T , sodass die Latenz kleiner oder gleich der Taktzeit ist und die gesamte Auswertung somit in Echtzeit zum Prüfprozess erfolgt. Dreier et al. demonstrieren, dass einzelne Prüfaufgaben an Batteriezellen mithilfe von CT-Scans innerhalb einer Sekunde durchgeführt werden können [109]. Aus Gesprächen mit CT-Experten wird zudem klar, dass die Durchführung eines CT-Scans einer ganzen Batteriezelle (Rundzelle, beispielsweise Typ 18650) innerhalb von etwa 10 s realistisch auf dem von Dreier et al. beschriebenen CT-Scanner¹ durchgeführt werden kann. Die Rekonstruktion der CT-Daten wird im Rahmen dieser Arbeit nicht näher betrachtet, sodass idealisiert angenommen wird, dass die Rekonstruktion innerhalb der geforderten Taktzeit von 10 s auf einem dedizierten Rechner erfolgt. Somit ergibt sich

¹ Konkret wird die kurze Messzeit durch den hohen Photonenfluss der von Dreier et al. verwendeten Röntgenquelle [109] erreicht.

für die Materialbestimmung, analog zur Rekonstruktionszeit, die maximale Inferenzzeit 10 s, um dem Takt der CT-Scans zu folgen. Die Materialbestimmung mit Deep Learning wird in der Regel (vergleiche Kapitel 4.2.4) schichtweise durchgeführt, sodass die Berechnungen auf unabhängigen CT-Schichten, senkrecht zur Drehachse, auf mehreren GPUs parallel ausgeführt werden können, wobei jede GPU die Zeit $T = 10$ s ausschöpfen kann. Im weiteren Verlauf des Kapitels wird sich zeigen, welche Architekturen nach dieser Definition *In-line-fähig* sind und insbesondere wirtschaftlich sinnvoll, das heißt unter Verwendung eines einzelnen Inferenzservers mit bis zu acht Graphikkarten, in realen Produktionsumgebungen genutzt werden können.

Eine weitere Rahmenbedingung für den industriellen Einsatz der entwickelten Verfahren zur Materialbestimmung besteht in der absoluten Präzision sowie der Bildqualität. Hierbei meint der Begriff *Präzision* die Vorhersagegenauigkeit der Methode in Bezug auf die ortsabhängige Materialzusammensetzung (Kernladungszahl Z und Dichte ρ) und der Begriff *Bildqualität* die klassische Bildqualität in Bezug auf Rauschen und Bildschärfe. Die Anforderungen an die Bildqualität werden direkt aus dem CT-Scan abgeleitet. Wie in Kapitel 2.3 beschrieben, wird das Rauschen der aufgenommenen Projektionen durch die Rekonstruktion in das Volumen zurückprojiziert. Steuwe et al. zeigen, dass die DL-basierte Bildverarbeitung das Rauschen in CT-Schichtbildern reduzieren kann [110]. Daher soll anhand der Ergebnisse dieser Arbeit die Rauschreduktion der verwendeten Methode stichprobenartig untersucht werden. Hierzu wird das Bildrauschen als effektive Schwankung der rekonstruierten Abschwächungskoeffizienten quantifiziert. Die Präzision der Methode muss im Kontext einer gegebenen Messaufgabe beurteilt werden. Mit dem Wissen, dass eine vorliegende Probe weder ein Gemisch noch eine Legierung ist, genügt zur Materialbestimmung von beispielsweise Aluminium die Vorhersage der Kernladungszahl $Z = 13 \pm 0.49$. Anders verhält es sich bei Legierungen, wie zum Beispiel Edelstählen, da hier die effektiven Kernladungszahlen verschiedener Edelstähle reellwertig sind [14] und sich gegebenenfalls erst in den Nachkommastellen unterscheiden. Gleichermaßen gilt für die Vorhersage der Dichte, sodass sich für beide Größen keine direkte Anforderung bezüglich der Präzision aus der Anwendungsperspektive ergibt – je genauer, desto besser. In Gesprächen mit CT-Anwendern wird häufig eine maximale Abweichung von 1 % (Abweichung vom jeweiligen Literaturwert) für die Materialbestimmung, das heißt konkret die Kernladungszahl und die Dichte, gefordert. Daraus ergibt sich beispielsweise für die Dichtevorhersage von Eisen mit dem Literaturwert ein Toleranzbereich $\rho \approx (7.80 \pm 0.08) \text{ g/cm}^3$. Zur Vereinfachung, das heißt unter Ausschluss von Legierungen, werden somit die Toleranzbereiche ± 0.5 für die Vorhersage der Kernladungszahl und $\pm 1\%$ für die Vorhersage der Dichte definiert. Innerhalb dieser Toleranzbereiche können sogenannte *Reinstoffe*² eindeutig identifiziert und anhand ihrer Dichte auf Porosität geprüft werden. Darauf aufbauend kann untersucht werden, welche maximale Präzision für die Vorhersage der Kernladungszahl und Dichte aus physikalischer Sicht überhaupt möglich ist. Die rekonstruierten DECT-Schichtbilder dienen als Eingabedaten für das vorgestellte Verfahren zur Materialbestimmung. Sie unterliegen einem Rauschen, welches die maximale Präzision der Materialbestimmung limitiert und zunächst bestimmt werden

² Gemeint ist hier ein Material, das nur aus einer Atomsorte besteht.

muss. Das absolute Rauschen entspricht der Standardabweichung der rekonstruierten Abschwächungskoeffizienten in einem homogenen Objekt. Problematisch ist jedoch, dass die rekonstruierten Abschwächungskoeffizienten, aufgrund der Strahlauflaufartefakte innerhalb einer stark abschwächenden Probe, positionsabhängig abgesenkt werden und somit die genannte Standardabweichung in einem Bildausschnitt signifikant beeinflussen. Um das absolute Rauschen von den Strahlauflaufartefakten (vergleiche Kapitel 2.4.1) innerhalb der Probe zu trennen, wird eine zweidimensionale Funktion an den gewählten Bildausschnitt $\mu^{(x,y)}$ im CT-Schichtbild angefittet, die näherungsweise den Einfluss der Strahlauflaufartefakte abbildet. Abbildung 6.2 zeigt einen eindimensionalen Schnitt durch den Kupferwürfel aus Abbildung 5.8 (realer Scan) sowie die angefittete Funktion. Die Standardabweichung der Differenz zwischen Signal und Fit

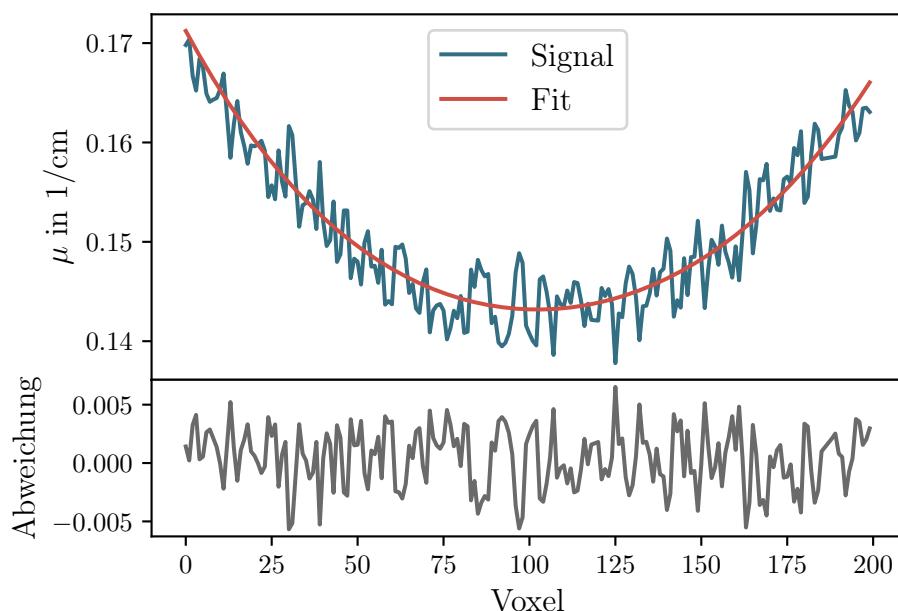


Abbildung 6.2: Schnitt durch den mittleren Teil des Kupferwürfels aus Abbildung 5.8 (Signal) mit dem Fit eines Polynoms vierten Grades (oben). Die absolute Abweichung zwischen dem Signal und dem Fit wird unten dargestellt und entspricht dem absoluten Rauschen in der Einheit $1/\text{cm}$. Der genaue Bildausschnitt innerhalb der Kupferprobe ist in Abbildung 6.3 eingezeichnet.

(Abbildung 6.2; unten) entspricht dem absoluten Bildrauschen in dem ausgewählten Bildausschnitt. Abbildung 6.3 stellt die DECT-Schichtbilder aus dem realen CT-Scan gemeinsam mit den ausgewählten Bildausschnitten zur Bestimmung des Rauschens dar. In allen Materialien ist das Rauschen im Low-Energy-Kanal (140 kV) stärker. Weiterhin ist das Rauschen in den Abschwächungskoeffizienten von Eisen und Kupfer deutlich höher als in den Abschwächungskoeffizienten des Aluminiums, da das Aluminium aufgrund der niedrigeren Abschwächung weniger Photonen absorbiert und somit die Zählstatistik am Detektor besser ausfällt. Zur Bestimmung des Einflusses des Bildrauschens $\Delta\mu_L$ auf

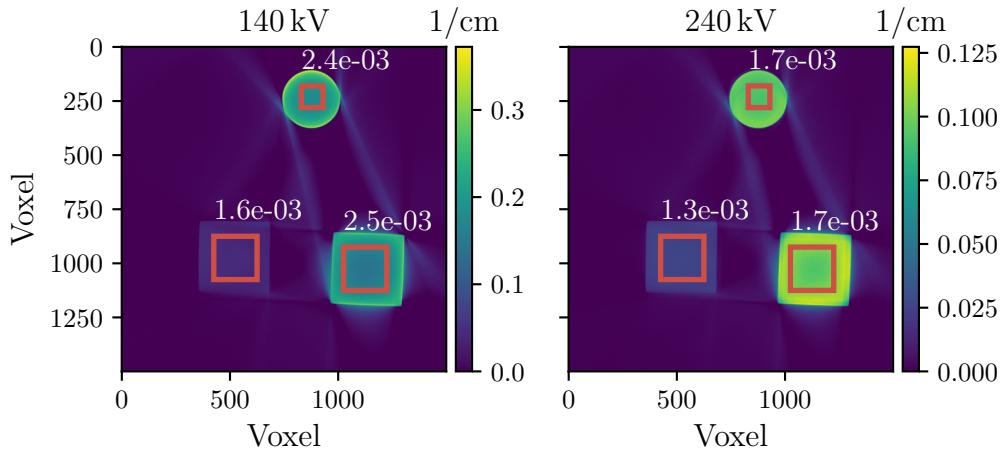


Abbildung 6.3: Rekonstruierte DECT-Volumina eines realen Scans mit berechnetem Rauschen in den verschiedenen Materialien. Die roten Boxen markieren die jeweilige Bildausschnitte zur Berechnung des Rauschens.

die vorhergesagte Kernladungszahl Z und Dichte ρ werden die Gleichungen nach Knoll [3, S. 54; G. 2.22] und Heismann [14] verwendet:

$$\begin{aligned} \mu_L &= \rho \cdot \mu_m && \text{(Knoll)} \\ \mu_L &= \alpha \rho \frac{Z^k}{E^l} + \beta \rho && \text{(Heismann)} \end{aligned} \quad (6.1)$$

wobei μ_m der Massenschwächungskoeffizient eines Materials m ist, α , β , k und l Konstanten nach Heismann [14] sind und E die effektive Photonenenergie (vergleiche Kapitel 2.6; Sprawls [34]) ist. Durch Umformen und Ableiten ergibt sich für die Fehlerfortpflanzung mit einem gegebenen Rauschen des rekonstruierten Volumens $\Delta\mu_L$

$$\begin{aligned} \mu_L &= \alpha \rho \frac{Z^k}{E^l} + \beta \rho \Leftrightarrow Z = \left(\frac{(\mu_L - \beta \rho) E^l}{\alpha \rho} \right)^{\frac{1}{k}} = \left(\frac{E^l}{\alpha \rho} \right)^{\frac{1}{k}} \cdot (\mu_L - \beta \rho)^{\frac{1}{k}} \\ \Rightarrow \frac{dZ}{d\mu_L} &= \left(\frac{E^l}{\alpha \rho} \right)^{\frac{1}{k}} \cdot \frac{1}{k} \cdot (\mu_L - \beta \rho)^{\frac{1}{k}-1} \\ \Rightarrow \Delta Z &= \left(\frac{E^l}{\alpha \rho} \right)^{\frac{1}{k}} \cdot \frac{1}{k} \cdot (\mu_L - \beta \rho)^{\frac{1}{k}-1} \cdot \Delta \mu_L \end{aligned} \quad (6.2)$$

für die Ungenauigkeit der Kernladungszahl ΔZ sowie

$$\begin{aligned}\mu_L = \rho \cdot \mu_m &\Leftrightarrow \rho = \frac{\mu_L}{\mu_m} \Rightarrow \frac{d\rho}{d\mu_L} = \frac{1}{\mu_m} \\ &\Rightarrow \Delta\rho = \frac{1}{\mu_m} \cdot \Delta\mu_L\end{aligned}\quad (6.3)$$

für die Ungenauigkeit der Dichte $\Delta\rho$.

Aus den berechneten Werten des Bildrauschens in Abbildung 6.3 und den Gleichungen 6.2 und 6.3 werden die bestmöglichen Präzisionen für die Materialbestimmung abhängig vom Material ermittelt und in Tabelle 6.1 aufgelistet. Da die Materialbestimmung auf beiden Energien des DECT-Scans basiert, wird das stärkere Rauschen (Low-Energy-Bild aus Abbildung 6.3 (links)) zur Abschätzung verwendet.

Tabelle 6.1: Theoretische Grenzauflösungen der Kernladungszahlvorhersage $Z \pm \Delta Z$ und Dichtevorhersage $\rho \pm \Delta\rho$ berechnet mithilfe von Gleichung 6.2 und 6.3 anhand der Daten des realen Scans aus Abbildung 5.8 und 6.3.

Material	$Z \pm \Delta Z$	$(\rho \pm \Delta\rho)$ in g/cm ³
Al ($Z = 13, \rho = 2.70$)	13 ± 0.033	2.70 ± 0.0096
Fe ($Z = 26, \rho = 7.80$)	26 ± 0.023	7.80 ± 0.0092
Cu ($Z = 29, \rho = 8.93$)	29 ± 0.021	8.93 ± 0.0083

6.2 SIMULIERTES PHANTOM ZUR QUANTITATIVEN AUSWERTUNG

Für einen quantitativen Vergleich zwischen den im weiteren Verlauf dieses Kapitels entwickelten DL-Architekturen werden Testdaten benötigt, deren Eigenschaften exakt bekannt sind. Die Phantome der Testdaten müssen verschiedene Materialien enthalten, deren Zusammensetzung und Position innerhalb der rekonstruierten Volumina exakt bekannt sind. Zur Abdeckung der in Kapitel 5.3 motivierten Bandbreite verwendeter Materialien, werden die Materialien Aluminium, Eisen und Kupfer exemplarisch ausgewählt. Aluminium zählt zu den leichtesten und Kupfer zu den schwersten ausgewählten Elementen in Tabelle 5.1 während Eisen mittig platziert ist. Zudem müssen die CT-Geometrie und das Quellspektrum exakt bekannt sein, um Abweichungen der DL-Modelle aufgrund von systematischen Modellierungsfehlern ausschließen zu können. Rückblickend auf die beobachtete Sim-To-Real-Gap in Kapitel 5.4.2 sollte daher eine Simulation zur Erzeugung des Phantoms verwendet werden, die den CT-Messprozess analog zu der implementierten Simulation modelliert. Somit kann beispielsweise der Einfluss der Streustrahlung auf die Ergebnisse der Materialbestimmung explizit ausgeklammert werden. Dennoch sollten die Testdaten in einem sinnvollen und realistischen Maß CT-Artefakte, konkret die Strahlauflährtung und das statistische (Bild-)Rauschen, enthalten. Die Stärke der Strahlauflährtung kann durch die Probengröße (genauer: die Schnittlänge durch ein Material) sowie durch das Quellspektrum beeinflusst werden. Basierend auf Experimenten an der realen CT-Anlage aus Kapitel 2.7 werden die DECT-Quellspektren mit den Beschleunigungsspannungen 140 kV (Low-Energy-Kanal) und 240 kV (High-Energy-Kanal) festgelegt. Während die höhere Beschleunigungsspannung

mit 240 kV die maximal einstellbare Beschleunigungsspannung an der Röntgenquelle der realen CT-Anlage ist, wird die niedrigere Beschleunigungsspannung so gewählt, dass sie einerseits einen möglichst großen Abstand zur höheren Beschleunigungsspannung aufweist und andererseits nicht so niedrig ist, dass die Strahlauflaufhärtingsartefakte die Bildqualität maßgeblich beeinträchtigen. Die Probengröße wird auf 1 cm (Kantenlänge) festgelegt, um deutliche Strahlauflaufhärtingsartefakte im Low-Energy-Kanal zu induzieren während der High-Energy-Kanal weitestgehend artefaktfrei bleibt. Zur Simulation des Testdatensatzes wird *aRTist* (vergleiche Kapitel 5.2.1) verwendet. Das simulierte Phantom wird in Abbildung 6.4 dargestellt. Die in der Simulation verwendeten CT-Scanparameter,

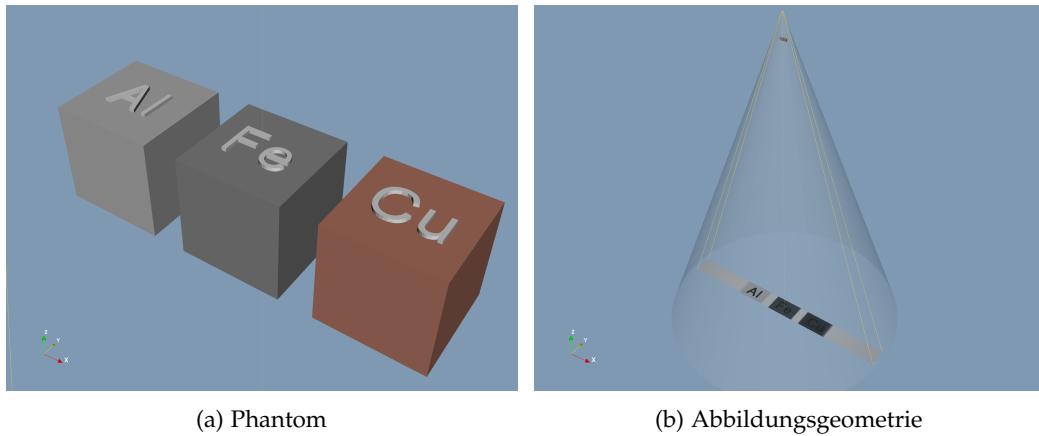


Abbildung 6.4: Screenshots (a) des Phantoms bestehend aus Aluminium (Al), Eisen (Fe) und Kupfer (Cu) sowie der gesamten Abbildungsgeometrie in (b) aus der Simulation *aRTist*. In (b) befindet sich die Röntgenquelle mit eingezeichnetem Strahlkegel oben im Bild. Direkt unterhalb ist das Phantom aus (a) zu sehen. Unten befindet sich ein Detektor, der die Projektion des Phantoms zeigt.

abgeleitet aus den Anforderungen, werden in Tabelle 5.2 aufgelistet. Abbildung 6.5 zeigt einen Schnitt durch das rekonstruierte Volumen des Phantoms aus Abbildung 6.4. Der Effekt der Strahlauflaufhärting kann anhand des deutlichen Cuppings (vergleiche Abschnitt 2.4.1), wie gefordert, in der oberen Zeile im Low-Energy-Kanal erkannt werden. Für die nachfolgenden Auswertungen dieses Kapitels wird ausschließlich das vorgestellte Phantom aus Abbildung 6.5 verwendet.

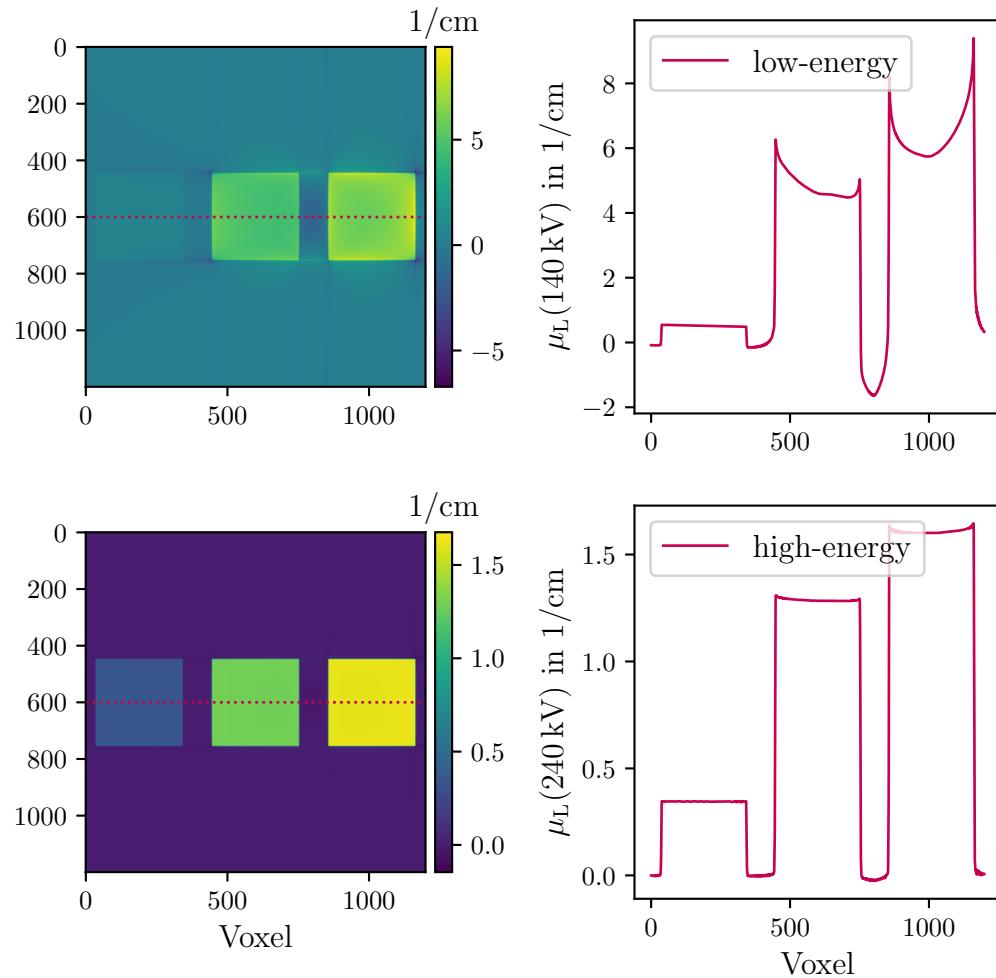


Abbildung 6.5: CT-Schnitbilder durch den Low-Energy-Kanal (oben) und den High-Energy-Kanal (unten) des Phantoms aus Abbildung 6.4. Die rechte Spalte stellt jeweils den Schnitt entlang der gestrichelten Mittellinien dar.

6.3 EINFLUSS DES GLOBALEN BILDKONTEXTES

CT-Artefakte beeinflussen im Allgemeinen die Bildqualität und insbesondere die quantitativ gemessenen Abschwächungskoeffizienten, wie in Abschnitt 2.4.1 beschrieben worden ist. Die Identifikation eines CT-Artefakts, in einem gegebenen CT-Volumen, ist für Domänenexperten in der Regel unproblematisch, da sie, im globalen Bildkontext, das gescannte Objekt von den charakteristischen Streifen, Linien und Grauwertverschiebungen unterscheiden können. Wird hingegen nur ein kleiner Ausschnitt des CT-Volumens betrachtet, ist diese Unterscheidung zunehmend anspruchsvoller und ab dem Unterschreiten einer gewissen Bildausschnittgröße unmöglich. DL-basierte Verfahren zur Verarbeitung von CT-Daten, oder allgemein Bilddaten, bieten gegenüber den meisten klassischen Algorithmen den Vorteil, dass sie dieses entscheidende Kontextwissen extrahieren können. Im Idealfall wird somit ein DL-Modell auf vollständigen Schnittbildern aus dem CT-Volumen trainiert, sodass der globale Kontext ebendieser Schicht vollständig verfügbar ist. Ein Problem mit diesem Ansatz ist, dass industrielle CT-Schnittbilder häufig eine hohe Auflösung bis zu 3000^2 Voxeln aufweisen, sodass das direkte Training auf einzelnen Schichtbildern dieser Auflösung aufgrund des immensen Rechenbedarfs nicht realistisch durchführbar ist. In Kapitel 5.1.2, den DL-induzierten Rahmenbedingungen, wurden daher die Bilddimensionen der simulierten CT-Schichtbilder vorläufig auf $(H, W) = (256, 256)$ festgelegt. Dennoch existieren Strategien, um mit großen Bildern zu trainieren und die Modelle auf ebendiesen auszuführen: Bei der sogenannten *Crop-Strategie* werden das Training und die Inferenz auf kleineren Bildausschnitten, sogenannten *Crops*, ausgeführt. Alternativ können die großen Bilder auch mittels Interpolation verkleinert werden (*Resize-Strategie*), sodass das gesamte CT-Schichtbild mit einer reduzierten Auflösung vorliegt. Beide Strategien werden im Folgenden erprobt und hinsichtlich ihrer Vorhersagegenauigkeit bei der Materialbestimmung verglichen. Basierend auf den Ergebnissen der Literaturrecherche in Kapitel 4.2.4 wird, der Mehrheit der Autoren folgend, die U-Net-Architektur nach Ronneberger et al. [36] verwendet. Da es sich bei der folgenden Untersuchung der Resize- und Crop-Strategie im Kern um die Bedeutung des globalen Bildkontextes für die lokale Materialbestimmung handelt, kann angenommen werden, dass sich andere Architekturen in diesem Punkt analog zum U-Net verhalten.

Um einen fairen Vergleich zu ermöglichen, wird ein Datensatz mit den Dimensionen $(C, H, W) = (2, 1024, 1024)$ der DECT-Schichtbilder simuliert. Da die Gesamtgröße des Datensatzes aufgrund der hohen räumlichen Auflösung signifikant ansteigt (vergleiche Kapitel 5.1), wird die Anzahl der simulierten Tupel auf 32.000 reduziert, sodass der Datensatz insgesamt etwa 260 GB umfasst. Für das Modell, das mit der Resize-Strategie trainiert wird ($\mathcal{M}^{\text{resize}}$), werden die Bilder mittels bilinearer Interpolation auf die Dimensionen $(C, H, W) = (2, 256, 256)$ verkleinert. Das andere Modell ($\mathcal{M}^{\text{crop}}$) wird mit, während des Trainings zufällig erzeugten, Crops mit den Dimensionen $(C, H, W) = (2, 256, 256)$ aus den DECT-Schichtbildern trainiert. Abbildung 6.6a zeigt den Training Loss und den Validation Loss des Models $\mathcal{M}^{\text{resize}}$, das auf vollständigen DECT-Schichtbildern trainiert worden ist. Nach etwa 150 Epochen erfährt das Modell Overfitting an die Trainingsdaten. Die Loss Curves des Modells $\mathcal{M}^{\text{crop}}$ sind in Abbildung 6.6b dargestellt. Anhand von

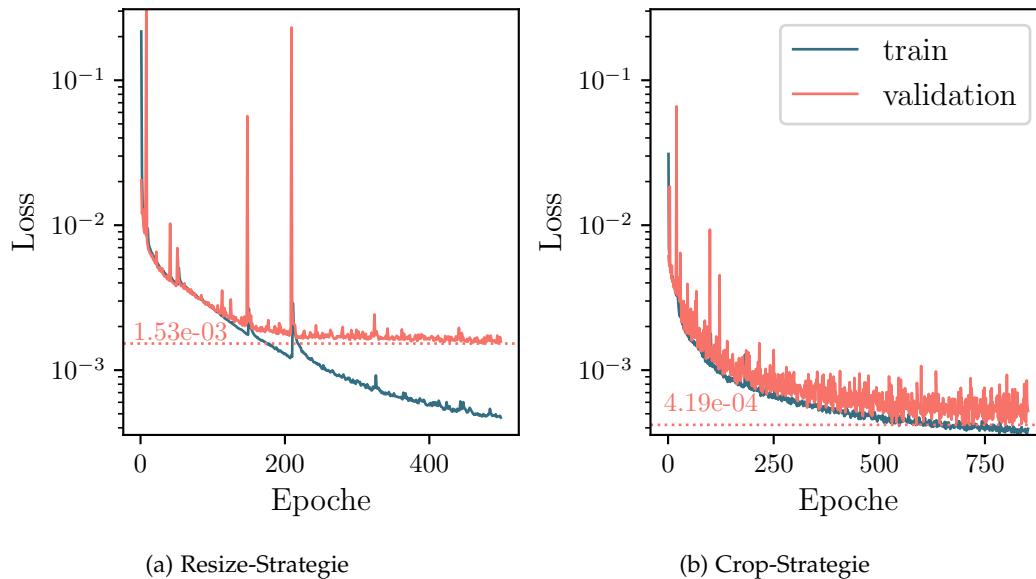


Abbildung 6.6: Training (*train*) und Validation Losses (*validation*) für (a) das Training auf ganzen CT-Schichtbildern mit der Resize-Strategie und (b) das Training auf zufälligen Ausschnitten der CT-Schichtbilder mit der Crop-Strategie. Die Originalauflösung der Schichtbilder ist (1024×1024) Pixel. Die Schichtbilder werden für das Training in (a) mittels bilinearer Interpolation verkleinert, sodass der gesamte Bildkontext vorliegt. Das Training in (b) wird mit Random Crops der DECT-Schichtbilder durchgeführt, sodass nur ein eingeschränkter, lokaler Bildkontext vorliegt. Die jeweils besten Validation Losses werden durch die gestrichelten Linien markiert. Beide Plots entsprechen der Legende in (b). Die Trainingsparameter sind in beiden Durchläufen bis auf die Anzahl der Epochen identisch.

Abbildung 6.6 scheint das Training mit der Crop-Strategie bessere Ergebnisse, das heißt konkret einen niedrigeren Validation Loss, zu liefern. Ein Grund für den niedrigeren Validation Loss kann in der relativen Größe der Regionen (vergleiche Abbildung 5.3) in Bezug zur Bildgröße, beziehungsweise im Fall der Crop-Bilder *Bildausschnittsgröße*, vermutet werden. Es besteht die Vermutung, dass größere, zusammenhängende Materialregionen, einfacher vom Modell vorhergesagt werden können. Dieser Zusammenhang zwischen der mittleren Regionsgröße relativ zur Bildgröße und der Materialvorhersage wird in Abschnitt 7.1 durch eine Variation der mittleren Regionsgröße spezifisch untersucht. Die quantitative Auswertung des Phantoms in Tabelle 6.2 bringt im Gegensatz zu Abbildung 6.6 hervor, dass die Bestimmung der Materialien in den meisten Fällen mit der Resize-Strategie bessere Ergebnisse liefert. Lediglich die Vorhersage der Kernladungszahl von Kupfer erzielt einen besseren Mittelwert mit der Crop-Strategie, wobei die berechnete Standardabweichung verhältnismäßig hoch ausfällt. Die Ergebnisse zeigen, dass die Resize-Strategie zur Optimierung der Materialbestimmungspräzision im Folgenden verwendet werden sollte. Offen bleibt die Frage nach einer sinnvollen Bildgröße infolge des Resizings. Wie bereits in Kapitel 5.1 beschrieben, hat die Bildgröße einen signifikanten Einfluss auf die Größe des gesamten Trainingsdatensatzes sowie die Dauer des Trainings. Vor dem Hintergrund, dass im weiteren Verlauf dieses Kapitels und auch in den weiteren Kapiteln zahlreiche Modelle trainiert werden müssen, um sowohl Hyperparameter zu optimieren als auch Grenzen der entwickelten Methodik zu untersuchen, werden die Bilddimensionen in den Trainingsdaten im Folgenden auf (256×256) Pixel

festgelegt. Diese Dimensionen orientieren sich weiterhin an den Ergebnissen der Literaturrecherche aus Kapitel 4.2.4 und sind ausreichend, um die Methodik zu untersuchen, ohne wertvolle Rechenzeit in einem unvernünftigen Maß zu verbrauchen. Eine Impli-

Tabelle 6.2: Vorhergesagte Kernladungszahlen in (a) und absolute Dichten in (b) für die verschiedenen Trainingsstrategien *resize* und *crop*. Die tabellierten Werte geben den Mittelwert μ und die Standardabweichung σ der Vorhersage in der Form $\mu \pm \sigma$ an. In der jeweils ersten Spalte werden die Materialien zusammen mit ihren Literaturwerten gelistet. Das jeweils bessere Ergebnis, pro Strategie und Material, ist grau hinterlegt. Alle Ergebnisse wurden mit einem U-Net [36] erzielt.

Material	<i>resize</i>	<i>crop</i>
Al (13)	13.03 ± 0.05	15.70 ± 1.08
Fe (26)	26.53 ± 0.08	26.81 ± 0.15
Cu (29)	28.69 ± 0.11	28.88 ± 0.62

(a) Kernladungszahl Z

Material	<i>resize</i>	<i>crop</i>
Al (2.70)	2.70 ± 0.01	2.65 ± 0.03
Fe (7.80)	7.81 ± 0.02	7.78 ± 0.03
Cu (8.93)	8.97 ± 0.02	9.14 ± 0.2

(b) Dichte ρ in g/cm³

kation aus dieser Wahl ist, dass die Materialbestimmung auch auf hochauflösenden CT-Schnittbildern mit deutlich größeren Bilddimensionen stets eine maximale Auflösung von (256×256) Pixel erreicht. Somit können kleine Partikel durch Anwendung der Resize-Strategie in den CT-Schichtbildern durch die Interpolation verfälschte Abschwächungskoeffizienten aufweisen oder vollständig verschwinden. Für einen spezifischen Anwendungsfall, beispielsweise die Materialbestimmung ebendieser Partikel, können aufbauend auf dieser Arbeit, mit der gleichen Methodik und genügend Rechenzeit, DL-Modelle auf deutlich größeren Bilddimensionen nach dem Resizing trainiert werden, um die auflösungsbedingte Präzision zu verbessern.

6.4 VERSCHIEDENE ARCHITEKTUREN: CONVOLUTION ODER TRANSFORMER?

Im Rahmen dieser Arbeit werden ausschließlich Architekturen für die Bildverarbeitung (CV) betrachtet, da die Materialbestimmung nach den Anforderungen aus Kapitel 5.1 als eine Bild-zu-Bild-Berechnung (Dense Prediction) formuliert wird. Mit der steigenden Popularität von Deep Learning ist in den letzten Jahren eine Vielzahl unterschiedlicher Architekturen für ebenso vielfältige Anwendungen vorgestellt und erprobt worden. Eine Architektur, die für vielfältige Anwendung aus verschiedensten Domänen angewandt wird [52, 53, 54, 55], ist das U-Net [36], das bereits in Abbildung 3.5 vorgestellt wurde. Das U-Net wurde im vorangehenden Abschnitt, motiviert durch die Ergebnisse der Literaturrecherche aus Kapitel 4.2.4, exemplarisch verwendet, um den Einfluss des globalen Bildkontextes innerhalb eines CT-Schichtbildes auf die Materialbestimmung zu untersuchen. In Abschnitt 6.4.1 werden die Kapazität des U-Nets (vergleiche θ Abschnitt 3.1) sowie die Trainingsdatenmenge systematisch optimiert, um die effizienteste Konfiguration im Kontext der formulierten Anforderungen für die Materialbestimmung zu finden.

Han et al. dokumentieren einen Trend, demzufolge Transformer-basierte Architekturen zur Bildverarbeitung zunehmend an Popularität gewinnen [111] und die Convolution-

basierten Architekturen durch ihre verbesserte Genauigkeit und Inferenzgeschwindigkeit in gängigen CV-Benchmarks zur Bildklassifikation [57], Objektdetektion und Segmentierung [37] verdrängen. Die Frage, ob Transformer-basierte Architekturen ebenfalls bei der Materialbestimmung anhand von CT-Schichtbildern die Convolution-basierten Architekturen übertreffen, ist bisher in der Literatur nicht quantitativ untersucht worden, sodass in ebendieser Untersuchung einer der Forschungsbeiträge dieser Dissertationsschrift besteht. In Abschnitt 6.4.2 wird eine Transformer-basierte Architektur vorgestellt, die für Dense Prediction geeignet ist und hinsichtlich der Materialbestimmung optimiert. Es erfolgt ein Vergleich zwischen dem U-Net und der Transformer-basierten Architektur, sodass eine geeignete Architektur für den weiteren Verlauf dieser Arbeit ausgewählt werden kann.

6.4.1 U-Net

Das U-Net wurde im vorangehenden Abschnitt in seiner ursprünglichen Form nach Ronneberger [36] verwendet. In diesem Abschnitt soll zunächst die Präzision der Materialbestimmung bei der Verwendung des U-Nets weiter verbessert werden, indem die Kapazität der U-Net-Architektur variiert wird. Einige Autoren berichten von Anpassungen der Feature-Map-Tiefe³ des U-Nets [65, 66, 74], sodass das U-Net für ihre spezifischen Anwendungen bessere Ergebnisse liefert oder Rahmenbedingungen, wie zum Beispiel eine reduzierte Inferenzzeit, besser erfüllt. Daher wird für die gegebene Fragestellung der Materialbestimmung die Tiefe der Feature-Map des U-Nets optimiert. Das U-Net wurde durch Ronneberger mit einer Feature-Map-Tiefe von 64 definiert [36]. Zur Vereinfachung wird die Notation U-Net^x verwendet, die ein U-Net mit der Feature-Map-Tiefe x bezeichnet. Tabelle 6.3 bietet einen Überblick über die im Folgenden getesteten Feature-Map-Tiefen. Die Feature-Map-Tiefen werden anhand von Zweierpotenzen gewählt, um ein effizientes Speicherlayout der zugrundeliegenden Rechenoperationen zu erreichen. Für jede Feature-Map-Tiefe werden zudem die Kapazität und die benötigte Rechenzeit für das Training über 100 Epochen angegeben. Beim U-Net führt eine Verdoppelung der

Tabelle 6.3: Kapazität des U-Nets und Rechenzeitbedarf τ zum Training über 100 Epochen für verschiedene Feature-Map-Tiefen. Die Angaben zum Rechenzeitbedarf wurden mit einem Rechenknoten bestehend aus vier NVIDIA H100 und einem Datensatz der Größe 250.000 (vergleiche Tabelle 5.1) ermittelt. Das Symbol † zeigt an, dass es sich um einen extrapolierten Wert aufgrund von Rechenzeitlimitierungen (maximal 24 Stunden pro Job) handelt.

Feature-Map-Tiefe	8	16	32	64	128	256
Kapazität / 1×10^6	0.5	1.9	7.5	30.0	120.0	480.0
τ in Stunden	2.4	2.6	3.4	4.8	9.2	25.8 [†]

Feature-Map-Tiefe zu einer Vervierfachung der Kapazität. Weiterhin ist zu beobachten, dass die benötigte Rechenzeit nicht proportional zur Kapazität (im tabellierten Bereich) ist, da mit steigender Kapazität der Datentransfer zwischen CPU- und GPU-Speicher

³ Beschrieben in Abschnitt 3.4.1: Die Anzahl der Kanäle des Tensors (der Feature-Map) vor der ersten Pooling-Layer im U-Net.

einen zunehmend kleineren Anteil der Rechenzeit ausmacht.⁴ Anhand der angegebenen Rechenzeiten in Tabelle 6.3 wird ersichtlich, dass die Modelle U-Net¹²⁸ und U-Net²⁵⁶, aufgrund der Rechenzeitbeschränkungen der Partition *WestAI* am Forschungszentrum Jülich, lediglich über 260 beziehungsweise 96 Epochen trainiert werden können. Die anderen Modelle werden bis zur Konvergenz – meistens über 500 Epochen – trainiert. Abbildung 6.7 fasst die Ergebnisse der Materialbestimmung für die verschiedenen Feature-Map-Tiefen nach erfolgtem Training der Modelle auf dem Trainingsdatensatz nach Tabelle 5.2 zusammen. Aluminium weist die niedrigsten Abweichungen (*MAE*) und gleichzeitig die

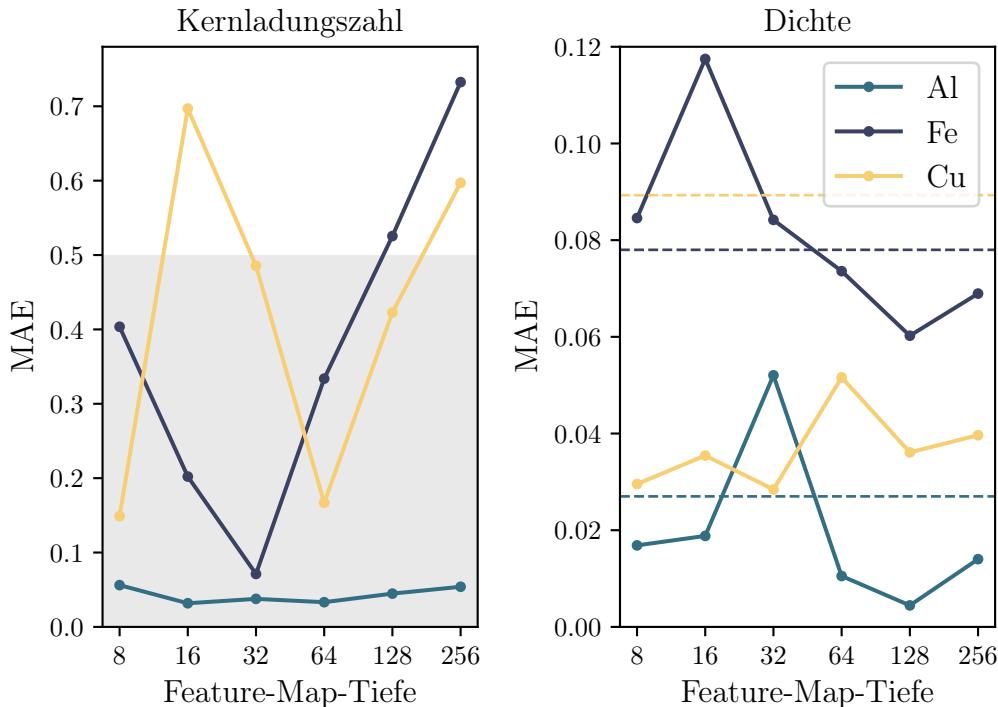


Abbildung 6.7: Mittlerer absoluter Fehler *MAE* der Vorhersage von Kernladungszahl (links) und Dichte (rechts) mit dem U-Net bei verschiedenen Feature-Map-Tiefen. Ausgewertet wurde mithilfe des in Abbildung 6.4 vorgestellten Phantoms anhand der Materialien Aluminium, Eisen und Kupfer. Der graue Bereich in der linken Abbildung markiert den Toleranzbereich (vergleiche Rahmenbedingungen in Abschnitt 6.1). Analog wird der obere Rand des Toleranzbereiches der Dichte für jedes Material im rechten Plot gestrichelt markiert.

höchste Stabilität der Kernladungszahlvorhersage zwischen den verschiedenen Modellen auf. Die *MAEs* der Materialien Eisen und Kupfer fluktuieren hingegen deutlich. Es ist zu beobachten, dass nur die Feature-Map-Tiefe 64 sowohl bei der Vorhersage der Kernladungszahl als auch bei der Dichte die geforderten Präzisionen von $\Delta Z = \pm 0.5$ (absolut) und $\Delta \rho = \pm 1\%$ (relativ) nach Abschnitt 6.1 für alle Testmaterialien erreicht. Somit wird die Feature-Map-Tiefe 64, in Übereinstimmung mit dem Vanilla U-Net nach

⁴ Im Allgemeinen werden Daten vom CPU-Speicher (auch: *Host Memory*) zunächst in den GPU-Speicher (auch: *Device Memory*) übertragen, bevor konkrete Berechnungen auf der GPU mit ebendiesen Daten ausgeführt werden können. In vielen Fällen stellt dieser Transfer einen Großteil der benötigten Berechnungsdauer dar, der, aus Sicht der GPU, reine Wartezeit ist. Mit zunehmender Kapazität des Modells steigt auch die benötigte Rechenzeit der Graphikkarte zur Verarbeitung eines Datenpakets und somit reduziert sich die Wartezeit auf das nächste Datenpaket. Mithilfe asynchroner Programmierung, die aufgrund ihrer Komplexität hier nicht weiter erläutert wird, werden die Wartezeiten der GPU signifikant reduziert und der Durchsatz verbessert.

Ronneberger [36], für die folgenden Untersuchungen verwendet.

Der zweite Optimierungsaspekt, der in diesem Fall ein Hyperparameter des Trainings ist, besteht in der Größe des Trainingsdatensatzes. Eine Verkleinerung des Trainingsdatensatzes kann die Simulation sowie das anschließende Training deutlich beschleunigen und somit effizienter gestalten. Aus der verwandten Forschung in Kapitel 4 geht hervor, dass einige Autoren Datensätze mit nur wenigen tausend Tupeln zum Training des U-Nets verwenden [65, 68, 73, 72]. Zunächst wird die Größenabhängigkeit des Trainingsdatensatzes auf die Materialbestimmung untersucht. Der zugrundeliegende Datensatz \mathcal{D} umfasst 250.000 Tupel (vergleiche Tabelle 5.2) – die verkleinerten Datensätze erfüllen die Relation

$$\mathcal{D}(10\text{ K}) \in \mathcal{D}(15\text{ K}) \in \mathcal{D}(25\text{ K}) \in \mathcal{D}(50\text{ K}) \in \mathcal{D}(100\text{ K}) \in \mathcal{D}(250\text{ K}) .$$

Aus dem ursprünglichen Datensatz $\mathcal{D}(250\text{ K})$ werden zufällig 100.000 Tupel ausgewählt, die den Datensatz $\mathcal{D}(100\text{ K})$ bilden. Analog wird für die kleineren Datensätze verfahren, sodass diese jeweils eine Teilmenge aller größeren Datensätze sind. Abbildung 6.8 stellt die Validation Losses mehrerer U-Nets dar, die jeweils auf den genannten Datensätzen trainiert worden sind. Im Gesamtbild von Abbildung 6.8 sinkt der finale Validation Loss

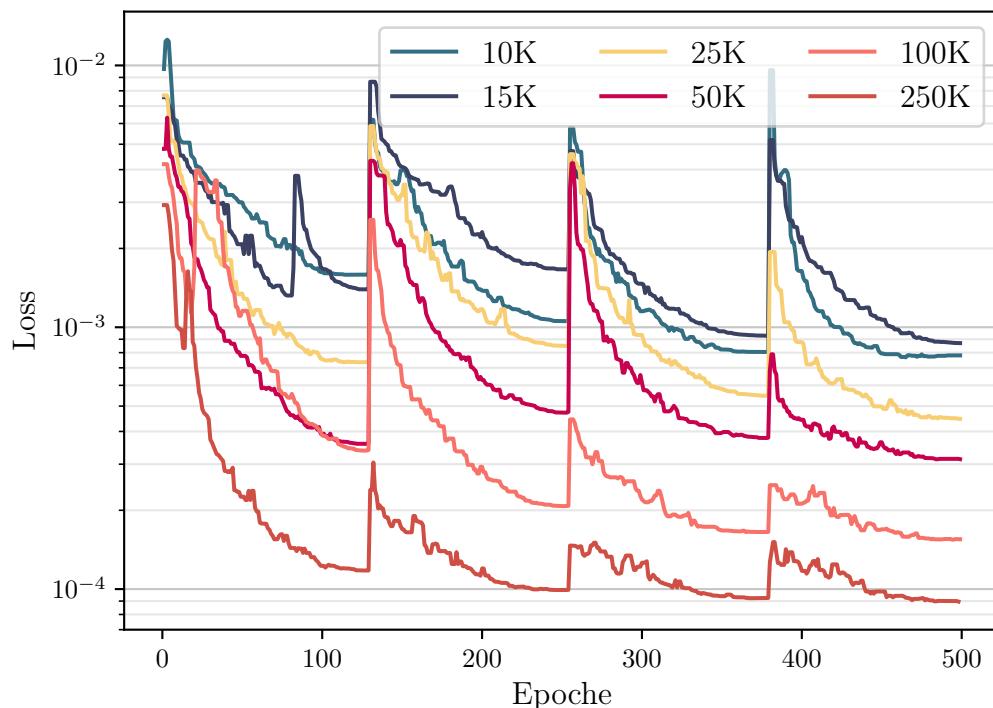


Abbildung 6.8: Validation Losses mehrerer U-Nets trainiert auf verschiedenen Datensatzgrößen (siehe Legende).

mit zunehmender Größe des Trainingsdatensatzes. Die einzige Ausnahme wird bei den Datensätzen der Größe 10K und 15K beobachtet, wobei es sich auch um statistische Schwankungen handeln kann. Grundsätzlich ist zu beobachten, dass die Validation Losses um bis zu eine Größenordnung im gezeigten Beispiel in Abbildung 6.8 variieren. Abbildung 6.9 bietet einen Überblick der Vorhersagen mit den soeben trainierten Modellen auf dem simulierten Testphantom aus Abbildung 6.5. Die Vorhersage der

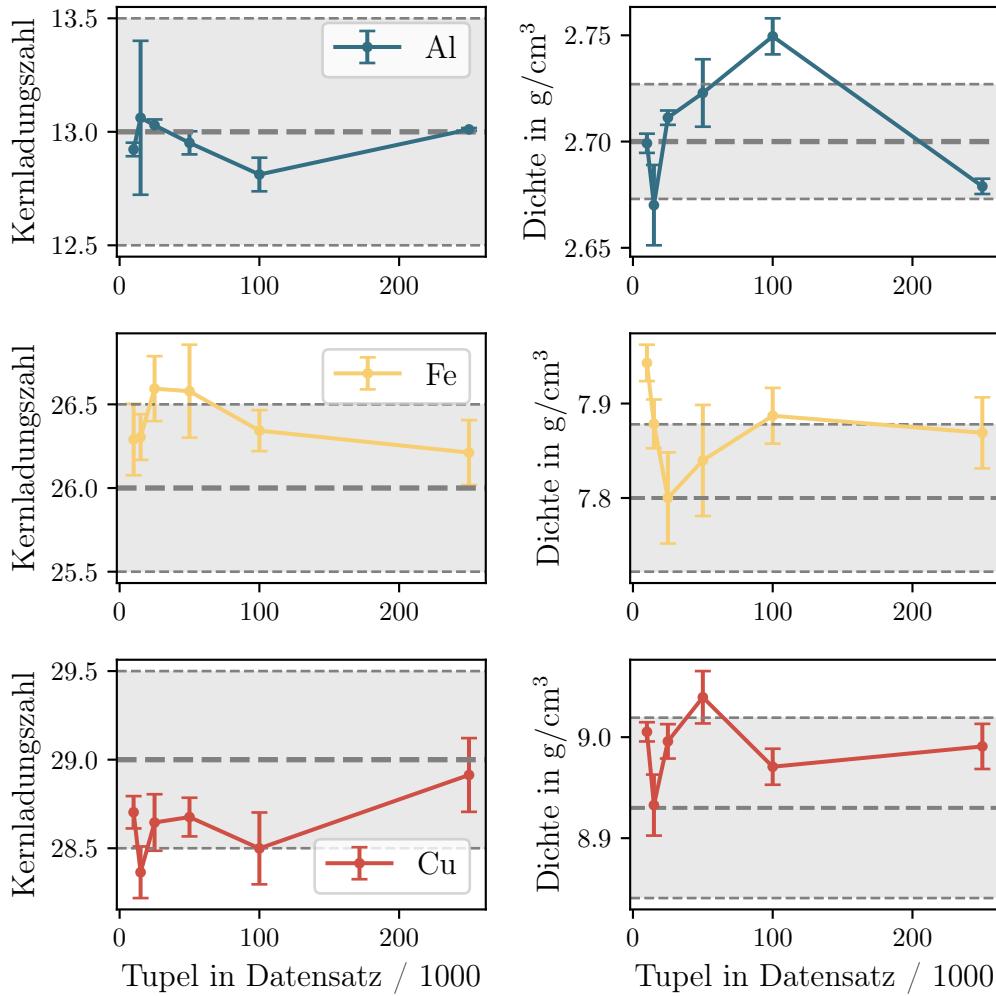


Abbildung 6.9: Vorhergesagte Kernladungszahlen und Dichten des Phantoms aus Abbildung 6.5 mit den U-Nets, trainiert auf verschiedenen Datensatzgrößen nach Abbildung 6.8. Die Literaturwerte sind grau und dick gestrichelt eingezeichnet. Die in Kapitel 6.1 definierten Präzisionen von ± 0.5 für die Kernladungszahl sowie $\pm 1\%$ für die Dichte sind grau und dünn gestrichelt als obere und untere Grenze des zulässigen Bereiches, grau hinterlegt, eingezeichnet.

Kernladungszahl liefert bei allen Modellen für die maximale Datensatzgröße (250K) die geringsten Abweichungen zu den Literaturwerten. Insbesondere die Vorhersage der Kernladungszahl des Aluminiums gelingt stabil innerhalb des Toleranzbereiches mit allen genannten Datensatzgrößen. Die Vorhersage der Dichte liegt für alle Materialien des Testphantoms nur bei den Datensatzgrößen 25K und 250K innerhalb des gültigen 1%-Bereiches. Bei allen anderen Datensatzgrößen liegt mindestens eine Vorhersage der Dichte außerhalb des genannten 1%-Bereiches. Im Folgenden wird zur besseren Vergleichbarkeit weiterhin die Standardgröße des Datensatzes mit 250.000 Tupeln verwendet. Eine weitere Vergrößerung des Trainingsdatensatzes wird aufgrund der bereits oben genannten Rechenzeitlimitierungen nicht weiter verfolgt, sollte jedoch aufbauend auf dieser Arbeit erprobt werden. Mit den benötigten Rechenzeiten aus Tabelle 6.3 wird ersichtlich, dass das U-Net⁶⁴ sehr genau innerhalb des Zeitlimits von 24 h über die geforderten 500 Epochen trainiert werden konnte. Durch die Vergrößerung des Datensatzes wird, analog zu den Trainingsdurchläufen der parameterstärkeren Modelle U-Net¹²⁸ und U-Net²⁵⁶, die Rechenzeitlimitierung überschritten. Es besteht dennoch die Möglichkeit, dass ein sehr viel größerer Datensatz bereits mit deutlich weniger Trainingsepochen eine robuste Materialbestimmung erlernen kann. Problematisch gestaltet sich in diesem Fall dennoch die Datenmenge des Trainingdatensatzes (vergleiche 5.1).

6.4.2 *Swin Transformer*

Nachdem in den vorangehenden Abschnitten dieses Kapitels ausschließlich das U-Net verwendet wurde, wird im Folgenden, motiviert durch die Erfolge des Transformers in anderen Anwendungen, eine Transformer-basierte Architektur für die Materialbestimmung erprobt. Es bietet sich an, eine Architektur zu untersuchen, die der Struktur des U-Nets im Kernkonzept folgt: ein Encoder, ein Decoder und Skip Connections zwischen den entsprechenden hierarchischen Ebenen des Encoders und Decoders. Die Idee, innerhalb der U-Net-Architektur den Convolution-basierten Encoder und Decoder durch einen Transformer-basierte Encoder und Decoder zu ersetzen, wurde bereits von Cao et al. für die Segmentierung klinischer CT-Schichtbilder [58] vorgeschlagen und ist somit als direkter Nachfolger des U-Nets, mindestens in der CT-Domäne, anzusehen. Abbildung 6.10 stellt die implementierte *Swin-Unet*-Architektur nach Cao et al. [58] schematisch dar. Der Encoder besteht in der Abbildung aus einem SwinV2-L nach Liu et al. [38], der bereits in Kapitel 3.4.2 eingeführt wurde. Das *Patch Embedding* berechnet für jedes Patch (vergleiche Kapitel 3.4.2) einen Vektor, dessen Länge *Embedding Dimension* genannt wird. Weitere Details zum Patch Embedding können bei Dosovitskiy [57] und Liu [37, 38] eingesehen werden. Im Decoder werden ebenfalls Swin-Transformer-Blöcke zusammen mit dem sogenannten *Patch Expanding* verwendet, das die Upconvolutions aus dem originalen U-Net (vergleiche Abbildung 3.5) ersetzt und das Swin-Unet hierdurch vollständig im Decoder auf Convolutions verzichten kann. Weitere Details zu der Architektur können in der zugehörigen Publikation [58] abgerufen werden. Analog zu der Feature-Map-Tiefe beim U-Net (vergleiche Tabelle 6.3) kann der Encoder des Swin-Unets mit unterschiedlichen Kapazitäten ausgelegt werden. Tabelle 6.4 fasst die im

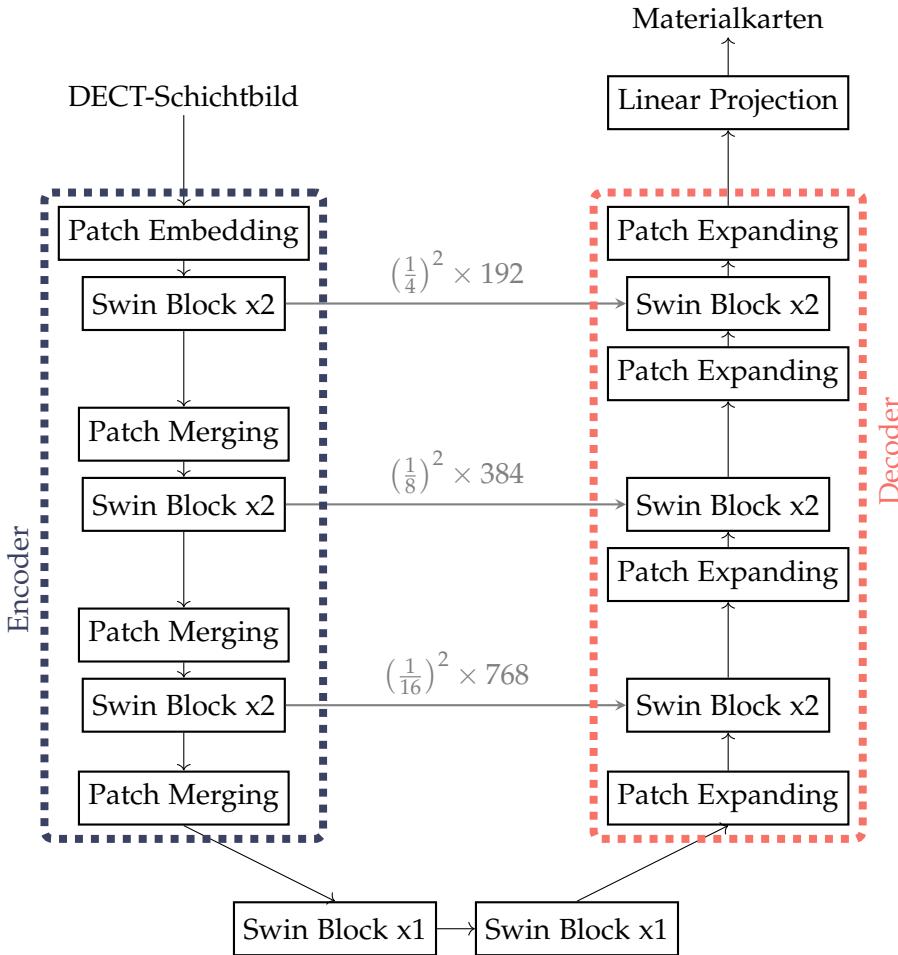


Abbildung 6.10: Architekturdiagramm des *Swin-Unet* nach Cao et al. [58] in leicht abgewandelter Form. Der dargestellte Encoder ist ein *SwinV2-L* nach Liu et al. [38] (vergleiche Tabelle 6.4).

Folgenden erprobten Swin-Unets mit verschiedenen Encodern zusammen. Aufgrund der Rechenzeitbeschränkungen der Partition *WestAI* am Forschungszentrum Jülich werden die parameterstarken Encoder *Swin-H* und *Swin-G* nicht betrachtet. In dem von Cao et al. vorgestellten Swin-Unet gibt es zudem die Bedingung, dass die Embedding Dimension ganzzahlig durch drei teilbar sein muss, sodass der Encoder *Swin-B* mit 128 Kanälen ausgeschlossen wird. Zur Vereinfachung wird im Folgenden ein Swin-Unet mit einem

Tabelle 6.4: Kapazität und Rechenzeitbedarf τ des Swin-Unets nach Cao et al. [58] zum Training über 100 Epochen für verschiedene Encoder in Form der Swin-Transformer-Modelle nach Liu et al. [38]. Der von Cao et al. verwendete Encoder wird nicht durch Liu et al. beschrieben und somit im Folgenden als *SwinV2-M* (*SwinV2-Minute*, gesprochen mar'njurt) bezeichnet. Die Angaben zum Rechenzeitbedarf wurden mit einem Rechenknoten bestehend aus vier NVIDIA H100 und einem Datensatz der Größe 250.000 (vergleiche Tabelle 5.1) ermittelt.

Encoder	SwinV2-M	SwinV2-S	SwinV2-L
Embedding Dimension	48	96	192
Kapazität / 1×10^6	21.4	84.4	336.2
τ in Stunden	6.1	6.3	10.4

SwinV2-L-Encoder mit *Swin-Unet-L* bezeichnet (analog *Swin-Unet-M* und *Swin-Unet-S*).

Cao et al. trainieren das Swin-Unet-M für die Segmentierung anhand eines Datensatzes, der 3779 annotierte Bilder umfasst und somit etwa drei Größenordnungen kleiner ist als der von Liu verwendete Datensatz [38]. Zudem initialisiert Cao die Parameter des Encoders mit Gewichten, die auf dem Datensatz *ImageNet* vorgenutzt worden sind. Die offiziell verfügbaren Swin-TransformerV2-Modelle wurden auf den Datensätzen *ImageNet-1K* ($\geq 10^6$ Bilder) und *ImageNet-22K* ($\geq 10^7$ Bilder) [94] vorgenutzt, die ein breites Spektrum an Bildern enthalten, jedoch keinen inhaltlichen Bezug zur CT aufweisen. Anschließend können die bereits vorgenutzten Modelle mithilfe eines verhältnismäßig kleinen Datensatz final an die Verteilung der neuen Daten angepasst werden – das sogenannte *Fine-Tuning*. Diese Strategie wird, in Anlehnung an die Arbeit von Cao [58], im Folgenden erprobt: Abbildung 6.11 stellt die Loss Curves zweier Trainingsdurchläufe mit dem Swin-Unet-S dar, wobei der Encoder einmal ohne das Vortraining initialisiert und einmal mit den vorgenutzten Gewichten des *ImageNet-22K*-Datensatzes geladen wurde. Es ist keine absolute Verbesserung des Validation Losses und keine Beschleunigung der

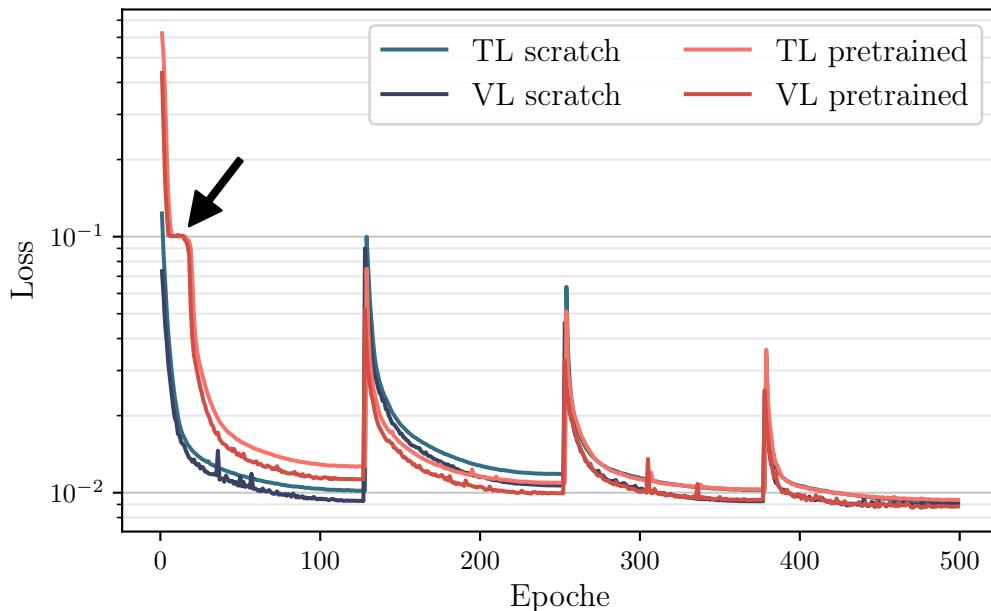


Abbildung 6.11: Train Loss *TL* und Validation Loss *VL* zweier Trainingsdurchläufe mit dem Swin-Unet-S nach Cao et al. [58]. Das Modell *scratch* verwendet keine vorgenutzten Gewichte, während das Modell *pretrained* die Gewichte für den SwinV2-S (Encoder) aus einem Vortraining auf dem *ImageNet-22K*-Datensatz verwendet. Der Pfeil kennzeichnet ein Plateau des Modells mit den vorgenutzten Gewichten.

Konvergenz durch den Einsatz der vorgenutzten Gewichte zu beobachten. Zusätzlich benötigt das Modell einige Epochen, gekennzeichnet durch den Pfeil in Abbildung 6.11, um sich an die neue Datenverteilung anzupassen. Wie bereits vermutet, unterscheidet sich die Datenverteilung der DECT-Schichtbilder im Rahmen dieser Arbeit zu stark von der Datenverteilung des *ImageNet-22K*-Datensatzes auf dem der Swin-Transformer vorgenutzt wurde. Es konnten keine Transformer-basierten Modelle gefunden werden, die auf CT-Daten vorgenutzt worden sind. Insgesamt konnte kein Mehrwert aus der Verwendung der vorgenutzten Gewichte des SwinV2-S gezogen werden. Die folgenden Swin-Unets werden ohne die Nutzung der vorgenutzten Gewichte direkt auf den simu-

lierten Daten für die Materialbestimmung trainiert.

Analog zu den Optimierungen des U-Nets im vorausgehenden Abschnitt, werden die verschiedenen Swin-Unets erprobt, um den besten Encoder für die gegebene Fragestellung zu ermitteln. Abbildung 6.12 stellt die mittleren absoluten Abweichungen der Materialvorhersagen des Swin-Unets mit verschiedenen Encodern dar. Die Modelle

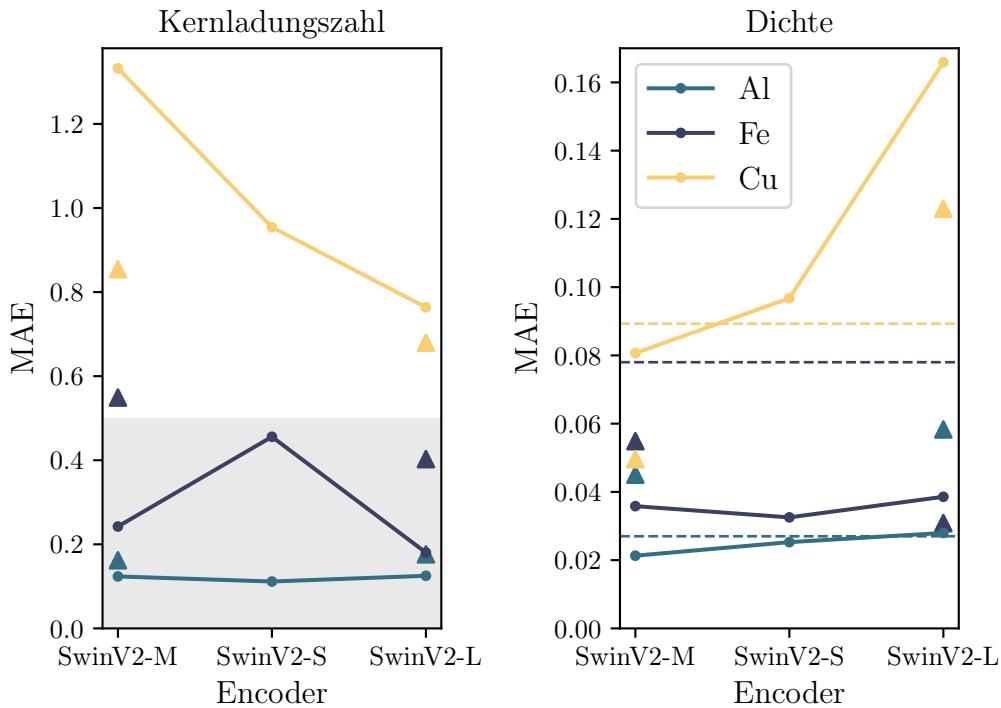


Abbildung 6.12: Mittlerer absoluter Fehler MAE der Vorhersage von Kernladungszahl (links) und Dichte (rechts) durch das Swin-Unet [58] mit verschiedenen Encodern aus Tabelle 6.4. Ausgewertet wurde mithilfe des in Abbildung 6.4 vorgestellten Phantoms anhand der Materialien Aluminium, Eisen und Kupfer. Der graue Bereich in der linken Abbildung markiert den Toleranzbereich (vergleiche Rahmenbedingungen in Abschnitt 6.1). Analog wird der obere Rand des Toleranzbereiches der Dichte für jedes Material im rechten Plot gestrichelt markiert. Die Swin-Unets mit den Encodern *SwinV2-M* und *SwinV2-L* wurden ergänzend mit einem vergrößerten Datensatz trainiert, der drei Millionen Tupel umfasst. Die entsprechenden Datenpunkte werden mit dreieckigen Markern dargestellt.

wurden unter maximaler Ausschöpfung der verfügbaren Rechenzeit trainiert, sodass die Anzahl der Trainingsepochen der ausgewerteten Modelle nicht gleich ist. Keinem der trainierten Swin-Unets gelingt die Materialbestimmung innerhalb der geforderten Toleranzbereiche. Die Kernladungszahl von Kupfer wird von keinem der trainierten Modelle innerhalb des Toleranzbereiches vorhergesagt. Für Aluminium und Eisen werden sowohl die Kernladungszahl als auch die Dichte innerhalb des Toleranzbereiches von den Modellen *Swin-Unet-M* und *Swin-Unet-S* korrekt vorhergesagt. Insgesamt liefert das Modell *Swin-Unet-M* die besten Ergebnisse in diesem Experiment, sodass es für den quantitativen Vergleich mit dem U-Net⁶⁴ aus dem vorangehenden Abschnitt verwendet wird. An diesem Punkt ist bereits klar, dass das Swin-Unet, aufgrund der fehlerhaften Vorhersage der Kernladungszahl von Kupfer, insgesamt schlechtere Ergebnisse bei der Materialbestimmung als das U-Net produziert – ein zunächst unerwartetes Ergebnis

vor dem Hintergrund der Überlegenheit der Transformer in anderen Disziplinen aus dem CV-Umfeld (vergleiche Han et al. [111]). Der zunächst offenkundigste Grund ist ein potentiell zu kleiner Trainingsdatensatz, sodass das Swin-Unet sein volles Potential nicht entfalten konnte (vergleiche Abschnitt 3.4.2). Daher wird zusätzlich mit einem deutlich vergrößerten Datensatz (3.000.000 Tupel), im Vergleich zu der in Tabelle 5.2 genannten Standardgröße, ein Training durchgeführt ohne eine signifikante Verbesserung verzeichnen zu können. Abbildung 6.12 enthält ebendiese Datenpunkte (dreieckige Marker). Weiterhin besteht die Vermutung, dass die Modelle aufgrund der Rechenzeitbeschränkungen nicht vollständig konvergieren konnten, sodass der Vergleich in Abbildung 6.12 mit den, auf kleineren Datensätzen, austrainierten Modellen unfair erscheint. Durch die Vergrößerung des Trainingsdatensatzes um den Faktor zwölf steigt gleichermaßen die benötigte Trainingszeit τ nach Tabelle 6.4 um ebendiesen Faktor an. Die weitere Untersuchung dieser Beobachtung sollte in einer zukünftigen Forschungsarbeit aus den oben genannten Gründen durchgeführt und mit den Ergebnissen von Han et al. [111] diskutiert werden.

6.4.3 Vergleich

Nach der Optimierung der beiden untersuchten Architekturen, dem U-Net [36] und dem Swin-Unet [58], werden die jeweils besten Modelle quantitativ in Tabelle 6.5 verglichen. Das Swin-Unet-M erzielt in vier von insgesamt sechs Testwerten (drei Materialien mit jeweils einer Kernladungszahl und einer Dichte) bessere Ergebnisse als das U-Net. Aus

Tabelle 6.5: Vorhergesagte Kernladungszahlen in (a) und absolute Dichten in (b) für das *U-Net* [36] und das *Swin-Unet-M* [58]. Die tabellierten Werte geben den Mittelwert μ , die Standardabweichung σ und die prozentuale Abweichung des Mittelwertes p von dem Literaturwert der Vorhersage in der Form $\mu \pm \sigma$ (p) an. In der ersten Spalte beider Tabellen werden die Materialien zusammen mit ihren Literaturwerten gelistet. Das jeweils bessere Ergebnis, pro Modell und Material, ist grau hinterlegt.

Material	<i>U-Net</i>	<i>Swin</i>
Al (13)	13.04 ± 0.01 (0.3 %)	13.06 ± 0.15 (0.5 %)
Fe (26)	26.32 ± 0.34 (1.2 %)	26.26 ± 0.18 (1.0 %)
Cu (29)	28.84 ± 0.17 (0.6 %)	27.76 ± 0.18 (4.3 %)

(a) Kernladungszahl Z

Material	<i>U-Net</i>	<i>Swin</i>
Al (2.70)	2.71 ± 0.00 (0.4 %)	2.70 ± 0.03 (0.0 %)
Fe (7.80)	7.85 ± 0.07 (0.6 %)	7.82 ± 0.04 (0.3 %)
Cu (8.93)	9.01 ± 0.02 (0.9 %)	8.87 ± 0.05 (0.7 %)

(b) Dichte in g/cm³

den Abbildungen 6.7 und 6.12 wurde bereits in den jeweiligen Abschnitten deutlich, dass nur das U-Net⁶⁴ alle Testmaterialien innerhalb der Toleranzbereiche korrekt identifiziert und somit die geforderten Rahmenbedingungen in Bezug auf die Präzision der Materialbestimmung erfüllt. Für die folgenden Untersuchungen wird demnach das U-Net⁶⁴ verwendet.

6.5 INFERENZ MIT DEM U-NET

Bisher wurden nur die Trainingsverläufe sowie die quantitativen Auswertungen mit Blick auf die Präzision der Materialbestimmung anhand des Testphantoms gezeigt. Die am Anfang dieses Kapitels genannten, industriellen Rahmenbedingungen umfassen neben der Präzision auch eine Anforderung an die Inferenzgeschwindigkeit der Architektur, die in diesem Abschnitt untersucht wird. Für eine feste Architektur, beispielsweise das ausgewählte U-Net⁶⁴ aus Abschnitt 6.4.1, ist die Inferenzzeit pro CT-Schichtbild lediglich von den räumlichen Dimensionen des Schichtbildes abhängig. Die Architektur des U-Nets basiert vollständig auf Convolutions (auch: *Fully-Convolutional*), sodass die Modelle, aus technischer Sicht, obwohl sie auf festen Bilddimensionen (256×256) Pixel trainiert worden sind, auf variablen Bilddimensionen ausgeführt werden können.⁵ Die im Folgenden ermittelten Inferenzgeschwindigkeiten sind somit auch für die U-Net⁶⁴-Modelle gültig, die, motiviert aus der Anwendungsperspektive (vergleiche Abschnitt 6.3), auf anderen Bilddimensionen trainiert worden sind. Abbildung 6.13 stellt die Inferenzzeit für unterschiedlich dimensionierte CT-Schichtbilder mit dem U-Net⁶⁴ dar. Zunächst ist

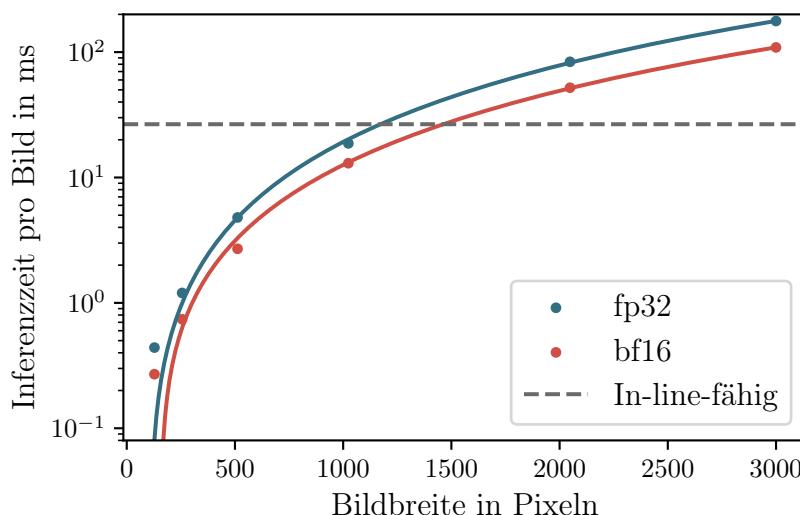


Abbildung 6.13: Benötigte Inferenzzeit pro (CT-Schicht-)Bild abhängig von der Bildbreite mit dem U-Net⁶⁴. Weiterhin sind die Bilder quadratisch, sodass die zu prozessierende Datenmenge quadratisch mit der Bildbreite ansteigt. Die durchgezogenen Linien sind quadratische Fits an die jeweiligen Datenpunkte. Die nach Abschnitt 6.1 maximale Inferenzzeit zur In-line-Fähigkeit beträgt für ein CT-Volumen bestehend aus 3000 Schichten und mit der Verwendung von acht Graphikkarten ≈ 27 ms und wird grau gestrichelt dargestellt. Für die gezeigten Ergebnisse wurde eine Nvidia H100 NVL verwendet.

gut zu erkennen, dass die Inferenzzeit ab einer Bildbreite von 256 Pixel quadratisch mit ebendieser skaliert.⁶ Das U-Net⁶⁴ ist nach Abbildung 6.13 bis zu einer Auflösung von (1024×1024) Pixel mit der gegebenen Hardware, das heißt acht Nvidia H100, In-

⁵ Der Effekt von unterschiedlichen Trainings- und Testdatenbilddimensionen wurde für die Bildklassifizierung von Touvron et al. [112] untersucht.

⁶ Durch die Verwendung quadratischer Bilder steigt die Gesamtzahl der Pixel quadratisch mit der Breite an. Somit ist die Inferenzzeit direkt proportional zur Gesamtzahl der Pixel.

Line-fähig. Bei voller Bildauflösung (3000×3000) Pixel beträgt die Berechnungszeit für die Materialbestimmung eines Volumens der Größe ($3000 \times 3000 \times 3000$) Voxel auf der genannten Hardware knapp 41 s.

6.6 ZUSAMMENFASSUNG DER ERGEBNISSE

In diesem Kapitel wurden zwei DL-basierte Architekturen, das U-Net [36] sowie das Swin-Unet [58], quantitativ anhand der Materialbestimmung optimiert und verglichen. Die systematische Untersuchung der Kapazitäten verschiedener U-Nets und Swin-Unets sowie der quantitative Vergleich dieser konzeptionell unterschiedlichen Architekturen für die Materialbestimmung anhand von CT-Bildern ist einer der Forschungsbeiträge dieser Dissertationsschrift. Es wird deutlich, dass die in der Literatur etablierte U-Net-Architektur für die Materialbestimmung, im Vergleich zum Swin-Unet, mit deutlich weniger Aufwand, in Form von Trainingsdaten und Rechenzeit, gute Ergebnisse erzielen kann. Weiterhin wird demonstriert, dass das Swin-Unet, obwohl es nicht alle Testmaterialien in dieser Arbeit erkennen konnte, das Potential birgt, präzisere Dichtevorhersagen als das U-Net zu produzieren (vergleiche Tabelle 6.5). Die in Abschnitt 6.1 geforderte Präzision wurde exklusiv durch das U-Net⁶⁴ erreicht, das ebenfalls bis zu einer Auflösung von (1024×1024) Pixel In-Line-fähig ist (vergleiche Abschnitt 6.5). Dennoch konnten die Anforderungen an die Präzision der Methode nur eingeschränkt im Rahmen dieses Abschnitts überprüft werden. Die vorausgehenden Abschnitte sowie Abbildung 6.7 geben einen Hinweis darauf, in welcher Größenordnung die Präzision der Vorhersage der Kernladungszahl und der Dichte unter idealen Bedingungen liegen kann. Für die Kernladungszahl ist, unter Ausschluss von Gemischen, eine ausreichende Genauigkeit mit einem Fehler von weniger als 0.5 (einheitenlos) nachgewiesen worden. Die Dichten der Testmaterialien wurden mit einem absoluten Fehler $< 0.1 \text{ g/cm}^3$ vorhergesagt. Alle Analysen bezüglich der Präzision wurden auf dem in Abschnitt 6.2 vorgestellten Phantom durchgeführt. Im folgenden Kapitel wird der Transfer in die reale CT-Anwendung überprüft. Viele der bisher gültigen Einschränkungen, beispielsweise die Fixierung der Probengröße (Kantenlänge), werden aufgehoben und ihr Einfluss auf die Vorhersagegenauigkeit geprüft. Ebenso ist die in Kapitel 5.4.2 beschriebene Sim-To-Real-Gap bisher nicht berücksichtigt worden, sodass die angegebenen Präzisionen, insbesondere in Tabelle 6.5, in einem realen Scan deutlich schlechter ausfallen werden. Am Ende des folgenden Kapitels wird die Diskussion zur Präzision der vorgestellten Methode daher erneut aufgegriffen.

7

TRANSFER IN DIE CT-PRAXIS

Nach der Realisierung einer Simulation zur Trainingsdatenerzeugung in Kapitel 5 sowie der Auswahl einer DL-Architektur in Kapitel 6 wird nun die Transferfähigkeit der gesamten Methodik zur Materialbestimmung in realistische CT-Scan-Szenarien überprüft. Die in Kapitel 5.1 abgeleiteten Anforderungen induzieren durch die gewählte Methodik der Trainingsdatenerzeugung, insbesondere durch die Fixierung einiger Simulationsparameter, messbare Einschränkungen für die Präzision der Materialbestimmung sobald von dem in der Simulation abgebildeten Szenario abgewichen wird. Diese Einschränkungen sind entweder im Phantomgenerator oder im implementierten CT-Projektor (vergleiche Abbildung 5.1) zu lokalisieren. Die Rekonstruktion erfolgt analog zu der Rekonstruktion realer Proben und induziert keine Abweichungen. Die Simulationsparameter werden im Folgenden in der durch den Algorithmus vorgegebenen Reihenfolge diskutiert.

Der Phantomgenerator erzeugt, pro Phantom, eine zufällige Anzahl an Regionen – die *Regionszahl*. Durch die Wahl des Intervalls aus dem die Regionszahl N (vergleiche Kapitel 5.3), beispielsweise $0 < N \leq 1000$ anhand von Tabelle 5.2, zufällig gezogen wird, ergibt sich indirekt die mittlere Größe der einzelnen Materialregionen innerhalb der simulierten Phantome. Der Einfluss ebendieser mittleren Regionsgröße hinsichtlich der Ortsauflösung der DL-basierten Materialbestimmung wird in Abschnitt 7.1 anschaulich motiviert und quantifiziert. Die durch den Phantomgenerator erzeugten Materialkarten fungieren als Ground Truths beim Training der Modelle. Sie weisen ideal scharfe Kanten zwischen den Materialregionen auf, sodass anzunehmen ist, dass die (Material-)Vorhersagen der trainierten Modelle ebenfalls ideal scharfe Kanten aufweisen. Dieser Aspekt wird in Abschnitt 7.2 exemplarisch, mithilfe einer in Abschnitt 7.1 eingeführten Methodik, untersucht. Die physische Größe des Phantoms, und damit auch die Größe des CT-Scanvolumens, ist ein weiterer Parameter des Phantomgenerators, der in praktischen Anwendungen, abhängig von der physischen Größe der zu scannenden Probe, stark variiert und einen deutlichen Einfluss auf die rekonstruierten Abschwächungskoeffizienten hat (vergleiche Strahlaufrärtungsartefakte; Kapitel 2.4.1). Zur Beschleunigung der Simulation¹ wird die physische Größe des Phantoms, die ebenso die CT-Geometrie beeinflusst, innerhalb der Simulation eines einzelnen Datensatzes konstant gehalten, sodass bei der Anwendung eines trainierten Modells auf eine Probe abweichender Größe die Vorhersagegenauigkeit potentiell sinkt. Die Vorhersagegenauigkeit, bezüglich der Materialbestimmung, wird in Abschnitt 7.3 abhängig von dem physischen Größenunterschied zwischen den Scanvolumen in den Trainings- und Testdaten analysiert. Nach den CT-spezifischen Anforderungen aus Kapitel 5.1 ist die Bestimmung poröser Materialien, beispielsweise für die Beurteilung einer Beschichtungen in einer Lithium-Ionen-Batterien,

¹ Durch das einmalige Anlegen der Datenstrukturen für die Rekonstruktion mit einer festen CT-Geometrie, ermöglicht durch das Fixieren der physischen Dimensionen des simulierten Phantoms, wird die Rekonstruktion signifikant beschleunigt und ermöglicht somit die in Kapitel 5.3 beschriebenen Simulationsgeschwindigkeiten.

relevant, wobei die Porosität, ausgedrückt durch die relative Dichte η , einen Einfluss auf die Vorhersagegenauigkeit der Materialbestimmung haben kann. Dieser Einfluss der relativen Dichte wird in Abschnitt 7.4 untersucht. Eine weitere Vereinfachung besteht in der Annahme der Fächerstrahl-CT, dem zweidimensionalen Grenzfall einer Kegelstrahl-CT, die als DL-spezifische Anforderung in Kapitel 5.1 genannt wurde. Bei realen CT-Scans werden, abgesehen von den selten eingesetzten Zeilendetektoren², typischerweise zweidimensionale Detektoren benutzt, die somit die Rekonstruktion eines dreidimensionalen Volumens, bestehend aus vielen übereinander gestapelten CT-Schichtbildern, ermöglichen. In Abschnitt 7.5 wird die Materialbestimmung auf allen CT-Schichtbildern, auch außerhalb der Mittelschicht, angewandt und ihre Genauigkeit überprüft.

Die bis zu diesem Punkt aufgezählten Einschränkungen existieren unabhängig von der in Kapitel 5.4.2 beschriebenen Sim-To-Real-Gap und können somit auf dem simulierten Testphantom aus Kapitel 6.2 analysiert werden. Abschließend wird der Einfluss der fehlenden Modellierung der Streustrahlung (vergleiche Kapitel 5.1 und 5.4.2) untersucht, indem ein trainiertes Modell zur Materialbestimmung auf realen, und damit streustrahlungsbehafteten, CT-Daten angewandt wird. Die verwendeten Daten wurden mit der CT-Anlage aus Kapitel 2.7 unter Verwendung der CT-Parameter aus Tabelle 5.2 aufgenommen.

Ziel dieses Kapitels ist primär die quantitative Untersuchung der Präzision der Materialbestimmung in Abhängigkeit der soeben genannten Einflussfaktoren, die von zentraler Bedeutung bei dem Transfer der Methodik in die reale CT-Anwendung sind. Weiterhin sollen Lösungen zur Reduktion der gefundenen Einschränkungen entwickelt oder mindestens diskutiert werden, sodass für zukünftige Forschungsarbeiten sinnvolle Ansatzpunkte angeboten werden.

7.1 GRÖSSE DER MATERIALREGIONEN

In Kapitel 5.3 wurde die implementierte Simulation beschrieben. Einer der zentralen Parameter bei der Erzeugung zufälliger Materialphantome ist die Anzahl der unterschiedlichen (Material-)Regionen pro Phantom (N in Tabelle 5.2). In der Simulation wird für jedes Phantom die Anzahl der Regionen N zufällig und gleichverteilt aus dem Intervall $[1, N_{max}]$ gezogen. Ebenso werden diese N Startpunkte der Regionen zufällig und gleichverteilt im verfügbaren Voxelgitter verteilt. Pro Zeitschritt wächst jede Region um einen festen Betrag in eine zufällige Richtung. Statistisch werden daher die einzelnen Regionen im Durchschnitt kleiner wenn N_{max} größer wird. Der Parameter N_{max} hat somit einen direkten Einfluss auf die mittlere Größe der einzelnen Regionen innerhalb aller Phantome, die zum Training des DL-Modells verwendet werden. Statistisch werden, beispielsweise für ein sehr kleines N_{max} , einzelne kleine Regionen im Trainingsdatensatz unterrepräsentiert, sodass einem trainierten Modell die Generalisierung auf derartig kleine Regionen vermutlich nur eingeschränkt gelingt. Ziel dieses Abschnittes ist die Untersuchung des räumlichen Auflösungsvermögens der trainierten Modelle

² Gemeint ist ein eindimensionaler Detektor, der nur eine Zeile an Pixeln besitzt und somit zur Fächerstrahl-CT genutzt wird.

auf Regionsgrößen, die von der mittleren Regionsgröße der Trainingsdaten abweichen. Zur quantitativen Bestimmung des Auflösungsvermögens kann die sogenannte *Modulationstransferfunktion* (kurz: *MTF*), oder auch *MTF-Kurve* als Funktion einer Frequenz, herangezogen werden [113, S. 8 f.]. Die MTF-Kurve gibt anschaulich den verbleibenden Kontrast eines Objektes in Abhängigkeit der Objektgröße, ausgedrückt als Ortsfrequenz f , an [113, S. 8 f.]. Eine häufig verwendete Einheit für die Ortsfrequenz ist lp/mm (*line pairs per millimeter*) oder cycles/Pixel – die Periode einer zyklischen Struktur. Anhand eines Siemenssterns, der eine variable, periodische Struktur abhängig von seinem Radius aufweist, kann das Auflösungsvermögen anschaulich bestimmt werden, wie die folgende Ausführung zeigt. Abbildung 7.1 zeigt einen rekonstruierten Siemensstern in kartesischen und polaren Koordinaten. Jedem Radius (rote Linie) kann ein Strichmuster, mit einer

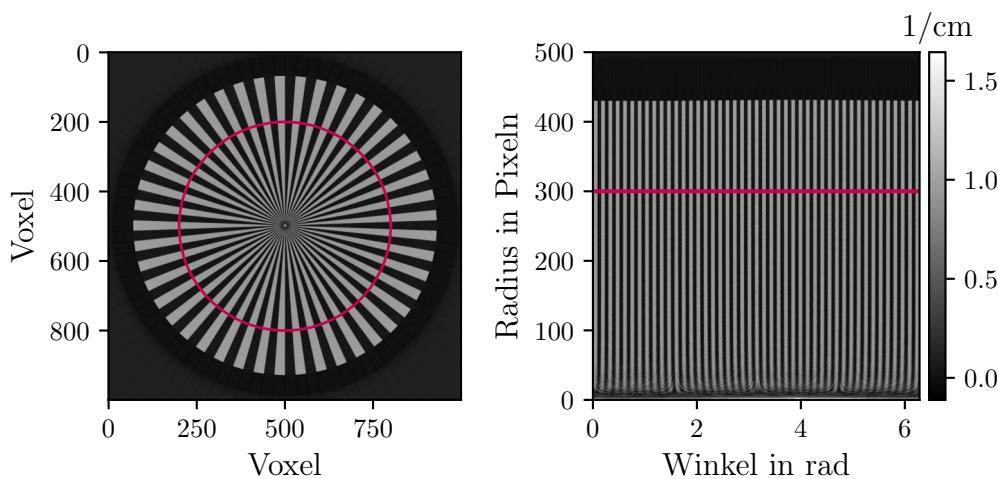


Abbildung 7.1: Rekonstruiertes Volumen eines simulierten Siemenssterns aus Aluminium in einem kartesischen Voxelgitter (links) und einem polaren Voxelgitter (rechts). Beide Abbildungen teilen sich die Grauwertskala rechts mit der Einheit cm^{-1} des Abschwächungskoeffizienten.

vom Radius abhängigen Ortsfrequenz, zugeordnet werden. Der Radius, bei dem noch 50 % des ursprünglichen Kontrastes des Strichmusters verbleiben, wird *MTF50-Radius* genannt. Anschaulich ist er genau der Radius, bis zu dem das Abbildungssystem (hier: das DL-Modell zur Materialbestimmung) noch die Hälfte des idealen Kontrastes ($\text{MTF} = 1$) darstellen kann.

Für die folgende Untersuchung wurden zwölf Datensätze basierend auf Tabelle 5.2 mit $N_{\max} \in \{25, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000\}$ simuliert und anschließend für das Training mehrerer U-Nets verwendet – jedes U-Net wird exklusiv mit einem Datensatz aus der genannten Menge trainiert. Die Modelle zur Materialbestimmung werden auf den Siemensstern aus Abbildung 7.1 (links) angewandt und berechnen Materialkarten, die anschließend durch eine Polartransformation in die Darstellung in Abbildung 7.2 (links) transformiert werden. Bereits ohne quantitative Berechnung des MTF50-Radius ist die Auflösungsgrenze der Materialkarte in Polarkoordinaten als Grenze zwischen dem eindeutig erkennbaren Strichmuster und der darunterliegenden, diffusen Fläche erkennbar. Abbildung 7.2 (rechts) stellt die normierte MTF-Kurve in Abhängigkeit des

Radius dar, sodass der MTF50-Radius (rote Linie in beiden Plots) quantitativ als Radius (x-Koordinate) für den Wert 0.5 der normierten MTF (y-Koordinate) bestimmt werden kann. Diese Methode zur Bestimmung der MTF50-Radien wird für die Materialvorhersa-

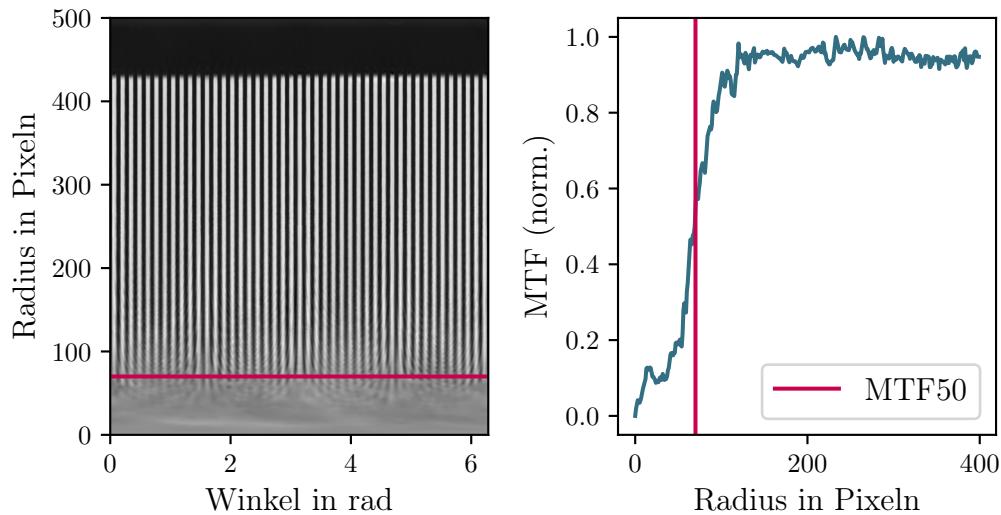


Abbildung 7.2: Dichtevorhersage eines Modells für den Siemensstern in Polarkoordinaten (links) und resultierende MTF-Kurve in Abhängigkeit vom Radius des Siemenssterns. Der MTF50-Radius ist in beiden Plots rot markiert.

gen der oben genannten Modelle durchgeführt. Somit wird für jedes Modell, trainiert auf einem Datensatz mit einem bestimmten N_{max} , der MTF50-Radius berechnet. Die Ergebnisse werden in Abbildung 7.3 graphisch zusammengefasst. Es ist kein eindeutiger

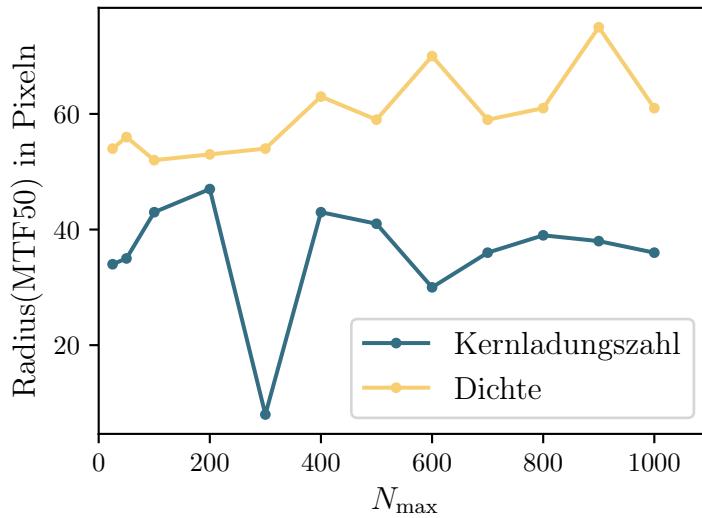


Abbildung 7.3: MTF50-Radien in Abhängigkeit von der maximalen Regionenzahl N_{max} für die Materialbestimmung auf dem Siemensstern aus Abbildung 7.1.

Zusammenhang zwischen den MTF50-Radien und der maximalen Regionenzahl N_{max} sowohl für die Bilder der Kernladungszahl als auch für die Bilder der Dichte zu beobachten. Somit besteht kein nachweisbarer Zusammenhang zwischen dem räumlichen Auflösungsvermögen des gewählten Materialbestimmungsansatzes und der maximalen

Regionszahl N_{max} in der Simulation. Die beobachteten Schwankungen sind vermutlich statistischer Natur und werden in der folgenden Auswertung der MTF50-Kurve mit einer alternativen Methode erneut beurteilt.

Eine weitere Methode zur Bestimmung der MTF-Kurve kann anhand eines gegebenen Profils, einem eindimensionalen Schnitt durch ein Bild, durchgeführt werden. Dieses Profil heißt *Edge-Spread-Function* (kurz: *ESF(x)*) mit der eindimensionalen Bildkoordinate x und es muss so gelegt werden, dass es die Kante eines Objektes im Bild, den maximalen Kontrastwechsel, erfasst. Anschaulich zeigt Abbildung 7.4 eine mögliche ESF und ihre Positionierung an der Grenzfläche zwischen dem Aluminiumwürfel und dem Hintergrund (Luft). Durch Differenzierung entlang der Bildkoordinate x wird die sogenannte

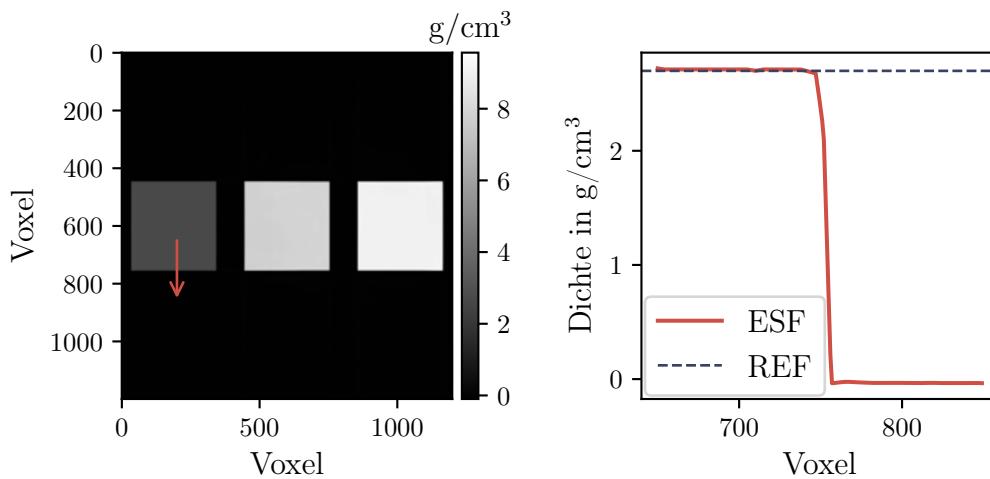


Abbildung 7.4: Vorhersage der Dichte durch ein trainiertes Modell auf dem Phantom aus Abschnitt 6.2 (links). Entlang des roten Pfeils wird die Edge-Spread-Function *ESF* extrahiert, die im rechten Plot dargestellt wird. Der Literaturwert für die Dichte des Aluminiums wird gestrichelt mit der Bezeichnung *REF* zusätzlich angegeben.

Line-Spread-Function (kurz: *LSF(x)*)

$$LSF(x) = \frac{d}{dx} ESF(x) \quad (7.1)$$

berechnet, die ein Maß für die Ortsunschärfe ist. Insgesamt ist die MTF einer Frequenz f aufbauend auf der $ESF(x)$ und der $LSF(x)$ durch

$$MTF(f) = \frac{|\mathcal{F}(LSF(x))(f)|}{|\mathcal{F}(LSF(x))(0)|} \quad (7.2)$$

mit der Fouriertransformation \mathcal{F} definiert [114, S. 225 f.]. Im Vergleich zu der phänomenologischen Bestimmung der MTF als Radius des Siemenssterns aus Abbildung 7.1 wird bei dieser Methode eine Ortsfrequenz in der Einheit 1/Pixel ermittelt – die sogenannte *MTF50-Frequenz*. Eine Umrechnung in die Einheit 1/mm kann unter Angabe der Voxelgröße des rekonstruierten Volumens trivial erfolgen.

Im Folgenden wird das Phantom aus Abbildung 6.4 verwendet, um die Materialbestimmung durchzuführen sowie die MTF50-Frequenz aus den Materialkarten zu ermitteln. Abbildung 7.5 (links) zeigt die MTF-Kurve eines extrahierten Kantenprofils am Rand des

Aluminiumsquaders nach erfolgter Materialbestimmung durch die trainierten Modelle. Die oben genannte Vermutung, einen Zusammenhang zwischen der maximalen Regi-

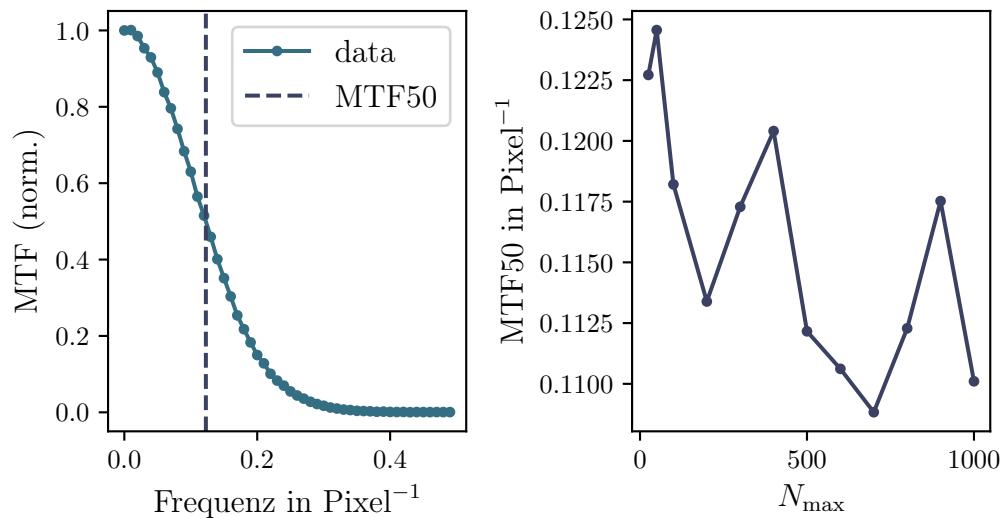


Abbildung 7.5: MTF-Kurve einer Kante mit eingezeichneter MTF50-Frequenz (links) für den Aluminiumwürfel aus dem Phantom in Abbildung 6.4. Die berechneten MTF50-Frequenzen für die verschiedenen maximalen Regionszahlen N_{\max} in den Trainingsdaten werden rechts dargestellt.

onszahl N_{\max} und der Auflösungsfähigkeit des Modells zu identifizieren, kann mittels Abbildung 7.5 (rechts) ebenfalls (vergleiche Abbildung 7.3) entkräftet werden. Entgegen der Intuition, dass ein Modell, welches mit einer höheren maximalen Regionszahl trainiert wurde, eine höhere MTF50-Frequenz besitzen sollte, zeigt die Auswertung einen gegenläufigen Trend. Die beste MTF50-Frequenz wird bei einer maximalen Regionszahl $N_{\max} = 50$ gemessen. Zur Erklärung dieser Beobachtung können mehrere Aspekte herangezogen werden. Die Strukturierung, oder Aufteilung der Regionen, innerhalb einzelner Phantome kann stark variieren. Statistisch ist es zwar wahrscheinlich, dass die verschiedenen Regionen im Mittel gleichgroß sind, jedoch können sich auch deutlich kleinere oder größere Regionen zufällig ausbilden. Zur Überprüfung dieser These werden die Regionsgrößen exemplarisch für Aluminium in einem der zuvor beschriebenen, simulierten Datensätze mit $N_{\max} = 100$ berechnet. Abbildung 7.6 stellt die gefundenen Regionsgrößen in Form eines Histogramms dar. Neben dem Erwartungswert werden kleinere und deutlich größere Regionen im gezeigten Datensatz beobachtet. Somit erscheinen Regionen verschiedenster Größe auch in einem Datensatz, der mit einem kleinen N_{\max} simuliert worden ist.³ Folglich hat der Parameter N_{\max} einen kleineren Einfluss auf die Größe der Materialregionen als zu Beginn dieses Abschnittes vermutet worden ist. Ein weiterer Aspekt ist, dass die Ground Truths (Materialkarten) aufgrund ihrer Erzeugung stets ideal scharfe Kanten aufweisen. Es ist daher sehr wahrscheinlich, dass die Modelle grundsätzlich lernen, Kanten scharf darzustellen, um die Materialkarten besser zu reproduzieren und so die Auswertung der MTF50-Frequenz anhand eines Kantenprofils nicht zur Quantifizierung des Auflösungsvermögens herangezogen werden

³ Die Voraussetzung ist lediglich, dass der Datensatz groß genug ist, sodass derartige, statistisch unterdrückte Effekte auftreten.

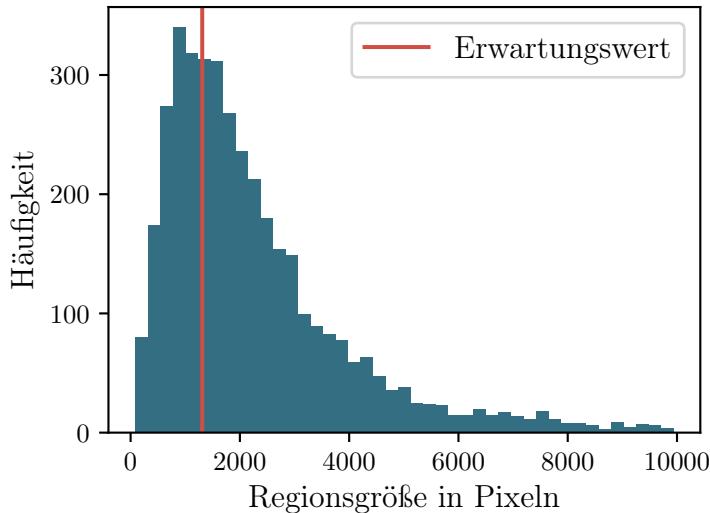


Abbildung 7.6: Histogramm der gefundenen Regionsgrößen von Aluminium (exemplarisch) in dem Trainingsdatensatz mit $N_{max} = 100$ anhand von 3200 zufällig ausgewählten Phantomen. Basierend auf der Auflösung jedes Phantoms (256×256) Pixel ergibt sich mit dem Erwartungswert der Regionenzahl $0.5 \cdot N_{max}$ der Erwartungswert der Regionsgröße zu 1311 Pixel.

kann. Dieser Aspekt wird im folgenden Abschnitt 7.2 untersucht. Weiterhin konnten keine Unterschiede der MTF-Kurven aus den Vorhersagen der Kernladungszahl und der Dichte ermittelt werden. Beide Kanäle verhalten sich, bezogen auf die MTF-Kurven, identisch.

Die meisten in diesem Abschnitt vorgestellten Modelle erreichen im Mittel eine Grenzfrequenz von 0.115 Pixel^{-1} , die mit der CT-Geometrie aus Tabelle 5.2 in eine Strukturgröße umgerechnet werden kann, bei der noch 50 % des Kontrastes abgebildet werden. Sie beträgt $\approx 174 \mu\text{m}$. Für eine praktische Anwendung, beispielsweise die Materialbestimmung kleiner Partikel innerhalb einer Probe, ist die MTF50-Frequenz nicht aussagekräftig, da sie ebenjene Ortsfrequenz ist, bei der noch die Hälfte des Kontrastes (Signal des Materials, beispielsweise die Dichte) abgebildet wird. Zur quantitativ korrekten Bestimmung der Dichte ($\pm 1\%$ nach Kapitel 5.1) eines Partikels mit einer Ortsfrequenz f_P , muss mindestens $\text{MTF}(f_P) \geq 0.99$ gelten.

7.2 KANTENSCHÄRFE

Die absolute Kantenschärfe ist für das Messen (Metrologie) an CT-Volumina von großer Bedeutung, sodass ein möglicher Verlust der Kantenschärfe durch die Anwendung der Materialbestimmung quantitativ analysiert werden muss. Aufbauend auf der im vorangehenden Abschnitt untersuchten Auflösungsfähigkeit der DL-basierten Materialbestimmung wird in diesem Abschnitt die Kantenschärfe der vorhergesagten Materialkarten im Vergleich zu den rekonstruierten Volumina ermittelt. Die bereits in Abschnitt 7.1 beschriebene MTF-Kurve, bestimmt anhand einer Objektkante, kann analog als Maß für die Kantenschärfe herangezogen werden. Scharfe Kanten werden im Gegensatz zu weichen Kanten aus höheren (Bild-)Frequenzanteilen dargestellt, sodass die MTF-Kurven, als Maß

für die Übertragung dieser (Bild-)Frequenzanteile, zum Vergleich der Kantenschärfen zwischen zwei Bildern genutzt werden kann. Abbildung 7.7 stellt die MTF-Kurven für die beiden DECT-Schichtbilder sowie die beiden Materialvorhersagen dar. Die Position der ESF, die den MTF-Kurven zugrundeliegt, entspricht der Markierung in Abbildung 7.4 (links). Bereits ohne quantitative Auswertung der MTF50-Frequenz wird anhand der

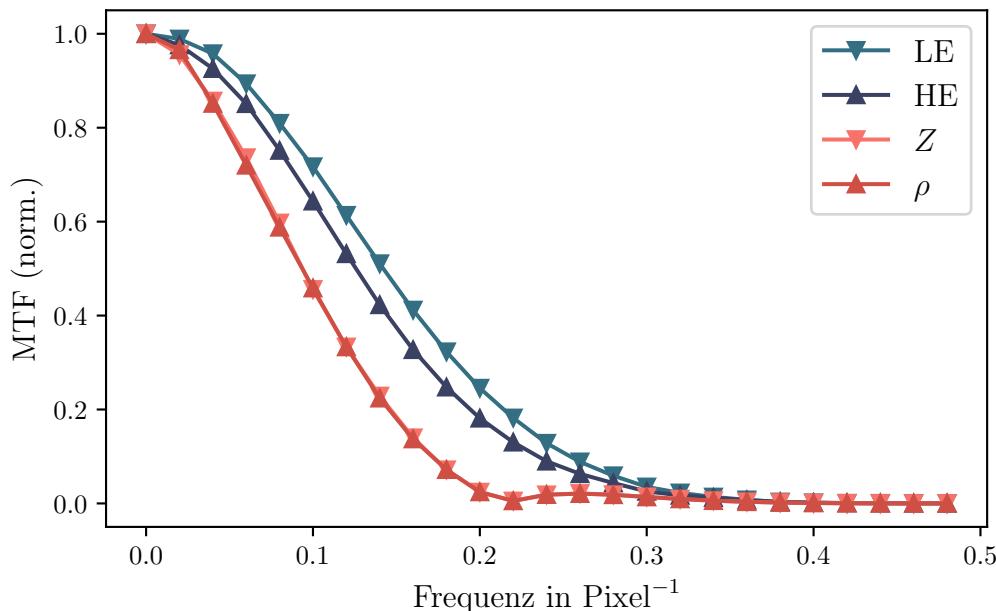


Abbildung 7.7: MTF-Kurven aus den beiden Kanälen des DECT-Schichtbildes (LE und HE) und der Materialvorhersagen (Kernladungszahl Z und Dichte ρ) über eine Kante des Aluminiumwürfels aus dem Phantom in Abbildung 6.4.

Lage der Kurven deutlich, dass die Materialbestimmung eine geringere Kantenschärfe als die beiden Kanäle des DECT-Schnitbildes *LE* und *HE* aufweist. Diese Beobachtung steht im Widerspruch zu den Trainingsdaten, da die Materialkarten absolut scharfe Kanten aufweisen und das Modell somit im Training die Darstellung ebendieser Kanten gelernt haben sollte. Zhou et al. zeigen für klinische CT-Schichtbilder, dass die Verwendung eines CNNs, konkret zur Rauschreduktion, einen negativen Einfluss auf die Ortsauflösung (MTF-Kurve) hat. Ein vergleichbarer Effekt kann anhand von Abbildung 7.7 für die Materialbestimmung dokumentiert werden. Zur Verbesserung der Kantenschärfe publizierten Mathieu et al. die Loss Function *Gradient Difference Loss*, die in Kombination mit dem, auch in dieser Arbeit verwendeten, MSE-Loss unscharfe (Bild-)Ausgaben reduziert [115]. Aufbauend sollte die von Mathieu et al. vorgestellte Loss Function für die Materialbestimmung erprobt werden, um dem Verlust der Kantenschärfe, der in diesem Abschnitt beobachtet werden konnte, entgegenzuwirken. Es wird festgehalten, dass die vorhergesagten Materialkarten eine verminderte Kantenschärfe im Vergleich zu den DECT-Schnitbildern aufweisen. Die Verbesserung der Kantenschärfe sollte Inhalt einer zukünftigen Forschungsarbeit sein, um insbesondere die Metrologie zu ermöglichen.

7.3 GRÖSSE DES SIMULIERTEN PHANTOMS

Eine weitere relevante Variable für den Transfer in die CT-Praxis ist die Größe, beziehungsweise Kantenlänge, des CT-Prüfvolumens. Wie bereits durch Abbildung 2.1 angedeutet, ist die CT-Geometrie ausschlaggebend für die Größe des Prüfvolumens. Die Größe des Prüfvolumens, beziehungsweise die Größe der zu prüfenden Probe, hat einen Einfluss auf die Ausprägung der Strahlaufhärtungsartefakte wie in Kapitel 2.4.1 beschrieben worden ist. Demnach wird vermutet, dass ein DL-basiertes Modell zur Materialbestimmung, das auf einer festen CT-Geometrie trainiert wurde, bei dem Transfer auf eine andere Geometrie aufgrund der veränderten Datenverteilung schlechte Vorhersagen produziert. Hierzu wird ein Testphantom, analog zu Abbildung 6.4, in aRTist simuliert, das insgesamt eine Kantenlänge von 1 cm aufweist. Es werden Modelle \mathcal{M} mit verschiedenen Kantenlängen des Simulationsvolumens trainiert und auf dem genannten Testphantom ausgewertet. Zur Vereinfachung wird die folgende Notation verwendet: $\mathcal{M}^{1\text{cm}}$ bezeichnet ein Modell, das mit Trainingsdaten der Kantenlänge 1 cm trainiert worden ist. Der mittlere absolute Fehler (MAE) über die Vorhersagen der Kernladungszahlen und Dichten der Materialien Aluminium, Eisen und Kupfer wird in Abbildung 7.8 abhängig von der Kantenlänge des Simulationsvolumens dargestellt. Erwartungsgemäß weist $\mathcal{M}^{1\text{cm}}$ den geringsten Fehler

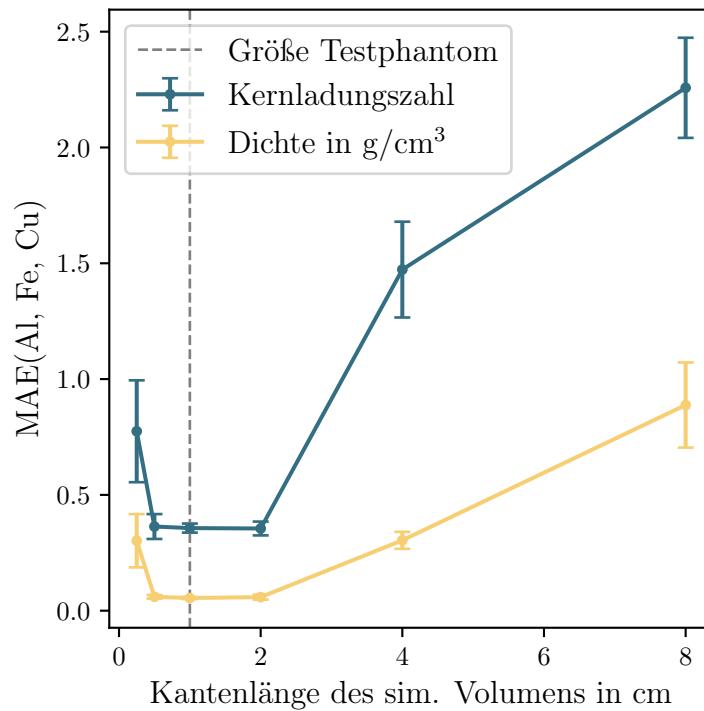


Abbildung 7.8: Mittlerer absoluter Fehler (MAE) der Kernladungszahl- und Dichtevorhersagen von Aluminium, Eisen und Kupfer in einem quadratischen Phantom mit der Kantenlänge 1 cm (übrige CT-Parameter siehe Tabelle 5.2). Mehrere Modelle wurden mit verschiedenen Kantenlängen des Simulationsvolumens trainiert (x-Achse). Die graue Linie zeigt die Kantenlänge des Testphantoms (1 cm) an. Die Fehlerbalken entsprechen einer gemittelten Standardabweichung über alle genannten Materialien.

bei der Auswertung des Testphantoms der gleichen Kantenlänge auf. Die Vorhersagen

der Modelle $\mathcal{M}^{0.5\text{cm}}$ und $\mathcal{M}^{2\text{cm}}$ weisen einen vergleichbar niedrigen MAE auf, während die verbleibenden Modelle mit zunehmend abweichender Kantenlänge ebenso zunehmende MAEs verzeichnen. Somit wird eine deutliche Einschränkung des vorgestellten Ansatzes zur Materialbestimmung deutlich: Für jede Kantenlänge des CT-Prüfvolumens wird ein Modell benötigt, das auf Trainingsdaten der entsprechenden Kantenlänge, oder einer sehr ähnlichen Kantenlänge, trainiert worden ist, um bestmögliche Ergebnisse zu erzielen. Da keine allgemeingültige Aussage zu dieser mindestens geforderten Ähnlichkeit der Kantenlängen getroffen werden kann, sollte in praktischen Anwendungen mit der exakt geforderten Kantenlänge trainiert werden, um Abweichungen dieser Art auszuschließen. Andererseits besteht theoretisch die Möglichkeit auf einem gemischten Datensatz, der viele verschiedene, simulierte Kantenlängen enthält, zu trainieren, sodass sich ein robustes DL-Modell ergibt. Dieser Ansatz wurde in erster Instanz nicht verfolgt, da die Rekonstruktion mit einer fixen CT-Geometrie signifikant schneller ist und somit die Rechenzeit der gesamten Simulation signifikant reduziert wird (vergleiche 5.4.4). Weiterhin ist zu beachten, dass verschiedene Kantenlängen zu unterschiedlich ausgeprägten Strahlaufläufungsartefakten führen und sich somit die Verteilungen der linearen Abschwächungskoeffizienten deutlich zerstreuen (vergleiche Kapitel 5.4.3). Durch die große Zahl an Mehrdeutigkeiten in den Trainingsdaten ist der Mehrwert des Trainings auf einem derartig universellen Datensatz gegenüber dem Training einzelner, auf eine Kantenlänge spezialisierter Modelle fraglich. Eine mögliche Lösung dieses Problems wird im folgenden Abschnitt durch die Anpassung eines spezialisierten Modells auf eine andere Kantenlänge diskutiert und erprobt.

7.3.1 Fine-Tuning zur Adaption der Kantenlänge

Zwar ist das U-Net in der gewählten Konfiguration verhältnismäßig schnell, das heißt günstig, zu trainieren, jedoch erscheint das Bereitstellen einer beliebig feingranularen Sammlung trainierter Modelle für jede CT-Geometrie, beziehungsweise Probengröße, nach den Resultaten aus dem vorausgehenden Abschnitt, insbesondere Abbildung 7.8, nicht wirtschaftlich. Eine bereits aus Kapitel 6.4 bekannte Lösung besteht in der Anwendung des Fine-Tunings. Kurz gefasst wird ein vortrainiertes Modell vorausgesetzt, das mithilfe eines kleinen Datensatzes an eine neue Datenverteilung angepasst wird. Wie bereits angemerkt und stichprobenhaft im Vorfeld dieser Arbeit erprobt, ist das Training eines einzelnen Modells auf einer großen Bandbreite an Kantenlängen aufgrund der variierenden Strahlaufläufungsartefakte wenig aussichtsreich. Im Folgenden wird das Fine-Tuning erprobt, um ein vortrainiertes Modell $\mathcal{M}^{4\text{cm}}$ mithilfe kleiner Datensätze an eine neue Datenverteilung, ein simuliertes Volumen mit der Kantenlänge 1 cm, anzupassen. Zunächst wird die benötigte Datenmenge, das heißt die Anzahl der Tupel ebendieses kleinen Datensatzes, untersucht, indem mehrere Modelle mit unterschiedlich großen Datensätzen dem Fine-Tuning unterzogen werden. Die verwendeten Datensatzgrößen betragen exemplarisch 1000, 2000, 4000 und 8000 Tupel. Eine detaillierte Optimierung dieser Datensatzgröße spielt für reale Anwendungen eine zu vernachlässigende Rolle, da die Simulation zur Trainingsdatenerzeugung aus Kapitel 5 in weniger als vier Mi-

nuten 8000 Tupel berechnen kann. Abbildung 7.9 stellt die MAEs der vorhergesagten Kernladungszahlen und Dichten abhängig von der jeweils verwendeten Größe des Trainingsdatensatzes für das Fine-Tuning dar. Es ist zu beobachten, dass 4000 Datentupel

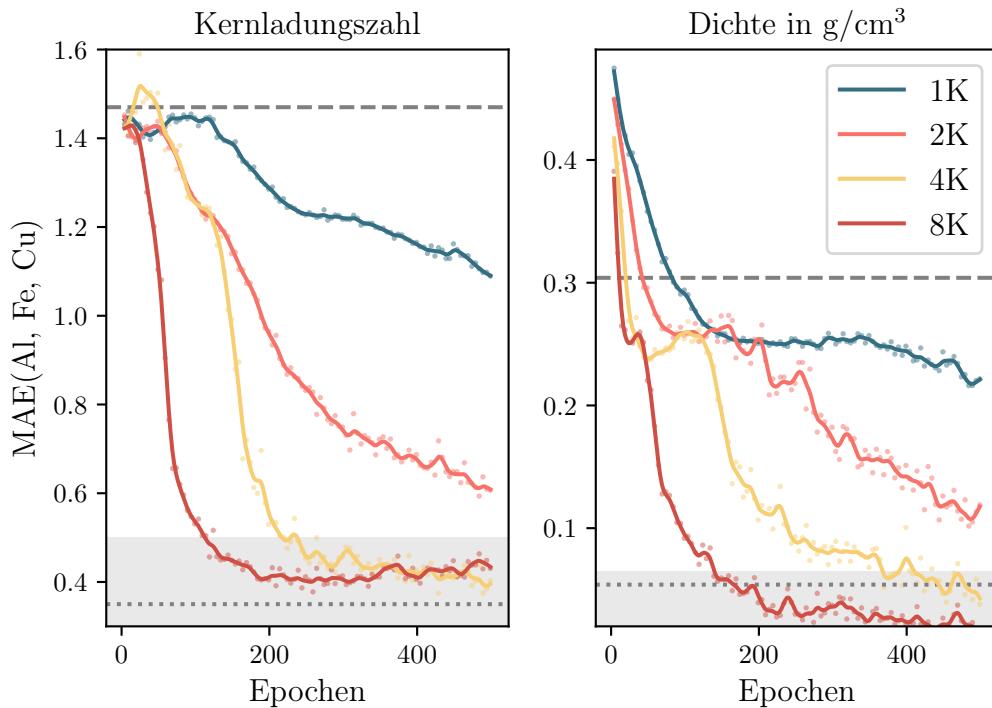


Abbildung 7.9: Mittlerer absoluter Fehler (MAE) der vorhergesagten Kernladungszahlen und Dichten von Aluminium, Eisen und Kupfer in einem quadratischen Phantom mit der Kantenlänge 1 cm (übrige CT-Parameter siehe Tabelle 5.2) während des Fine-Tunings des Modells $\mathcal{M}^{4\text{cm}}$ mit verschiedenen Datensatzgrößen (Legende), die auf einer Kantenlänge 1 cm basieren. Die grauen Linien geben die *Baseline* (obere Linie), das heißt den Fehler des vortrainierten Modells $\mathcal{M}^{4\text{cm}}$ ohne Fine-Tuning, sowie den MAE des Modells $\mathcal{M}^{1\text{cm}}$ (untere Linie) an. Die in den Rahmenbedingungen geforderte Präzision der Kernladungszahl- und Dichtevorhersage wird als grauer Toleranzbereich eingezeichnet.

ausreichen, um das Modell so anzupassen, dass die geforderte Präzision erreicht wird (gelbe Kurve). Die Trainingszeit über 500 Epochen beträgt für ebendiese Datensatzgröße knapp 27 min auf vier Nvidia H100. Zudem besteht die Möglichkeit, dass ein derartig kleiner Trainingsdatensatz vollständig im GPU-Speicher vorgehalten werden kann und die angegebene Trainingszeit somit signifikant reduziert wird.⁴ Diese Optimierung kann für eine praxisnahe Anwendung implementiert werden und ist nicht Teil dieser Arbeit. In Summe zeigen die Ergebnisse, dass mithilfe des Fine-Tunings kleine Anpassungen bezüglich der CT-Geometrie⁵ nachträglich auf vortrainierten Modellen in einer kurzen Zeitspanne durchgeführt werden können. Weiterhin können mit den gegebenen sechs

⁴ Pro GPU müssen in diesem Rechenbeispiel 1000 Tupel verarbeitet werden, die etwa 1 GB Speicher benötigen. In der aktuellen Implementierung werden die Batches, bestehend aus 128 Tupeln, von einer Festplatte in den CPU-Speicher geladen und anschließend in den GPU-Speicher kopiert. Dieser Schritt kann bei einem Datensatz, der vollständig in den GPU-Speicher passt, einmalig zu Beginn des Trainings durchgeführt werden, sodass die Berechnungszeit pro Epoche deutlich reduziert werden kann.

⁵ In diesem Abschnitt wurde gezielt das Fine-Tuning eines Modells auf eine neue Kantenlänge des rekonstruierten Volumens untersucht. Im Allgemeinen werden die Maße des rekonstruierten Volumens durch die CT-Geometrie festgelegt.

Stützpunkten ($\mathcal{M}^{0.1\text{cm}}$, $\mathcal{M}^{0.5\text{cm}}$, $\mathcal{M}^{1\text{cm}}$, $\mathcal{M}^{2\text{cm}}$, $\mathcal{M}^{4\text{cm}}$ und $\mathcal{M}^{8\text{cm}}$) mittels Fine-Tuning die Kantenlängen zwischen 1 mm und 8 cm abgedeckt werden – ein realistischer Einstellbereich des Prüfvolumens einer hochauflösenden CT-Anlage.

7.4 RELATIVE DICHTEN

Neben massiven Proben können auch poröse Werkstoffe, beispielsweise aufgrund ihrer vergrößerten Oberfläche⁶, von technischer Relevanz sein. Ein bekanntes Beispiel hierfür ist die Elektrodenbeschichtung in modernen Lithium-Ionen-Batterien, die häufig aus einem porösen Material aufgebaut wird. Im Bestfall ist die räumliche Auflösung des CT-Systems so hoch, dass die einzelnen Poren innerhalb einer derartigen Probe aufgelöst werden können. Somit werden Voxel rekonstruiert, die vollständig (massiv) mit Material gefüllt sind und andere Voxel rekonstruiert, die nur Luft enthalten und somit leer sind. Abseits dieses Idealfalls tritt der in Kapitel 2.4.1 beschriebene Partial Volume Effect auf. Manche Voxel sind teilausgefüllt und ihr linearer Abschwächungskoeffizient ist proportional zur relativen Dichte η – dem relativen Volumenanteil eines Materials an einem Voxel. In Kapitel 5.3 wurde die implementierte Simulation vorgestellt, wobei der Faktor η die Porosität der simulierten Phantome bestimmt (vergleiche Abbildung 5.1). Es wurde zunächst festgelegt, dass $\eta \in [0.8, 1.0]$ gilt – die zu untersuchenden Proben sind vermutlich auf ebendiesen Wertebereich der relativen Dichte eingeschränkt. Für die Allgemeingültigkeit des vorgestellten Ansatzes zur Materialbestimmung wäre es darüber hinaus vorteilhaft, wenn die relative Dichte beliebige Werte aus dem Intervall $[0, 1]$ annehmen könnte. In diesem Abschnitt wird untersucht, welchen Einfluss die untere Schranke dieses Intervalls auf die Präzision der Materialbestimmung hat. Zunächst werden die Verteilungen zweier Trainingsdatensätze verglichen, die mit unterschiedlichen relativen Dichten simuliert wurden. Abbildung 7.10 zeigt den Unterschied der Verteilungen der Abschwächungskoeffizienten im High-Energy-Kanal für zwei Simulationen mit unterschiedlichen Wertebereichen der relativen Dichte η . Im Fall $\eta = 1$ werden nur massive Materialien simuliert. In Abbildung 7.10 ist zu erkennen, dass die Simulation poröser Materialien die Verteilung der Abschwächungskoeffizienten insgesamt glättet beziehungsweise aufweicht (blaue Kurve). Dieser Effekt ist anhand von Gleichung 2.3 sofort ersichtlich, da durch den größeren Wertebereich von ρ_m insgesamt ein breiteres Spektrum an Abschwächungskoeffizienten $\mu_{L,m}$ pro Material gemessen werden kann. Weiterhin führt dieser Effekt zu einer Verstärkung der Mehrdeutigkeiten bei der Materialbestimmung (vergleiche Abbildung 5.10). Im Vergleich dazu weist die Verteilung ohne Porosität (rote Kurve in Abbildung 7.10) deutlich voneinander getrennte Häufungspunkte auf, die somit die verschiedenen Materialien zumindest teilweise voneinander trennen. Es folgt die Vermutung, dass das Erlernen der Materialbestimmung anhand eines Datensatzes mit einem größeren Intervall der relativen Dichten η für ein DL-Modell schwieriger ist. Diese Vermutung wird im Folgenden quantitativ untersucht. Es werden Trainingsdatensätze mit verschiedenen unteren Grenzen der relativen Dichten

⁶ Die Vergrößerung einer (Kontakt-)Oberfläche spielt häufig bei der Optimierung chemischer Reaktionen eine wichtige Rolle.

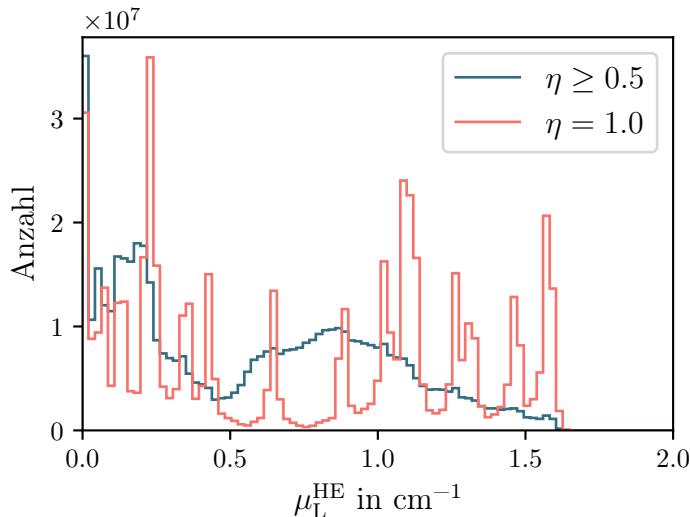


Abbildung 7.10: Histogramme aus 8000 zufällig gewählten Datentupeln aus den Simulationen mit einer relativen Dichte $\eta \geq 0.5$ und einer relativen Dichte $\eta = 1.0$. Dargestellt werden die High-Energy-Kanäle mit einer Beschleunigungsspannung $U_B = 240 \text{ kV}$ und den Simulationsparametern aus Tabelle 5.2.

$\eta \geq \{0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$ simuliert und zum Training mehrerer U-Nets [36] verwendet. Ein Ausschnitt der beim Training ermittelten Validation Losses ist in Abbildung 7.11 abgebildet. Gemäß der obigen Hypothese erreichen die trainierten Modelle mit einer

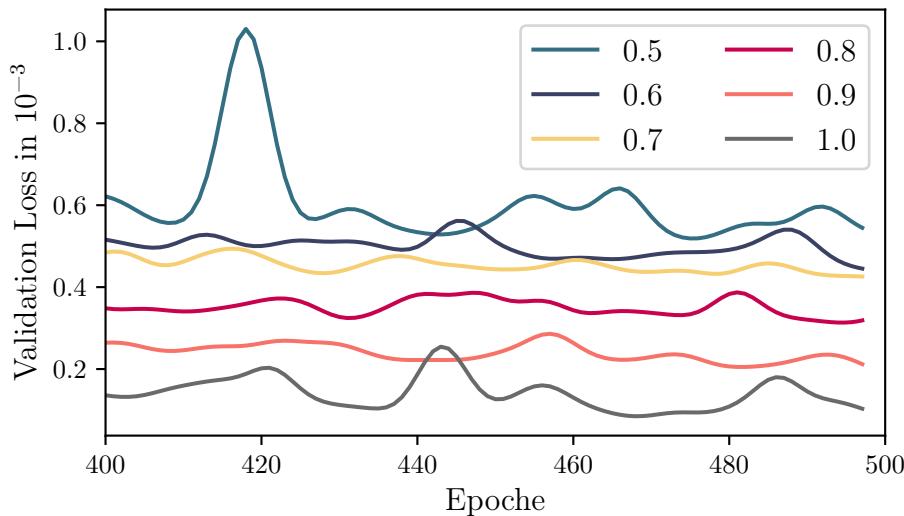


Abbildung 7.11: Validation Losses der U-Nets ab der 400. Trainingsepoke mit verschiedenen $\eta \geq \{0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$ im Trainingsdatensatz (Legende). Die Validation Losses wurden zur Verbesserung der Darstellung mit einem eindimensionalen Gaussfilter (Standardabweichung $\sigma = 3$) geglättet.

Abnahme der Intervallgröße von η höhere Genauigkeiten – einen niedrigeren Validation Loss. Darauf aufbauend zeigt Abbildung 7.12 die besten Validation Losses der Modelle für verschiedene, minimale, relative Dichten η . Es ist ein grob linear fallender Trend zu beobachten. Er unterstützt die Vermutung, dass kleinere Intervalle (weniger Porosität) einfacher für das Modell zu erlernen sind. Auf die Berechnung der extrapolierten Da-

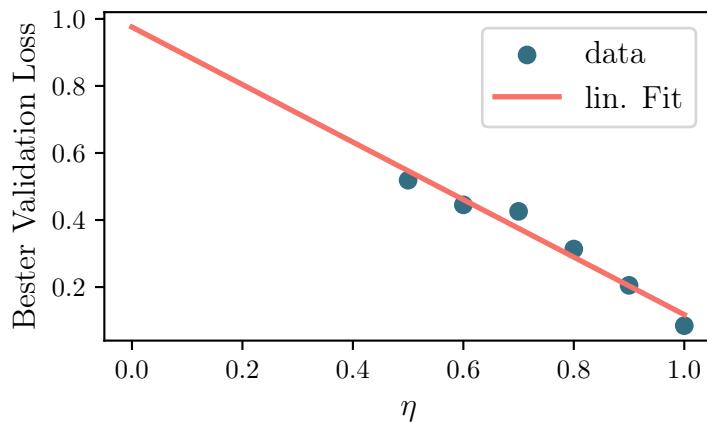


Abbildung 7.12: Beste Validation Losses (multipliziert mit 10^3) der U-Nets aus den Trainings mit verschiedenen $\eta \geq \{0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$. Eine lineare Funktion ist an die Daten angepasst und auf den gesamten Definitionsbereich von η extrapoliert.

tenpunkte für $\eta < 0.5$ wird zugunsten der verfügbaren Rechenzeit bewusst verzichtet. Zudem stehen derzeit keine Proben, die eine derart niedrige relative Dichte aufweisen, im Fokus der Anwendung. Die Ausweitung der Methodik auf beispielsweise hochporöse, technische Metallschäume mit relativen Dichten $\eta < 0.5$ und Porengrößen, die unterhalb der CT-Auflösung liegen, ist möglich und kann für einen gegebenen Anwendungsfall durchgeführt werden.

Zuletzt ist die Materialbestimmung auf einem geeigneten, simulierten Phantom zu überprüfen. Mithilfe der CT-Parameter aus Tabelle 5.2 wird ein Aluminiumwürfel mit den relativen Dichten $\eta \in \{0.8, 0.85, 0.9, 0.95, 1.0\}$ simuliert. In den simulierten Trainingsdaten erscheint poröses Aluminium beispielsweise Al^{80%} mit der Kernladungszahl $Z = 13$ und der Dichte $\rho = 0.8 \cdot 2.7 \text{ g/cm}^3 = 2.16 \text{ g/cm}^3$. Somit ist zu erwarten, dass die Vorhersage der Kernladungszahl stabil den Literaturwert 13 ermittelt, während die Vorhersage der absoluten Dichte proportional zur relativen Dichte η erfolgt. Zwar ist der im soeben genannten Testphantom simulierte Dichtebereich $\eta \geq 0.8$, dennoch wurden zur Verbesserung der Generalisierungsfähigkeit auch deutlich größere Dichtebereiche zum Training einiger Modelle verwendet. Abbildung 7.13 stellt die vorhergesagten Kernladungszahlen und Dichten der Modelle, trainiert mit $\eta \geq \{0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$, dar. Zur Vereinfachung wird die folgende Notation verwendet: $\mathcal{M}^{\eta \geq 0.8}$ bezeichnet ein Modell, das auf Trainingsdaten mit einer relativen Dichte $\eta \geq 0.8$ trainiert worden ist. Die Auswertung der vorhergesagten Kernladungszahlen zeigt, dass die Modelle entsprechend ihrer relativen Dichte aus den Trainingsdaten Aluminium korrekt erkennen. Das Modell $\mathcal{M}^{\eta > 0.9}$ erkennt die Kernladungszahl des Aluminiums bei $\eta < 0.9$ nicht korrekt. Analog erkennt das Modell $\mathcal{M}^{\eta=1}$ Aluminium für $\eta \geq 0.95$. Die verbleibenden Modelle erkennen die Kernladungszahl im gesamten Phantom korrekt. In den Dichtevorhersagen zeichnet sich ein ähnliches Bild ab. Die Dichtevorhersage des Modells $\mathcal{M}^{\eta=1}$ weicht stark von den anderen Modellen ab. Es wird vermutet, dass dieses Modell das Konzept der relativen Dichte nie erlernen konnte und somit die gezeigten Ergebnisse hervorbringt. Im Gegensatz dazu extrapoliert das Modell $\mathcal{M}^{\eta > 0.9}$ korrekt die Dichte des porösen Aluminiums

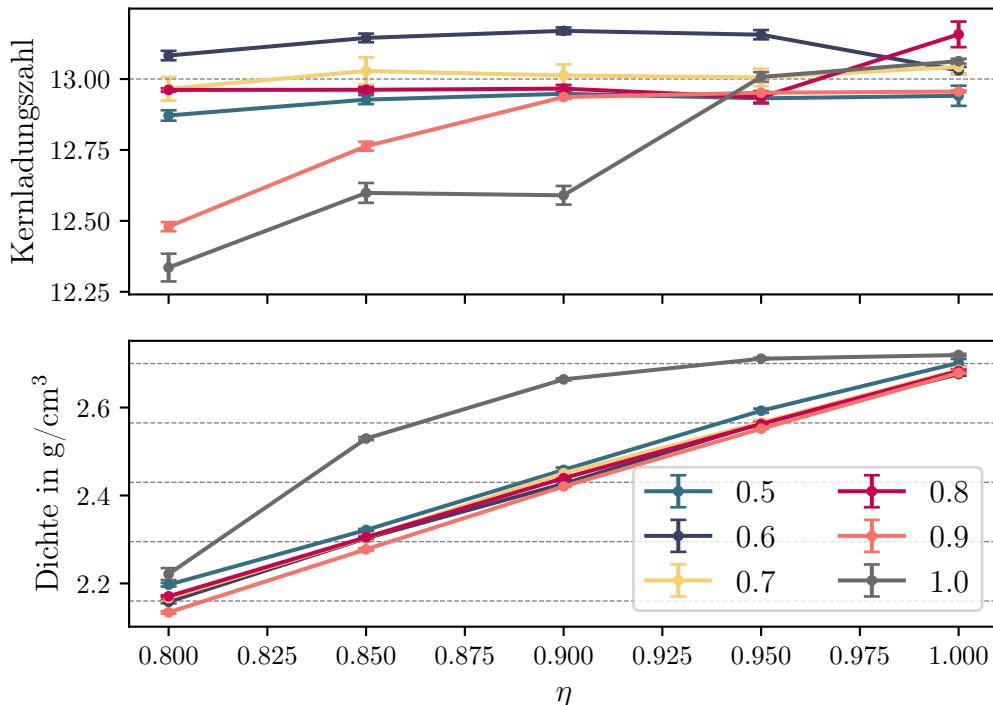


Abbildung 7.13: Vorhersagen der Kernladungszahl und Dichte für ein Aluminiumphantom mit verschiedenen relativen Dichten η (x-Achse) durch Modelle, die mit unterschiedlichen unteren Grenzen der relativen Dichte (Legende) trainiert worden sind. Die gestrichelten Linien geben die Literaturwerte an. In dem unteren Plot müssen die absoluten Dichten für $\eta \in \{0.8, 0.85, 0.9, 0.95, 1.0\}$ vorhergesagt werden, sodass fünf Literaturwerte eingezeichnet sind.

bis hin zu $\eta = 0.8$.

Es ist festzuhalten, dass die Bestimmung der Kernladungszahl poröser Materialien explizit für die in den Materialien vorliegende Porosität in den Trainingsdaten vorliegen muss, während die Dichtevorhersage, mindestens teilweise, durch die Modelle extrapoliert werden kann. Im Allgemeinen muss der Dichtebereich der Probe mindestens in den Trainingsdaten abgebildet werden, sodass die Verwendung eines Modells $M^{\eta > 0.5}$ für alle Proben, die innerhalb dieses Dichtebereiches liegen, möglich ist. Für einen gegebenen Anwendungsfall sollte demnach die relative Dichte der Probe grob abgeschätzt und in der Simulation der Trainingsdaten, beziehungsweise der Auswahl eines bereits trainierten Modells, berücksichtigt werden. Analog zu dem vorausgehenden Abschnitt ist die Anwendung des Fine-Tunings auch hier denkbar, um vortrainierte Modelle auf anwendungsspezifische, relative Dichten zu optimieren.

7.5 VORHERSAGEN AUSSERHALB DER CT-MITTELSCHICHT

Eine weitere Einschränkung besteht in der Erzeugung eines zweidimensionalen Phantoms (vergleiche DL-spezifische Anforderungen in Abschnitt 5.1) und die daraus resultierende, rekonstruierte Mittelschicht, die für das Training verwendet wird (vergleiche Kapitel 5.3). Demnach liegen die simulierten CT-Schichtbilder in der Mittelebene (Fächerstrahl-CT), sodass keine Feldkamp-Artefakte (vergleiche Kapitel 2.4.1) auftreten. Anschaulich be-

schrieben kann ein dreidimensionales, rekonstruiertes Volumen als CD-Stapel⁷ dargestellt werden, der entlang der CT-Drehachse verläuft – sozusagen an ebendieser aufgefädelt ist. Die mittlere CD im Stapel symbolisiert in diesem Vergleich die CT-Mittelschicht. In realen CT-Scans werden hingegen die allermeisten Proben mit der dreidimensionalen Kegelstrahl-CT gescannt und Schnittebenen, die entlang der Drehachse versetzt zur Mittelebene liegen, enthalten wichtige Informationen und müssen, auch in der Materialbestimmung, korrekt abgebildet werden. Ein bekanntes Problem der Kegelstrahl-CT kann mittels einer Simulation in *aRTist* visualisiert werden: Die rekonstruierten Abschwächungskoeffizienten μ ändern sich mit dem Abstand zur Mittelebene wie in Abbildung 7.14 dargestellt wird. Sie sinken mit zunehmendem Abstand zur Mittelebene für beide

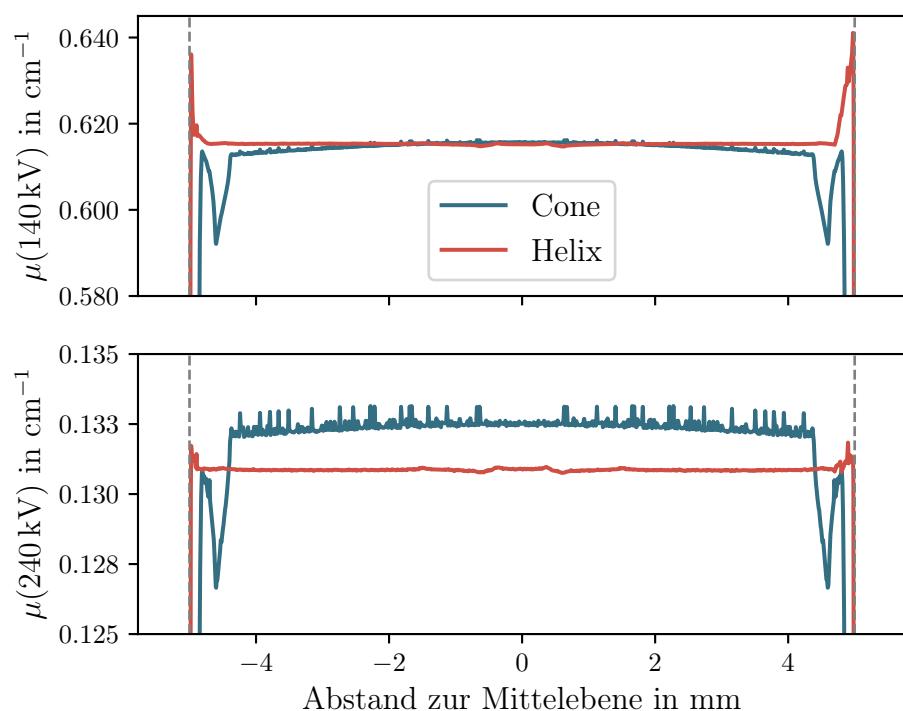


Abbildung 7.14: Schnitt entlang der Drehachse (z-Achse) durch das rekonstruierte Volumen eines Eisenstabs der Länge 10 cm für die DECT-Spektren nach Tabelle 5.2: 140 kV und 240 kV. Gezeigt werden die rekonstruierten Abschwächungskoeffizienten μ , die aus einem simulierten Scan mit einer Kegelstrahl-CT (Cone) und einer Helix-CT (Helix) aufgenommen worden sind. Die physischen Start- und Endpunkte des Eisenstabes sind grau gestrichelt eingezeichnet.

Spektren bei der Kegelstrahl-CT (blaue Kurve) ab. Zusätzlich ist an den Rändern zu erkennen, dass die Kegelstrahl-CT den Eisenstab etwa 0.5 mm zu kurz rekonstruiert. Mithilfe der ebenfalls in Abbildung 7.14 dargestellten Helix-CT (rote Kurve) werden die beiden genannten Probleme der Kegelstrahl-CT weitestgehend behoben. Das Profil erscheint deutlich konstanter und die Endpunkte des Eisenstabes werden korrekt, wenn auch mit überschwingenden Abschwächungskoeffizienten, rekonstruiert. Allein anhand dieser Erkenntnis aus Abbildung 7.14 kann gefolgert werden, dass die Helix-CT für die Materialbestimmung vorzuziehen ist, da sie weniger Artefakte im Volumen, insbesondere in den Randbereichen, erzeugt und somit die Verschiebung der Datenverteilung im

⁷ Gemeint ist ein Stapel sogenannter *Compact Discs*.

Vergleich zur simulierten Mittelschicht im Trainingsdatensatz verringert wird.

Aufbauend auf dieser Erkenntnis wird in Abbildung 7.15 die Materialbestimmung konkret auf einem dreidimensionalen, simulierten Volumen durchgeführt und in Abhängigkeit des Abstandes zur Mittelebene entlang der CT-Drehachse ausgewertet. Die Vorher-

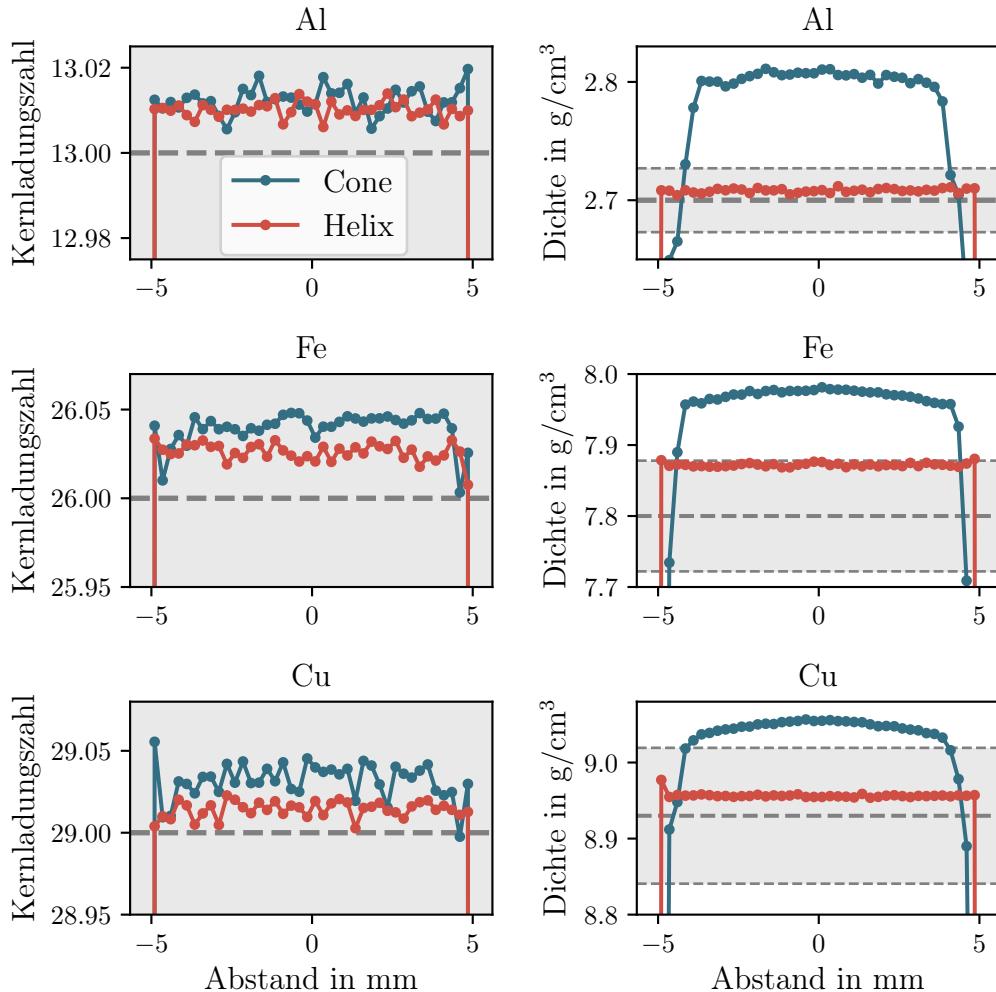


Abbildung 7.15: Vorhergesagte Kernladungszahlen und Dichten abhängig vom Abstand zur Mittelebene eines Phantoms angelehnt an Abbildung 6.4 verlängert entlang der Drehachse auf 10 cm. Die Literaturwerte sind grau dick gestrichelt eingezeichnet. Die in Kapitel 6.1 definierten Präzisionen von ± 0.5 für die Kernladungszahl sowie $\pm 1\%$ für die Dichte sind dünn gestrichelt als obere und untere Grenze des zulässigen Bereiches, grau hinterlegt, eingezeichnet.

sage der Kernladungszahl gelingt dem trainierten Modell sowohl bei der Kegelstrahl-CT als auch bei der Helix-CT im Toleranzbereich der in Kapitel 6.1 abgeleiteten Rahmenbedingungen problemlos. Zudem ist im Profil der Vorhersage der Kernladungszahl im Volumen aus der Kegelstrahl-CT (linke Spalte, blaue Linie) keine Abstandsabhängigkeit zu beobachten. Bei der Vorhersage der Dichte auf dem Volumen der Kegelstrahl-CT liegen alle Vorhersagen außerhalb des grau markierten Toleranzbereiches (vergleiche Abschnitt 6.1). Die reduzierten Abschwächungskoeffizienten im High-Energy-Volumen der Helix-CT (vergleiche Abbildung 7.14; unten) senken die Vorhersage der Dichte ab, sodass die vorhergesagten Dichten innerhalb der Toleranz liegen. Weiterhin ist zu ver-

merken, dass sich bei der Vorhersage der Dichte im Volumen der Kegelstrahl-CT das bereits in Abbildung 7.14 beobachtete Profil beobachten lässt, welches mit zunehmendem Abstand zur Mittelebene abfällt. Das Modell interpretiert den genannten Verlauf der Abschwächungskoeffizienten als Variation der relativen Dichte, da die Vorhersage der Kernladungszahlen konstant ist. Im Einklang mit der Erkenntnis aus Abbildung 7.14 wird festgehalten, dass die Helix-CT der Kegelstrahl-CT für die Materialbestimmung dreidimensional ausgedehnter Proben vorzuziehen ist. Die in Kapitel 5.1 gewählte Methodik – das exklusive Training auf CT-Mittelschichten – stellt keine Einschränkung für die Anwendung der Materialbestimmung auf dreidimensionale Proben dar, sofern die Helix-CT zur Bildaufnahme verwendet wird.

7.6 ANWENDUNG AUF REALE CT-DATEN

Abschließend wird der entwickelte Ansatz zur Materialbestimmung auf realen CT-Scandaten erprobt. Im Rahmen der Evaluation der implementierten Simulation in Abschnitt 5.4.2 wurde bereits gezeigt, dass die reale CT-Anlage, im Vergleich zur implementierten Simulation, geringfügig abweichende Abschwächungskoeffizienten misst. Somit besteht die Vermutung, dass die Materialvorhersage ebenfalls keine idealen Ergebnisse liefern wird. Abbildung 7.16 (linke Spalte) zeigt die rekonstruierten CT-Mittelschichten eines DECT-Scans an der CT-Anlage aus Abschnitt 2.7 mit den Scanparametern aus Tabelle 5.2. In der rechten Spalte werden die Materialvorhersagen eines DL-Modells dargestellt. Zunächst ist zu beobachten, dass die CT-Artefakte zwischen den einzelnen Proben teilweise, und fälschlicherweise, als Material mit einer niedrigen Kernladungszahl und Dichte erkannt werden. Weiterhin ist, insbesondere in der Dichtevorhersage des Kupfers (rechts unten), zu erkennen, dass die Strahlaufhärtungsartefakte (das Cupping, vergleiche Kapitel 2.4.1) nicht vollständig vom DL-Modell als solche erkannt und entfernt werden. Der homogene Kupferquader (im Schnittbild: Kupferquadrat) enthält somit nach der Vorhersage ein falsches Dichteprofil. Ein möglicher Grund für diese Abweichung besteht in der Sim-To-Real-Gap, die bereits in Kapitel 1.1 angesprochen wurde und auf eine unzureichende Modellierung des realen CT-Messprozesses durch die Simulation zurückzuführen ist. Wie in Kapitel 5.2 beschrieben, bildet eine Raytracing-basierte Simulation den realen Messprozess nicht ideal nach, sodass systematische Modellierungsfehler zwangsläufig in die Ergebnisse eingetragen werden.

Neben der qualitativen Auswertung von Abbildung 7.16 können die Materialvorhersagen quantitativ anhand von Tabelle 7.1 bewertet werden. Die vorhergesagten Kernladungszahlen für Aluminium, Edelstahl und Kupfer stimmen unter Einbezug der Unsicherheiten mit den Literaturwerten überein. Dennoch befindet sich der Mittelwert der Kernladungszahl für Edelstahl mit 26.73 außerhalb des tolerierten Fehlerbereiches ± 0.5 (vergleiche Abschnitt 6.1). Bei den Dichtevorhersagen liegen alle Mittelwerte deutlich unterhalb der Literaturwerte. Der Grund für diese einheitliche Absenkung der vorhergesagten Dichten kann, mindestens teilweise, in der fehlenden Modellierung der Streustrahlung gefunden werden. Abbildung 7.17 zeigt das CT-Schichtbild einer Raytracing-basierten Simulation und das Differenzbild zu einer Monte-Carlo-basierten Simulation. Die Streustrahlung

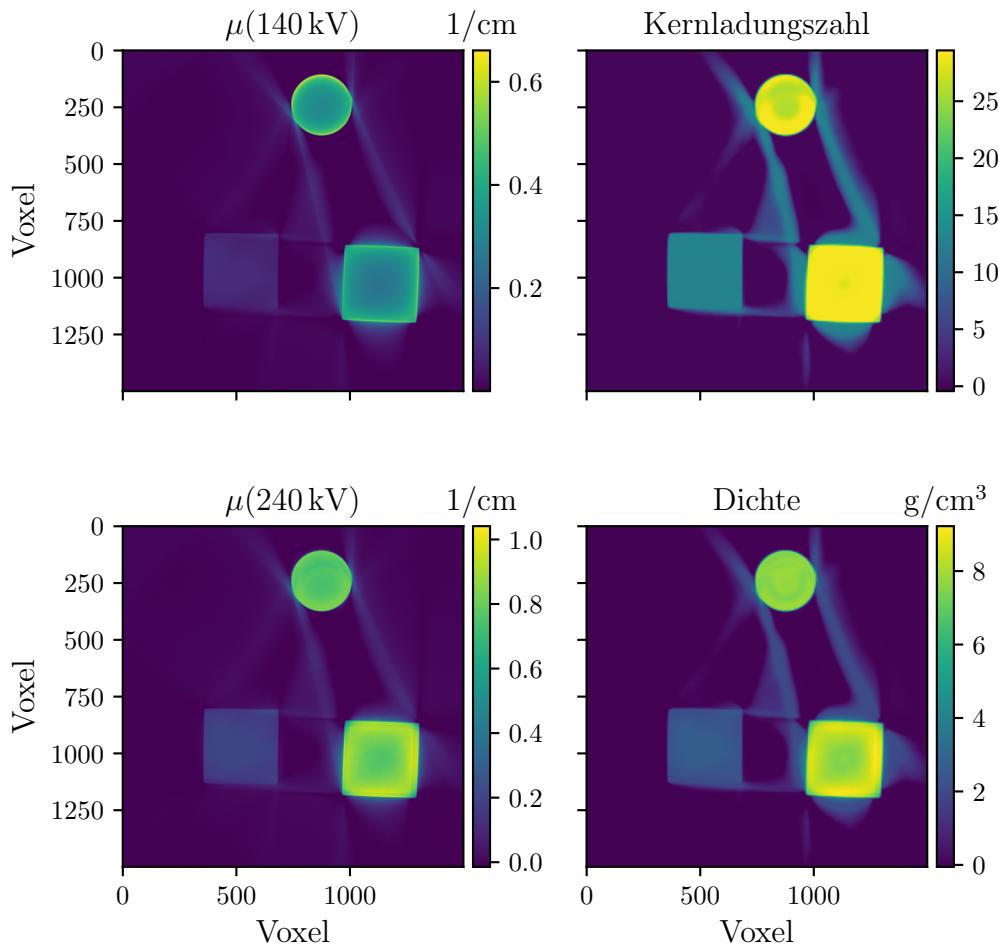


Abbildung 7.16: Rekonstruierte DECT-Schichtbilder (links) und berechnete Materialkarten (rechts) eines realen Scans an der CT-Anlage aus Kapitel 2.7. Alle x- und y-Achsen tragen die Einheit Voxel.

trägt aufgrund ihrer Winkelverteilung ebenfalls zum gemessenen Detektorsignal bei, sodass insgesamt eine niedrigere Abschwächung rekonstruiert wird (vergleiche Abschnitt 2.1 und 2.3). Nach der Auswertung in Abbildung 7.17 senkt die Berücksichtigung der Streustrahlung die rekonstruierten Abschwächungskoeffizienten um etwa 10 % für das genannte Beispiel (Aluminiumwürfel, 240 kV) gegenüber der Raytracing-basierten Simulation ab. Daraus folgt, dass auch die Dichte niedriger vorhergesagt wird – die beobachtete Abweichung in Tabelle 7.1.

Die einzelnen Bestandteile der Probe aus Abbildung 7.16 sind räumlich klar voneinander separiert und in ihrem Inneren nicht strukturiert, sodass die Vorhersage der Materialien verhältnismäßig einfach erscheint. Eine komplexer strukturierte Probe wurde mit den oben genannten CT-Parametern aus Tabelle 5.2 gescannt und ist in Abbildung 7.18 dargestellt – es handelt sich um eine Lithium-Ionen-Batterie (LIB). Die gezeigte LIB ist eine sogenannte *Rundzelle*, die aus aufgewickelten Lagen besteht. In der Dichtevorhersage in Abbildung 7.18 (rechts unten) ist zu erkennen, dass die LIB ein Gehäuse mit einer hohen Dichte ($\approx 6 \text{ g/cm}^3$) besitzt und in ihrem Inneren Dichten zwischen 1 g/cm^3 und

Tabelle 7.1: Vorhergesagte Kernladungszahlen und Dichten durch ein U-Net [36] für das Experiment aus Abbildung 7.16. Die tabellierten Werte geben den Mittelwert μ , die Standardabweichung σ und die prozentuale Abweichung des Mittelwerts p vom Literaturwert der Vorhersage in der Form $\mu \pm \sigma$ (p) an. In der ersten Spalte werden die Materialien zusammen mit ihren Literaturwerten gelistet.

Material	Kernladungszahl Z	Dichte ρ in g/cm^3
Al ($Z = 13, \rho = 2.70$)	13.00 ± 0.01 (0.0 %)	2.59 ± 0.10 (4.1 %)
V2A ($Z = 25.9, \rho = 7.77$)	26.73 ± 0.90 (3.2 %)	7.73 ± 0.12 (0.5 %)
Cu ($Z = 29, \rho = 8.93$)	28.95 ± 0.19 (0.2 %)	7.92 ± 0.37 (11.3 %)

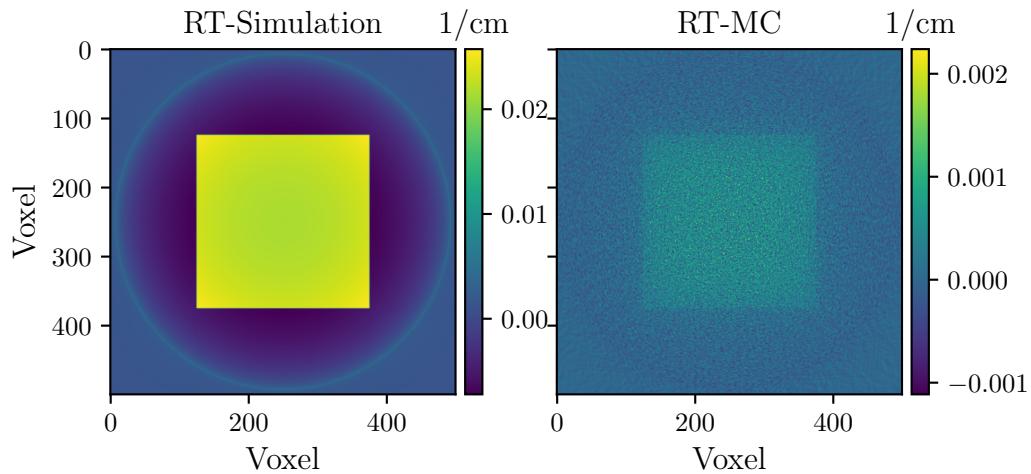


Abbildung 7.17: Links: Simuliertes CT-Schichtbild eines Aluminiumwürfels bei einer Quellspannung $U_B = 240 \text{ kV}$ mit der Verwendung eines Raytracing-basierten Projektors (RT). Rechts: Differenzbild zu dem simulierten CT-Schichtbild eines Monte-Carlo-basierten Projektors (MC). Der Mittelwert innerhalb des Aluminiums im Differenzbild beträgt 0.002 cm^{-1} – in guter Näherung 10 % des insgesamt gemessenen Signals. Die Berechnungsdauern der Bilder betragen in der genannten Reihenfolge $T_{\text{RT}} = 1.2 \text{ min}$ und $T_{\text{MC}} \approx 120 \text{ h}$.

$4 \text{ g}/\text{cm}^3$ vorliegen. Allgemein ähneln die Kontrastverhältnisse der Dichtevorhersage den DECT-Schichtbildern in Abbildung 7.18 (links). Auffällig ist bezüglich der Vorhersage der Kernladungszahl in Abbildung 7.18 (rechts oben), dass einige Lagen im Inneren der LIB eine hohe Kernladungszahl und gleichzeitig niedrige Dichte aufweisen. Hierbei handelt es sich um ein Beschichtungsmaterial aus einer Gruppe namens Lithium-Nickel-Mangan-Cobalt-Oxid (NMC). Eine häufig verwendete Mischung mit einer Dichte von $3.4 \text{ g}/\text{cm}^3$ heißt NMC811, die im rekonstruierten Volumen nach der Gleichung von Heisemann [14] mit einer effektiven Kernladungszahl $Z_{\text{eff}} \approx 26.4$ erscheint. Da die genaue Zusammensetzung der meisten kommerziell erhältlichen LIBs nicht öffentlich bekannt ist, wird der gezeigte Scan zusammen mit der Materialbestimmung der LIB nur qualitativ ausgewertet. Anhand des Scans wird dennoch deutlich, dass die vorgestellte Methode zur Materialbestimmung prinzipiell für die Untersuchung von LIBs geeignet ist. Unklar ist zu diesem Zeitpunkt, ob typische Fehlerbilder, nach denen in den CT-Scans von Batterien

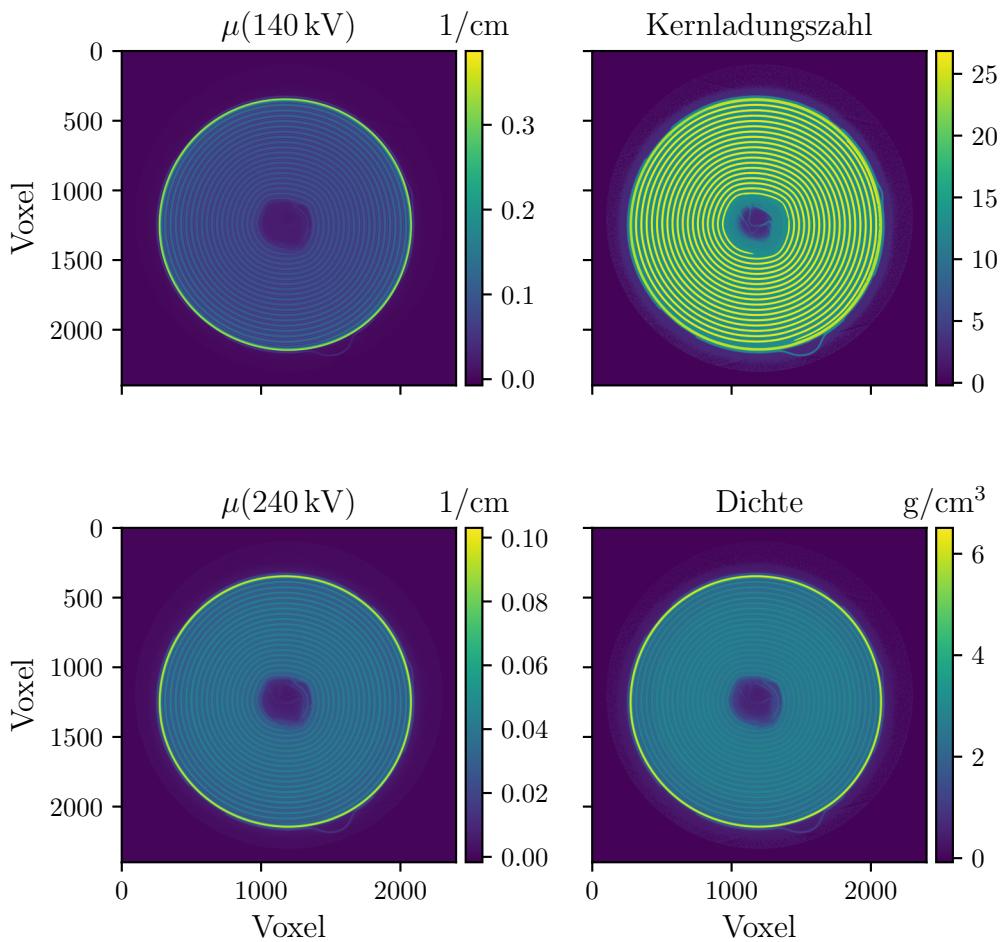


Abbildung 7.18: Rekonstruierte DECT-Schichtbilder (links) und berechnete Materialkarten (rechts) eines CT-Scans einer Lithium-Ionen-Batterie (LIB), durchgeführt an der CT-Anlage aus Kapitel 2.7.

gesucht wird⁸, mithilfe der vorgestellten Methode gefunden werden können. Mit einer gezielten Auswertung einer bekannten und vollständig charakterisierten LIB kann das vorgestellte Verfahren zur Materialbestimmung aufbauend auf dieser Arbeit evaluiert und gegebenenfalls entsprechen qualifiziert werden.

Zusammenfassend kann festgehalten werden, dass die Übertragung der trainierten Modelle, für quantitative Aussagen, auf reale CT-Daten nicht trivial ist. Die Datenverteilungen der realen CT-Anlage unterscheiden sich zu stark von den simulierten Verteilungen, sodass weitere Schritte notwendig sind, um die Materialvorhersagen zu verbessern. Eine weitere Optimierung der Simulation zur Reduktion der Sim-To-Real-Gap erscheint wenig aussichtsreich, da die angesprochenen, systematischen Modellierungsfehler einer Raytracing-basierten Simulation nicht umgehbar sind ohne Monte-Carlo-basierte Verfahren zu nutzen. Hingegen erscheint das in Abschnitt 7.3.1 erprobte Fine-Tuning

⁸ Zu den wichtigsten Untersuchungen zählt die Identifikation freier Partikel, beispielsweise aus Kupfer, die als Verunreinigung einen negativen Einfluss auf die Leistung und Sicherheit der LIB ausüben können [116, 117]. Nach der Untersuchung des Auflösungsvermögens in Abschnitt 7.1 ist klar, dass das Auflösungsvermögen der Materialbestimmung nicht nur die Sichtbarkeit, sondern ebenso die quantitativen Materialvorhersagen kleiner Partikel beeinflusst.

als vielversprechende Methoden zur Überwindung der Sim-To-Real-Gap. Nach dem erfolgreichen Fine-Tuning eines Modells zur Anpassung an eine veränderte CT-Geometrie (vergleiche Abschnitt 7.3.1), könnte das gleiche Verfahren für die Anpassung an reale CT-Daten verwendet werden. Durch das Vortrainieren der Modelle genügt ein deutlich kleinerer Datensatz zum Fine-Tuning, der anhand eines geeigneten, realen Phantoms erzeugt werden kann (Kalibrationsscan). Ein Nachteil dieses Verfahrens besteht jedoch darin, dass das Fine-Tuning, analog zu Abschnitt 7.3.1, für jeden CT-Parametersatz erneut anhand eines Kalibrationsscans durchgeführt werden muss. Es besteht derzeit keine Möglichkeit, den Einfluss der CT-Geometrie oder des Quellspektrums getrennt von der Abweichung zwischen der Simulation und der realen CT-Anlage zu beschreiben, da sowohl die CT-Geometrie als auch das Quellspektrum die simulierten und real gemessenen Datenverteilungen in komplexer Form beeinflussen. Idealerweise bestünde die DL-Architektur aus getrennten und voneinander unabhängigen Modulen, die jeweils einen dieser Freiheitsgrade abdecken. Somit wäre das Fine-Tuning, beispielsweise an eine neue CT-Geometrie, möglich, ohne für jede Anwendung einen Kalibrationsscan durchführen zu müssen.

Ein alternativer Ansatz kann durch die Verwendung eines zweiten DL-Modells realisiert werden, das, abhängig von den CT-Parametern, explizit den Unterschied zwischen der Datenverteilung der Raytracing-basierten Simulation und einem realen Scan (oder einer Monte-Carlo-basierten Simulation) erlernt und ebendiesen effektiv kompensieren kann. Dieser Ansatz wird in Kapitel 9 im Kontext zukünftiger Forschungsarbeiten ausführlich diskutiert.

7.7 ZUSAMMENFASSUNG DER ERGEBNISSE

In diesem Kapitel wurde die entwickelte Methodik, bestehend aus der implementierten Simulation (Kapitel 5) sowie der ausgewählten DL-Architektur (Kapitel 6), in realistische CT-Scanszenarien überführt und ihre Einschränkungen bezogen auf die Präzision der Materialbestimmung untersucht. Die meisten untersuchten Einschränkungen treten unabhängig von der beobachteten Sim-To-Real-Gap auf, sodass sie anhand von simulierten Phantomen untersucht werden konnten. In Abschnitt 7.1 wurde gezeigt, dass der Erwartungswert der Regionsgröße, bestimmt durch den Parameter N_{max} , in den simulierten Phantomen (vergleiche Phantomgenerator in Kapitel 5.3) keinen Einfluss auf die Ortsauflösung der DL-basierten Materialbestimmung hat. Weiterhin wurde erkannt, dass die MTF-Kurve, eine häufig genutzte Darstellungsform für die Ortsauflösung von CT-Systemen, nicht zur Quantifizierung der Ortsauflösung bei der Materialbestimmung geeignet ist, sodass diesbezüglich eine domänenpezifische Methodik entwickelt werden sollte. Die Analyse der Kantenschärfe der Materialvorhersagen in Abschnitt 7.2 bestätigt die Ergebnisse von Steuwe et al. [110], dass CNNs, die mit dem MSE-Loss auf CT-Schichtbildern trainiert werden, die Kantenschärfe der Bilder reduzieren.

In Abschnitt 7.3 konnte gezeigt werden, dass das Training eines Modells auf einer festen Kantenlänge des simulierten Volumens die Generalisierungsfähigkeit des Modells auf andere Kantenlängen einschränkt. Das Fine-Tuning eines vortrainierten Modells auf eine

neue Kantenlänge, die in einer vergleichbaren Größenordnung der ursprünglichen Kantenlänge liegt, wurde als Lösungsvorschlag aufgezeigt und erfolgreich für den Transfer eines Modells $\mathcal{M}^{4\text{cm}}$ (Kantenlänge im Training 4 cm) auf die neue Kantenlänge 1 cm anhand eines kleinen Datensatzes demonstriert.

Neben der Materialbestimmung massiver Proben konnten in Abschnitt 7.4 die Kernladungszahlen und Dichten von porösem Aluminium (Porosität $\leq 20\%$) mit Modellen durchgeführt werden, die auf entsprechenden Porositäten, beziehungsweise relativen Dichten, innerhalb der Trainingsdaten konditioniert wurden.

Eine weitere relevante Einschränkung der entwickelten Methode, das Training auf zweidimensionalen Phantomen (Fächerstrahl-CT), konnte im Rahmen von Abschnitt 7.5 bei der Anwendung auf dreidimensionale Proben analysiert werden. Durch den Einsatz der Helix-CT, als Alternative zur Kegelstrahl-CT, werden aufgrund der Verschiebung entlang der CT-Achse alle Schichten des dreidimensionalen Volumens wie eine Mittelschicht betrachtet, sodass ein auf simulierten Mittelschichten trainiertes Modell stabile Materialvorhersagen über ein dreidimensionales Volumen liefert.

Die sicherlich gravierendste Einschränkung ist durch die mehrfach angesprochene Sim-To-Real-Gap gegeben. In Abschnitt 7.6 wurde die Materialbestimmung auf einem realen CT-Schichtbild erprobt, wobei die Ergebnisse entsprechend der ausgeprägten Sim-To-Real-Gap ausfallen. Es ist festzuhalten, dass die vorgestellte Methode zur Materialbestimmung nur stark eingeschränkt direkt in die CT-Praxis übertragbar ist. Weiterführende Schritte wurden in Abschnitt 7.6 diskutiert, um die Sim-To-Real-Gap zu reduzieren und so den Transfer in die reale Anwendung zu ermöglichen. Das schlichte Training auf annotierten Realdaten, das teilweise in der verwandten Forschung in Kapitel 4.2.4 durchgeführt wurde, ist für industrielle Anwendungen aufgrund der enormen Vielfalt bezogen auf die Prüfteile (Form und Material) und die CT-Parameter für ein gut generalisiertes Modell nicht durchführbar. Prinzipiell wird ein Ansatz benötigt, der das Grundkonzept der Materialbestimmung aus simulierten Daten erlernt und lediglich den Versatz zur realen Datenverteilung in den CT-Scans anhand weniger, im Bestfall nicht annotierter, Daten erlernen kann. Die Konzeption und Entwicklung eines derartigen Ansatzes ist Gegenstand zukünftiger Forschung.

8

ANWENDUNGSPOTENTIAL AUF ANDERE PROBLEMSTELLUNGEN IN DER CT

Die entwickelte Methodik zur Erzeugung von Trainingsdaten für die Materialbestimmung in CT-Schichtbildern sowie die implementierten DL-Architekturen können, in leicht abgewandelter Form, neben der Materialbestimmung für weitere Anwendungen in der CT verwendet werden und somit zur Verbesserung der CT-Datenverarbeitung im Allgemeinen abseits der Materialbestimmung beitragen. Ziel dieses Kapitels ist die exemplarische Ausführung oder konzeptionelle Identifikation einiger Anwendungsmöglichkeiten, die ohne komplexe Weiterentwicklungen der, in den drei vorausgehenden Forschungsfragen erarbeiteten, Methodik umgesetzt werden können. Mit der beobachteten Sim-To-Real-Gap wurde ein zentrales Problem der Trainingsdatenerzeugung identifiziert, das zu einer deutlichen Verschlechterung der Materialbestimmung in realen CT-Scans führt, wie in Kapitel 7.6 aufgezeigt wurde. Dennoch besteht ein zu identifizierendes Potential für andere Anwendungsfälle, in denen die Sim-To-Real-Gap, konkret die fehlende Modellierung der Streustrahlung, eine untergeordnete Rolle spielt und somit die entwickelte Methodik direkt verwendet werden kann. Die Einflüsse der Sim-To-Real-Gap werden daher in den jeweiligen Abschnitten anwendungsbezogen diskutiert.

Für die meisten Prüfaufgaben, die mithilfe der CT durchgeführt werden, genügt die Messung mit einem einzelnen Spektrum, sodass DECT im Gesamtfeld der CT-Prüfung in einer Nische zu lokalisieren ist. Daher basieren die in diesem Kapitel vorgestellten Anwendungen nicht auf DECT, sodass sie per Definition das Potential besitzen, auf jeden industriellen CT-Scan angewandt zu werden.

Aufbauend auf den Beobachtungen aus Kapitel 7.6 (vergleiche Abbildung 7.16) wird zunächst ein Verfahren zur Reduktion der Strahlaufhärtungsartefakte in Abschnitt 8.1 vorgestellt und auf reale CT-Daten angewandt. Abschnitt 8.2 untersucht die implizit erlernte Rauschunterdrückung des Verfahrens aus Abschnitt 8.1 ebenfalls auf realen CT-Daten. Abschließend wird in Abschnitt 8.3 eine Methodik konzeptionell erläutert, die zur Erkennung von Poren und Partikeln in CT-Scans verwendet werden kann.

8.1 REDUKTION DER STRAHLAUFHÄRTUNGSArTEFAKTE

Die Reduktion der Strahlaufhärtungsartefakte wird seit vielen Jahren untersucht, um die Bildqualität in CT-Scans zu verbessern. In klinischen CT-Scans stellt beispielsweise implantiertes Metall in Form von Schrauben oder Schienen eine Hauptquelle für die sogenannten *Metallartefakte* dar, die als Extremfall der Strahlaufhärtung definiert werden können. Nach Selles et al. lassen sich Verfahren zur Reduktion der Metallartefakte anhand der verwendeten Datendomäne in zwei Kategorien unterteilen: *Sinogram-based* und *Image-based* [118]. Im Rahmen dieser Dissertationsschrift wurde die Materialbestim-

mung auf CT-Schichtbildern (*Image-based*) untersucht, sodass die in den vorangehenden Kapiteln entwickelte Methodik vorzugsweise auf die zuletzt genannte Datendomäne, das heißt im rekonstruierten Volumen, angewandt wird. Weiterhin wurde in Kapitel 7.6 beobachtet, dass die trainierten Modelle die Strahlauflaufhärtung in den Trainingsdaten verarbeiten und in den Materialvorhersagen teilweise kompensieren (vergleiche Abbildung 7.16). Abbildung 8.1 stellt die Kompensation der Strahlauflaufhärtung exemplarisch und qualitativ anhand der Materialbestimmung innerhalb der Materialien Edelstahl und Kupfer aus dem realen CT-Scan nach Abbildung 7.16 dar. Aluminium weist aufgrund

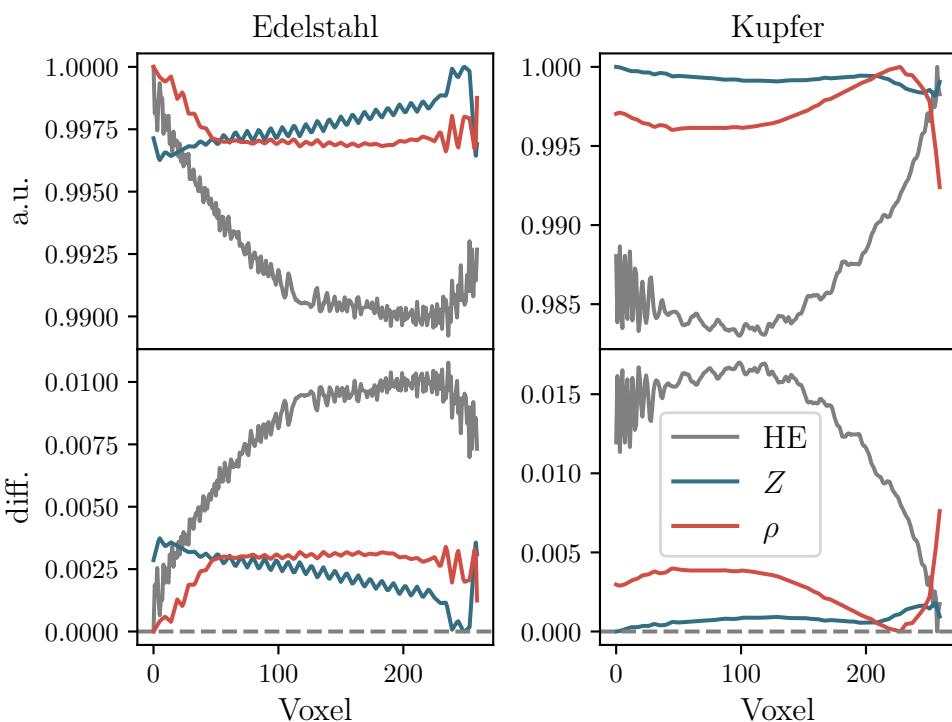


Abbildung 8.1: Teilschnitte durch den Edelstahlstab (links) und den Kupferwürfel (rechts) im High-Energy-Schichtbild mit 240 kV (graue Linie) sowie durch die Materialvorhersagen (blaue und rote Linie). Gezeigt wird jeweils ein Teilprofil (obere Zeile) durch die Objekte zur normierten Darstellung des Cuppings, hauptsächlich bedingt durch die Strahlauflaufhärtung, sodass die y-Achse einheitenlos ist. In der unteren Zeile wird die absolute Abweichung von einem idealen Objekt (gestrichelte Linie), das heißt ohne Strahlauflaufhärtungsartefakte, dargestellt. Ziel der Abbildung ist die qualitative Veranschaulichung, dass die Materialvorhersagen weniger Cupping aufweisen als das High-Energy-Schichtbild.

seiner niedrigen Abschwächung grundsätzlich weniger Strahlauflaufhärtungsartefakte auf, sodass die Materialien Edelstahl und Kupfer zu diesem Zweck anschaulicher sind. Das Vergleichprofil wurde dem High-Energy-Schichtbild entnommen, da es weniger Strahlauflaufhärtungsartefakte aufweist als das Low-Energy-Schichtbild und die Verwendung des letzteren somit ein unfairener Vergleich wäre.¹

Wie einleitend beschrieben, ist die Durchführung der DECT beziehungsweise der Materialbestimmung für viele CT-Prüfaufgaben mit der Verwendung eines EIDs zu zeitauf-

¹ Das trainierte Modell berechnet die Materialbestimmung basierend auf den Low- und High-Energy-Schichtbildern. Somit ist die qualitative Beurteilung der Artefaktkompensation nur fair, wenn das artefaktärmer Eingabebild, das High-Energy-Schichtbild, zum Vergleich herangezogen wird.

wändig und es bedarf einer vereinfachten Methode zur Reduktion der Strahlaufhärtingsartefakte, die ohne eine DECT auf bestehende CT-Scanprotokolle angewandt werden kann. Mithilfe der implementierten Simulation lässt sich eine einfachere und zugleich effektivere Methode zur Kompensation der Strahlaufhärtingsartefakte realisieren: Wie in Kapitel 2.4.1 beschrieben, besteht der Grund für die Strahlaufhärtingsartefakte in der Nutzung eines polychromatischen Röntgenspektrums. Bisher wurde die Simulation mit zwei verschiedenen Quellspektren verwendet, um die DECT-Scans darzustellen. Durch Verwendung zweier Spektren, beispielsweise 140 kV (polychromatisch) und 100 keV (monochromatisch), berechnet die Simulation automatisch ein rekonstruiertes Volumen mit dem polychromatischen Spektrum und ein rekonstruiertes Volumen mit dem monochromatischen Spektrum, das keine Strahlaufhärtingsartefakte aufweist. Mithilfe der simulierten Daten kann somit ein neuronales Netz, in diesem Fall exemplarisch ein U-Net [36], trainiert werden, um die polychromatischen CT-Schichtbilder in monochromatische CT-Schichtbilder umzurechnen. Abbildung 8.2 stellt die Vorhersage eines nach dieser Methodik trainierten U-Nets anhand von realen CT-Daten dar. Anhand der CT-

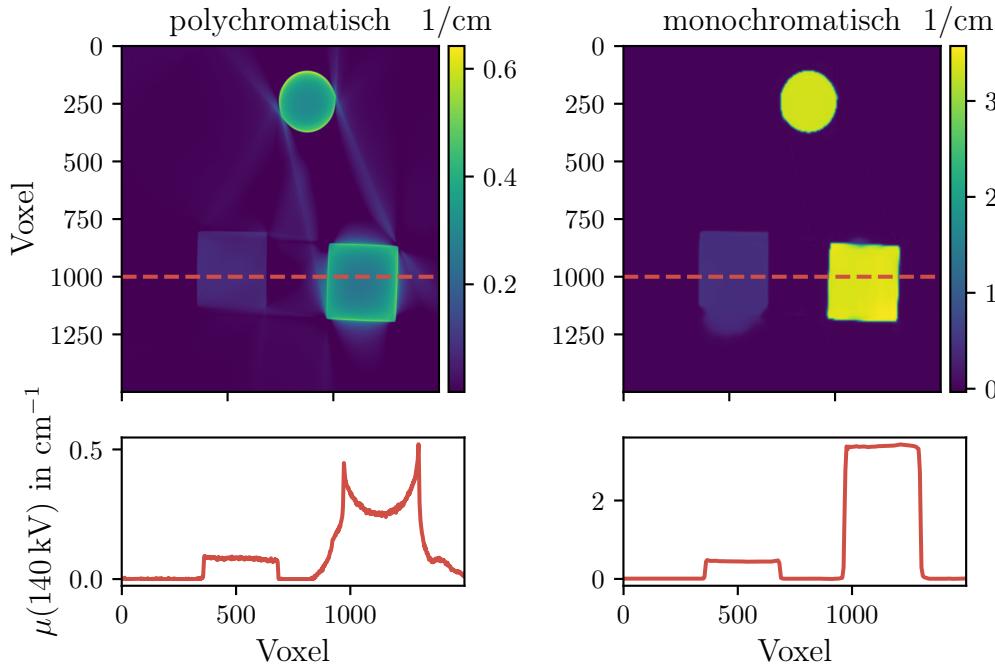


Abbildung 8.2: Polychromatisches CT-Schichtbild (links) mit einer Beschleunigungsspannung $U_B = 140 \text{ kV}$ des CT-Scans aus Kapitel 7.6 und Vorhersage des monochromatischen CT-Schichtbildes durch ein U-Net [36]. Die untere Zeile zeigt die Profile entlang der roten, gestrichelten Linie.

Schichtbilder wird deutlich, dass die Strahlaufhärtingsartefakte deutlich unterdrückt werden (insbesondere untere Zeile anhand des reduzierten Cuppings). Die Außenkante des Aluminiumquaders wird allerdings nicht korrekt dargestellt – im unteren Teil wird fälschlicherweise Material vorhergesagt, obwohl an dieser Stelle lediglich Luft erscheinen sollte. Dieses Beispiel soll exemplarisch zeigen, dass die Realisierung einer Artefaktkorrektur für die Strahlaufhärtingsartefakte mit diesem Ansatz möglich ist.

Der Einfluss der Sim-To-Real-Gap, hauptsächlich induziert durch die fehlende Modellie-

rung der Streustrahlung, wird anhand dieses Beispiels ebenfalls in verminderter Form deutlich: Formell werden durch das beschriebene Verfahren die polychromatischen Abschwächungskoeffizienten aus dem realen CT-Scan in monochromatische Abschwächungskoeffizienten umgerechnet. Durch die fehlende Streustrahlung in den simulierten Trainingsdaten werden die durch das DL-Modell berechneten, monochromatischen Abschwächungskoeffizienten verfälscht, da die Inputs des realen Scans der Streustrahlung, und damit einer leicht abweichenden Verteilung, unterliegen. Wie einleitend beschrieben, ist die Reduktion der Strahlaufhärtingsartefakte nicht rein quantitativ motiviert, sondern dient ebenfalls einer Verbesserung der (subjektiven) Bildqualität. Bei der strukturellen CT-Prüfung, beispielsweise zur Identifikation von Poren und Rissen, sind die absoluten Werte der rekonstruierten Abschwächungskoeffizienten für den Anwender irrelevant. Entscheidend ist einzig der (relative) Bildkontrast, sodass die Anwendung der in diesem Abschnitt beschriebenen Methode keinen negativen Einfluss auf die Durchführung der Prüfaufgabe hat.

8.2 RAUSCHREDUKTION

In diesem Abschnitt wird das trainierte Modell aus dem vorausgehenden Abschnitt 8.1 auf reale CT-Daten angewandt und das Rauschen des Inputs (CT-Schichtbild) mit dem Rauschen der Prediction (monochromatisches CT-Schichtbild) verglichen. Neben DL-basierten Verfahren, die explizit zur Rauschunterdrückung in CT-Bildern entwickelt worden sind [119], besteht die Vermutung, dass das trainierte Modell aus Abschnitt 8.1 bereits aufgrund der Trainingsstrategie ein gewisses Potential zur Rauschreduktion aufweist. Das Modell wurde auf rauschfreien CT-Schichtbildern trainiert, sodass es kein realistisches Verständnis des CT-spezifischen Rauschmusters im Rahmen des Trainings erlangen konnte. Wie in Kapitel 3.4.1 beschrieben, enthält das verwendete U-Net Dropout-Layer, die, einzig motiviert aus der Perspektive des DLs, zur Stabilisation des Trainings genutzt werden [45]. Die Dropout-Layer können formell als binäres (Bild-)Rauschen interpretiert werden, das vom Modell unterdrückt wird.

Zur Quantifizierung des Rauschens in einem CT-Bildausschnitt wird die Standardabweichung in ebendiesem nach dem in Kapitel 6.1 (Abbildung 6.2) beschriebenen Verfahren ermittelt. Abbildung 8.3 stellt ein real gemessenes, polychromatisches CT-Schichtbild sowie das monochromatische CT-Schichtbild durch die Vorhersage eines U-Nets trainiert im Rahmen von Abschnitt 8.1 dar. Anhand der roten Boxen ist zu erkennen, dass das Bildrauschen durch Anwendung des U-Nets, sowohl innerhalb der Proben als auch im leeren Raum², reduziert wird. Auch wenn diese stichprobenartige Untersuchung der Rauschreduktion keineswegs eine Allgemeingültigkeit für alle CT-Scans aufweist, ist ein gewisses Potential der Methodik exemplarisch zu verzeichnen.

Ein positiver Nebeneffekt in Bezug auf eine verbesserte Robustheit gegenüber Bildrauschen kann intuitiv aus der Verwendung der Dropout-Layer abgeleitet, wenn auch nur schwierig für CT-Bilder quantifiziert, werden. Nach Srivastava et al. führt die Verwen-

² Gemeint ist der Teil des rekonstruierten Volumens, der keine Probe, sondern nur Luft enthält, die in guter Näherung vollständig transparent für die Photonen im CT-relevanten Energiebereich ist.

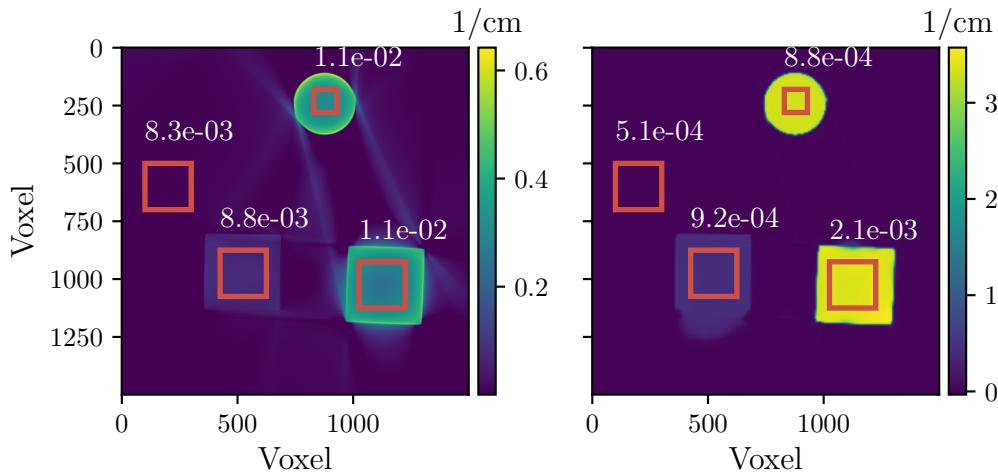


Abbildung 8.3: Rekonstruiertes CT-Schichtbild bei einer Quellspannung $U_B = 140\text{ kV}$ eines Scans mit der CT-Anlage aus Kapitel 2.7 und den CT-Parametern nach Tabelle 5.2 (links). Das monochromatische CT-Schichtbild (rechts) wurde mit einem U-Net nach der Methodik aus Abschnitt 8.1 berechnet. Die Zahlenwerte oberhalb der roten Boxen geben die normierten Standardabweichungen der rekonstruierten Abschwächungskoeffizienten innerhalb der roten Boxen wieder.

dung der Dropout-Layer während des Trainings anschaulich zu der zufälligen Aufspaltung des neuronalen Netzes in Teilinstanzen des gesamten Modells (*Sub-Networks*) [45]. Bei der Inferenz, bei der die Dropout-Layer im Normalfall deaktiviert werden, agieren nach obiger Sichtweise alle Sub-Networks additiv und mitteln ihre (Zwischen-)Ergebnisse, sodass ein Rauschen in den Eingabedaten statistisch unterdrückt wird.

8.3 DETEKTION VON POREN UND PARTIKELN

Wie in Kapitel 7.6 beschrieben, kann die Computertomographie verwendet werden, um Lithium-Ionen-Batterien zu prüfen. Im Kontext von Kapitel 7.1 wurde erläutert, dass die Materialbestimmung von Partikeln mit dem vorgestellten Ansatz nur bis zu einer gewissen Partikelgröße zuverlässig funktionieren kann. Die implementierte Simulation kann mit kleinen Anpassungen zur Erzeugung von Trainingsdaten für die Detektion von Partikeln oder Poren in CT-Scans genutzt werden: Zunächst werden zufällige Phantome nach dem in Kapitel 5.3 beschriebenen Verfahren erzeugt. Anschließend können gezielt kleine Regionen eines bestimmten Materials hinzugefügt werden, deren Positionen, analog zu den Materialkarten, als Annotationen dienen. Für das Training einer Segmentierung dieser kleinen Objekte können somit die CT-Schichtbilder als Inputs und die binär maskierten Partikel als Ground Truths verwendet werden. Alternativ können die bekannten Positionen der eingefügten Partikel für die Positionierung so-

genannter *Bounding Boxes* verwendet werden (Box-basierte Detektion). Im Kontext der beobachteten Sim-To-Real-Gap kann die Argumentation aus Abschnitt 8.1 aufgegriffen werden. Auch für die Detektion von Poren und Partikeln spielt die Streustrahlung eine stark untergeordnete Rolle. Die Partikel werden nicht anhand ihrer quantitativen Abschwächungskoeffizienten, sondern anhand ihres Kontrastunterschiedes gegenüber ihrer Umgebung erkannt – analog benötigt ein CT-Experte nicht die absoluten Werte der Abschwächungskoeffizienten zur Identifikation eines Partikels oder einer Pore innerhalb eines sonst homogenen Materials. Diese beispielhafte Anwendung der implementierten Simulation wird im Rahmen dieser Arbeit nicht erprobt, kann jedoch aufbauend zur Entwicklung eines neuronalen Netzes zur Partikel- oder Porenidentifikation genutzt werden.

8.4 ZUSAMMENFASSUNG DER ERGEBNISSE

In diesem Kapitel wurden weitere Anwendungsfälle der entwickelten Methodik aus den ersten drei Forschungsfragen identifiziert. Insbesondere die Methodik der Trainingsdatenerzeugung bietet allein durch den Austausch einiger Parameter das Potential, Trainingsdaten für alternative Anwendungsfälle, abseits der Materialbestimmung, zu generieren. Die vorgestellten Anwendungsfälle besitzen die Gemeinsamkeit, dass sie nicht auf quantitativen CT-Daten basieren und somit, im Vergleich zur Materialbestimmung, robuster gegenüber der in Abschnitt 7.6 diskutierten Sim-To-Real-Gap sind. Durch die Substitution eines der beiden Quellspektren in den Simulationsparametern können mit der implementierten Simulation Trainingsdaten für die Umrechnung zwischen poly- und monochromatischen CT-Schichtbildern erzeugt werden. Die Ergebnisse eines mit ebendiesen Daten trainierten U-Nets wurden in Abschnitt 8.1 anhand realer Scandaten präsentiert und weisen eine deutliche Reduktion der Strahlaufhärtungsartefakte auf. In Abschnitt 8.2 wurde die implizit erlernte Rauschunterdrückung des trainierten U-Nets aus Abschnitt 8.1 untersucht. Abschließend wurde die Möglichkeit diskutiert, mithilfe des Phantomgenerators (vergleiche Abschnitt 5.3) Trainingsdaten für die Box-basierte Detektion oder Segmentierung von Partikeln und Poren zu erzeugen. Alle genannten Anwendungsfälle weisen das Potential auf, auch ohne die weitere Reduktion der Sim-To-Real-Gap, erfolgreich mit simulierten Daten trainiert und in die realen CT-Anwendung übertragen werden zu können.

9

KRITISCHE WÜRDIGUNG

Im Rahmen der Einleitung in Kapitel 1 wurde die Verwendung eines DL-basierten Verfahrens zur Materialbestimmung anhand von CT-Daten motiviert. Hieraus wurden vier Forschungsfragen abgeleitet und in den Kapitel 5 bis 8 beantwortet. Ziel dieses Kapitels ist die Bewertung und kritische Diskussion ebendieser Antworten auf die Forschungsfragen.

Zur Beantwortung der ersten Forschungsfrage (Kapitel 5) wurde eine simulationsbasierte Methodik zur Trainingsdatenerzeugung für die Materialbestimmung in CT-Schichtbildern entwickelt. Die einzige, heutzutage realistisch durchführbare, Simulationstechnik für die Erzeugung großer Trainingsdatenmengen ist die Raytracing-basierte Projektion virtueller Prüfteile (vergleiche Kapitel 5.1). Sie birgt jedoch eine systematische Ungenauigkeit: die Streustrahlung, ein Produkt der Compton-Streuung (vergleiche Kapitel 2.1), wird nicht modelliert. Rückblickend auf die Wechselwirkungsprozesse von Photonen mit Materie in Abbildung 2.4 wird deutlich, dass der relative Anteil der Compton-Streuung an der gesamten Wechselwirkung von der Photonenergie abhängt. Abbildung 9.1 stellt den Anteil der Compton-Streuung an dem totalen Abschwächungskoeffizienten μ_L energieabhängig dar. Die Compton-Streuung ist, in großen Teilen des

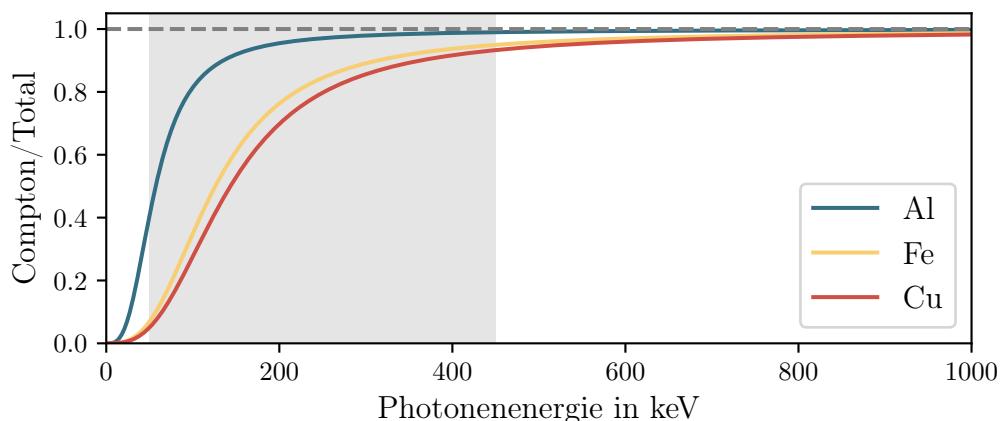


Abbildung 9.1: Anteil der Compton-Streuung am totalen Abschwächungskoeffizienten für verschiedene Materialien abhängig von der Photonenergie. Der relevante Energiebereich für die industrielle CT ist, analog zu Abbildung 2.4, grau hinterlegt.

industriell relevanten Energiebereiches, der einflussreichste Wechselwirkungsprozess, sodass ein signifikanter Anteil des gemessenen (Detektor-)Signals auf die vorwärtsgerichtete Streustrahlung (vergleiche Kapitel 2.1, Klein-Nishina-Gleichung [3, S. 50f.]) zurückzuführen ist. Durch die fehlende Modellierung der Streustrahlung in der implementierten Simulation ist eine deutliche Sim-To-Real-Gap zu beobachten (vergleiche Abbildung 5.8 und Abbildung 7.17). Ein Lösungsansatz besteht in dem expliziten Lernen der Sim-To-Real-Gap mit einem DL-Modell, sodass die simulierten Bilder, nach Pro-

zessierung durch ein entsprechend trainiertes Modell, den realen Bildern stark ähneln. Dieser Ansatz wird von Maier konkret für die fehlende Modellierung der Streustrahlung in simulierten CT-Bildern beschrieben: Er trainiert ein DL-Modell, das die absolute Differenz zwischen einer Raytracing-basierten Simulation und einer (realistischen) Monte-Carlo-basierten Simulation vorhersagt [120]. Obwohl Maier diese Methodik für einen anderen Anwendungsfall konzipiert hat, bietet sie eine Möglichkeit, um die Verteilung der Trainingsdaten im Rahmen dieser Arbeit deutlich an die real gemessenen CT-Daten anzugeleichen und um somit den Transfer auf reale CT-Daten zu verbessern. Da die Streustrahlung ein dreidimensional auftretender Effekt ist, sollte die implementierte Simulation zur Datenerzeugung ebenfalls dreidimensionale Phantome und somit zweidimensionale Projektionen berechnen, sodass die von Maier entwickelte Methode ihr volles Potential entfalten kann. Dennoch ist zu betonen, dass die von Maier trainierten Modelle für das Training stets die Projektionen aus einer Monte-Carlo-basierten Simulation benötigen, sodass diese Methode ebenfalls mit einem erheblichen Berechnungsaufwand verbunden ist. Ein alternatives Verfahren wurde durch Tercan für industrielle Sim-To-Real-Gaps beschrieben und besteht in dem Vortrainieren eines DL-Modells auf simulierten Daten und dem anschließenden Fine-Tuning auf wenigen, annotierten Realdaten [121]. Nach Tercan lernt das DL-Modell während des Vortrainings die grundlegenden Zusammenhänge der simulierten Trainingsdatenverteilung und während des Fine-Tunings lediglich den, potentiell trivialen, Versatz zur realen Datenverteilung. Beide Ansätze nach Maier und Tercan sollten in einer zukünftigen Arbeit grundsätzlich für die Verarbeitung von CT-Daten mit DL, auch abseits der Materialbestimmung, untersucht werden.

Die im Rahmen der zweiten Forschungsfrage (Kapitel 6) ausgewählte DL-Architektur, das U-Net [36], wird in der Literatur für die Materialbestimmung anhand von CT-Bildern häufig verwendet (vergleiche Tabelle 4.1). Im quantitativen Vergleich mit dem Swin-Unet [58] konnte sich das U-Net für die Materialbestimmung im Rahmen dieser Arbeit durchsetzen. Die deutliche Überlegenheit der Transformer-basierten Architekturen gegenüber den Convolution-basierten Architekturen in den großen CV-Benchmarks [111, 57, 37, 38] steht in einem Widerspruch zu den Ergebnissen von Kapitel 6. Es ist die Frage zu stellen, aus welchem Grund das U-Net bessere Ergebnisse bei der Materialbestimmung erzielen konnte. Der offenkundigste Grund besteht in einem Mangel an Trainingsdaten, da das Swin-Unet ohne vortrainierte Gewichte initialisiert wird. In den Originalpublikationen nach Liu et al. wird der Swin Transformer auf Datensätzen mit vielen Millionen Bildern zunächst vortrainiert und anschließend für die speziellen Anwendungen mittels Fine-Tuning angepasst [38]. Cao et al. geben für ihr Segmentierungsproblem aus der klinischen CT an, auf dem ImageNet-Datensatz vortrainierte Gewichte für den Swin Transformer zu verwenden [58], jedoch wird keine konkrete Referenz angegeben. Die Versuche in dieser Arbeit, einen vortrainierten Swin Transformer (ImageNet-22K) zu nutzen, führten zu keiner Verbesserung der Materialbestimmung (vergleiche Kapitel 6.4.2; Abbildung 6.11). Der Grund für dieses Resultat kann in den domänenspezifischen Unterschieden zwischen dem ImageNet-22K-Datensatz und den DECT-Schichtbildern gefunden werden. Mit einem offenen und ausreichend dimensionierten CT-Datensatz könnten parameterstarke Modelle, beispielsweise der Swin Transformer aber auch etablierte CNNs wie das ResNet,

vortrainiert werden und so den Zugang für Forschende aus dem CT-Bereich signifikant vereinfachen. Darüber hinaus können auf einem derartigen Datensatz offene Benchmarks definiert werden, sodass die für die CT entwickelten Modelle untereinander vergleichbar werden. Insbesondere für spezifische und sicherheitskritische CT-Anwendung bietet ein offener Datensatz nicht nur Vergleichbarkeit sondern ebenso Transparenz. Ein Beispiel hierfür wäre die Identifikation von Partikeln in Lithium-Ionen-Batterien anhand von CT-Schichtbildern (vergleiche Kapitel 8.3). Die absoluten Detektionswahrscheinlichkeiten, der potentiell sicherheitskritischen Partikel, werden momentan von verschiedenen CT-Systemherstellern anhand eigener, interner Datensätze angegeben. Anhand eines offenen Testdatensatzes könnten die verschiedenen Algorithmen quantitativ verglichen und eventuell sogar qualifiziert werden. Allgemein existieren im CV-Bereich bekannte Trainings- und Testdatensätze beispielsweise für die Bildklassifikation [122, 94], die zum quantitativen und anwendungsbezogenen Vergleich verschiedener DL-Architekturen genutzt werden. Simpson et al. stellen mehrere offene Datensätze bereit, um die Segmentierung trainierter Modelle anhand von klinischen Magnetresonanztomographie- und CT-Scans zu beurteilen. Demnach sollten ebenfalls offene Datensätze aus der industriellen CT publiziert werden, um einerseits eine gemeinsame Datenbasis für das Training von Basismodellen (vergleiche Swin Transformer [38]) zu schaffen und andererseits auf ebendiesen Daten verschiedene Prüfaufgaben evaluieren zu können, sodass die vorgestellten DL-Architekturen untereinander vergleichbar sind. Condon et al. stellen einen offenen Datensatz bereit, der die CT-Schichtbilder von LIBs beinhaltet [123], jedoch keine Annotationen bereitstellt. Für diesen Datensatz wäre die Ergänzung entsprechender Annotationen zur Ausführung der Segmentierung oder Detektion von Partikeln denkbar und sollte durchgeführt werden. Die Erweiterung auf andere Prüfteile und alternative CT-Parameter kann ebenfalls zielführend sein, um ausdrucksstarke Basismodelle für die Verarbeitung von industriellen CT-Daten zu entwickeln.

Die dritte Forschungsfrage (Kapitel 7) befasst sich mit der Übertragbarkeit des entwickelten Ansatzes zur Materialbestimmung in die reale CT-Praxis. In den Anforderungen an die Simulation (Kapitel 5.1) und den Rahmenbedingungen des DL-Modells (Kapitel 6.1) wurden einige Annahmen getroffen, deren Auswirkungen in der realen CT-Anwendung aufgezeigt und quantifiziert werden. Die implementierte Simulation erzeugt stets einen Trainingsdatensatz, der mit einem festen Satz an CT-Parametern (beispielsweise Tabelle 5.2) erzeugt worden ist. Diese Annahme wurde getroffen, um einerseits die Präzision der Materialbestimmung zu optimieren und andererseits die Laufzeit der Simulation zu reduzieren. Jedoch benötigt eine spezifische Anwendung nach diesem Vorgehen einen ebenso spezifischen Trainingdatensatz. In Kapitel 7.3 konnte gezeigt werden, dass das Fine-Tuning genutzt werden kann, um vortrainierte Modelle an neue CT-Geometrien anhand eines verhältnismäßig kleinen, simulierten Datensatzes mit moderatem Berechnungsaufwand anzupassen. Analog können die anderen CT-Parameter, beispielsweise die Einstellungen der Röntgenquelle, mithilfe eines kleinen Trainingsdatensatzes angeleert werden. Ein alternativer Ansatz besteht in der Entwicklung einer DL-Architektur, die neben den CT-Schichtbildern ebenfalls die CT-Parameter verarbeitet und somit eine bessere Generalisierung aufweist.

Problematischer verhält sich das räumliche Auflösungsvermögen der DL-basierten Materialbestimmung, das in Kapitel 7.1 untersucht wurde. Es konnte gezeigt werden, dass die Größe der Materialregionen, entgegen der Erwartungen, keinen Einfluss auf das räumliche Auflösungsvermögen der DL-basierten Materialbestimmung hat. Darüber hinaus ist die Messung der MTF-Kurve – eine etablierte Metrik zur Bestimmung des Auflösungsvermögens von CT-Systemen – nur bedingt zur Messung des räumlichen Auflösungsvermögens bei der Materialbestimmung geeignet (vergleiche Kapitel 7.1). Es bedarf einer angepassten Metrik, die die Eigenschaft der Materialbestimmung als Regressionsproblem berücksichtigt. Die Entwicklung einer derartigen Metrik, die sowohl die Präzision der Materialbestimmung als auch die Ortsauflösung innerhalb der CT-Schichtbilder bewerten kann, muss Gegenstand zukünftiger Forschung sein. Abschließend wurde in Kapitel 7.6 demonstriert, dass die entwickelte Methode zur Materialbestimmung aufgrund der Sim-To-Real-Gap nur bedingt auf reale CT-Daten anwendbar ist und somit weiterer Forschungsbedarf besteht. Neben den bereits diskutierten Methoden zur Reduktion der Sim-To-Real-Gap, die auf eine Verbesserung der Simulationsdaten (nach Maier [120]) beziehungsweise das Fine-Tuning mit Realdaten (nach Tercan [121]) abzielen, kann ergänzend seitens der DL-Architektur und der Trainingsstrategie eine konzeptionelle Änderung erprobt werden: das Training mit gemischten Daten aus der Simulation und der realen CT-Anlage. Konkret ist die Annotation realer CT-Daten, wie mehrfach beschrieben, sehr zeitaufwändig und daher nicht zu favorisieren, während die Annotation bei den Simulationsdaten automatisch aus der entwickelten Methodik folgt (vergleiche Kapitel 5.1.4). Es bedarf somit einer Trainingsstrategie, die annotierte Simulationsdaten und nicht annotierte Realdaten verarbeitet – ein Verfahren, das als sogenannte *Adversarial Domain Adaption* von Tzeng et al. publiziert wurde [90]. Bei der Adversarial Domain Adaption werden zwei Encoder, die jeweils die Bilder aus einer Quell- und einer Zieldomäne in einen Latent Space kodieren, gezwungen, einen sehr ähnlichen, das heißt für einen Discriminator ununterscheidbaren, Latent Space zu kodieren [90]. Übertragen auf die in dieser Arbeit gegebene Problemstellung könnte mit dieser Methode ein Encoder trainiert werden, der die realen CT-Daten und die simulierten CT-Daten in einen (nahezu) identischen Latent Space kodiert. Aus diesem Latent Space kann, mithilfe des auf simulierten Daten trainierten Decoders (unter Zuhilfenahme der annotierten Materialien), die Materialbestimmung durchgeführt werden. Anhand dieser Beschreibung erscheint die Methode der Adversarial Domain Adaption aus theoretischer Sicht für das gegebene Problem gut geeignet zu sein, wobei beachtet werden muss, dass bereits kleine morphologische Unterschiede zwischen den simulierten und realen CT-Daten als Unterscheidungsmerkmal durch den Discriminator verwendet werden können und so die gesamte Methodik destabilisieren können. Eine für CT-Daten allgemeingültige Lösung der beschriebenen Sim-To-Real-Gap beherbergt das Potential, vielfältige Anwendungen, auch abseits der Materialbestimmung, auf Basis simulierter CT-Daten zu trainieren und anschließend in die praktische CT-Anwendung zu übertragen. Somit bieten sowohl die simulationsbezogene als auch die DL-bezogene Reduktion der Sim-To-Real-Gap bei CT-Daten ein erhebliches Forschungspotential für zukünftige Arbeiten.

Trotz der im Rahmen dieser Arbeit signifikanten Sim-To-Real-Gap, die den Transfer der

Materialbestimmung auf reale CT-Daten behindert, konnten in Kapitel 8 im Rahmen der vierten Forschungsfrage exemplarisch einige Anwendungspotentiale der entwickelten Methodik für andere Aufgabenfelder innerhalb der CT aufgezeigt werden. Zusammenfassend ist die entwickelte Methode zur Trainingsdatenerzeugung für Anwendungen geeignet, die nicht der quantitativen CT zugeordnet sind – das heißt, die absoluten Abschwächungskoeffizienten eine für die Prüfaufgabe untergeordnete Rolle spielen. Zu diesen Prüfaufgaben zählt beispielhaft die räumliche Detektion von Partikeln innerhalb eines CT-Schichtbildes, für die der Einfluss der Streustrahlung zu vernachlässigen ist.

10

ZUSAMMENFASSUNG

Die industrielle CT ist eine zerstörungsfreie Prüftechnik, die dreidimensionale, digitale Abbilder, die sogenannten *CT-Volumen*, der Prüfteile mithilfe von Röntgenstrahlung erzeugt. Physikalisch betrachtet ist die CT somit ein ortsaufgelöstes Messverfahren der linearen Abschwächungskoeffizienten μ_L der Röntgenstrahlung, die zur Bildaufnahme verwendet wird. Verschiedene Materialien weisen unterschiedliche und energieabhängige lineare Abschwächungskoeffizienten auf, deren Verlauf charakteristisch für das jeweilige Material ist. Somit besteht die Möglichkeit, aus den mithilfe der CT gemessenen linearen Abschwächungskoeffizienten, ortsaufgelöste Rückschlüsse auf die Materialien zu ziehen, aus denen das gescannte Prüfteil besteht – die Materialbestimmung. Neben vielfältigen Anwendungen im Umfeld der klinischen CT, die auch den Ursprung der CT-basierten Materialbestimmung darstellt, ist die Materialbestimmung in der industriellen CT ein wertvolles Werkzeug zur Charakterisierung und Qualitätssicherung von Prüfteilen.

Seit Jahrzehnten werden Methoden zur Materialbestimmung anhand von CT-Volumen einerseits durch Anwendung der physikalischen Gesetze (klassische Algorithmen) und andererseits durch datengetriebene Verfahren (künstliche neuronale Netze), zu denen auch das Deep Learning zählt, erforscht [13, 14, 22]. Im Rahmen dieser Dissertations-schrift wird ein DL-basiertes Verfahren zur Materialbestimmung vorgestellt.

Zunächst werden im Rahmen der ersten Forschungsfrage die Anforderungen an die simulationsbasierte Trainingsdatenerzeugung sowie die Implementierung ebendieser Simulation vorgenommen. Die implementierte Simulation basiert auf einem Raytracing-basierten Ansatz und kann demnach einen der wichtigsten Wechselwirkungsprozesse, die Compton-Streuung, nur bedingt modellieren. Somit unterliegen die simulierten CT-Daten einer Sim-To-Real-Gap, die durch den Vergleich mit realen CT-Daten ersichtlich wird. Dennoch bietet die implementierte Simulation aufgrund ihrer Geschwindigkeit und Skalierbarkeit das Potential, Datensätze zu berechnen, die mit mehreren Millionen Einträgen eine wichtige Grundlage für das Training parameterstarker DL-Modelle auch für zukünftige Forschungsarbeiten bilden.

In der zweiten Forschungsfrage wird eine geeignete DL-Architektur für die zu lernende Materialbestimmung gesucht. Formell ist die Materialbestimmung eine sogenannte *Dense Prediction*, die für jedes Pixel eines Eingabebildes eine (Material-)Vorhersage berechnet. Aufgrund der Ergebnisse einer Literaturrecherche im Rahmen dieser Dissertationsschrift wird zunächst das U-Net [36] implementiert und für die Materialbestimmung optimiert. Weiterhin ist der Erfolg von Transformer-basierten Architekturen in der Bildverarbeitung [57, 37, 38] unbestreitbar, sodass auch für die Materialbestimmung eine Transformer-basierte Architektur erprobt werden muss. Hierzu wird das Swin-Unet [58] implementiert und quantitativ mit dem U-Net verglichen. Obschon die Materialbestimmung einzelner, ausgewählter Materialien mit dem Swin-Unet bessere Ergebnisse hervorbringt, sind lediglich die Vorhersagen mit dem U-Net auf allen betrachteten Testmaterialien innerhalb

der von CT-Anwendern geforderten Fehlertoleranzen, sodass das U-Net für die weitere Untersuchungen verwendet wird. Die Ergebnisse geben ein Hinweis auf das in der Literatur bisher unausgeschöpfte Potential Transformer-basierter Architekturen für die Verarbeitung von CT-Daten, welches in dieser Dissertationsschrift teilweise aufgedeckt, jedoch nicht vollständig genutzt werden konnte.

In der dritten Forschungsfrage werden die Auswirkungen der in den ersten beiden Forschungsfragen getroffenen Annahmen beim Transfer in die reale CT-Anwendung untersucht. Es wird gezeigt, dass die auf festen Probengrößen trainierten Modelle mittels Fine-Tuning an andere Probengrößen angepasst werden können und die geforderte Präzision der Materialbestimmung erreichen. Weiterhin wird demonstriert, dass die Annahme der Fächerstrahl-CT in den simulierten Daten (es werden zweidimensionale Proben auf einen eindimensionalen Detektor projiziert) ebenfalls auf dreidimensionale CT-Daten schichtweise angewandt werden kann, sofern die Helix-Trajektorie zur Datenaufnahme verwendet wird. Die Materialbestimmung poröser Proben ist ebenfalls möglich, wenn die verwendeten Trainingsdaten ebendiese Porositäten enthalten. Bei der Beurteilung des räumlichen Auflösungsvermögens wurde festgestellt, dass die in der CT verbreitete MTF-Kurve nicht mit der Materialbestimmung kompatibel ist und an dieser Stelle weiterer Forschungsbedarf besteht. Ebenfalls konnte der Einfluss der bereits beobachteten Sim-To-Real-Gap auf die Materialbestimmung anhand eines realen CT-Scans mit bekannten Probenmaterialien bestätigt werden: Die fehlende Modellierung der Streustrahlung senkt die rekonstruierten Abschwächungskoeffizienten ab, sodass insbesondere die Dichtevorhersagen unterhalb der Literaturwerte zu verordnen sind. Klar ist, dass die Reduktion der Sim-To-Real-Gap ein zentraler Bestandteil zukünftiger Forschung sein muss, um einerseits die Materialbestimmung und andererseits die DL-basierte Datenverarbeitung von CT-Daten zu verbessern.

Abschließend werden in der vierten Forschungsfrage weitere Anwendungsmöglichkeiten der entwickelten Methodik, abseits der Materialbestimmung, genannt und im Kontext der beobachteten Sim-To-Real-Gap hinsichtlich ihrer Realisierbarkeit eingeordnet. So besteht die Möglichkeit, DL-Modelle beispielsweise auf die Detektion von Partikeln zu trainieren, da die Erkennung der Partikel nicht auf absoluten Werten, sondern auf dem relativen Bildkontrast in einer lokalen Umgebung erfolgen kann. Im Allgemeinen sind Anwendungsfälle, die nur geringfügig von der fehlenden Modellierung der Streustrahlung beeinflusst werden, mit der entwickelten Methodik realisierbar. Hierzu zählt gegebenenfalls auch die Materialbestimmung bei sehr viel niedrigeren Photonenergien und mit leichten Materialien – zum Beispiel für die klinische CT.

LITERATURVERZEICHNIS

- [1] Röntgen, Wilhelm C.: Ueber eine neue Art von Strahlen. In: Sitzungsberichte Physikal. Institut der Universität Würzburg (1896). (Zugriff: 2024-06-13)
- [2] Nobel Prize Outreach: The Nobel Prize in Physics 1901. <https://www.nobelprize.org/prizes/physics/1901/roentgen/facts/>. (Zugriff: 2024-06-18). Version: 1901
- [3] Knoll, Glenn F.: Radiation Detection and Measurement. 3. ed. New York, NY Weinheim : Wiley, 2000. – ISBN 978–0–471–07338–3
- [4] Dally, J.F.Halls: On the Use of the Roentgen Rays in the Diagnosis of Pulmonary Disease. In: The Lancet 161 (1903), Nr. 4165, 1800–1806. [http://dx.doi.org/10.1016/S0140-6736\(01\)86540-5](http://dx.doi.org/10.1016/S0140-6736(01)86540-5). (Zugriff: 2024-10-21). – DOI 10.1016/S0140-6736(01)86540-5. – ISSN 0140–6736
- [5] Moseley, H.G.J.: LXXX. The High-Frequency Spectra of the Elements. Part II. In: The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 27 (1914), Nr. 160, 703–713. <http://dx.doi.org/10.1080/14786440408635141>. (Zugriff: 2024-10-21). – DOI 10.1080/14786440408635141
- [6] Nobel Prize Outreach: The Nobel Prize in Physics 1914. <https://www.nobelprize.org/prizes/physics/1914/summary/>. (Zugriff: 2024-10-21). Version: 1914
- [7] Watson, J. D. ; Crick, F. H. C.: Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. In: Nature 171 (1953), April, Nr. 4356, 737–738. <http://dx.doi.org/10.1038/171737a0>. (Zugriff: 2024-10-21). – DOI 10.1038/171737a0. – ISSN 1476–4687
- [8] Nobel Prize Outreach: The Nobel Prize in Physiology or Medicine 1979. <https://www.nobelprize.org/prizes/medicine/1979/press-release/>. (Zugriff: 2024-05-21). Version: 1979
- [9] Schreiber, Joseph J. ; Anderson, Paul A. ; Hsu, Wellington K.: Use of Computed Tomography for Assessing Bone Mineral Density. In: Neurosurgical Focus 37 (2014), Juli, Nr. 1, E4. <http://dx.doi.org/10.3171/2014.5.FOCUS1483>. (Zugriff: 2024-08-13). – DOI 10.3171/2014.5.FOCUS1483. – ISSN 1092–0684
- [10] Nagpal, Prashant ; Guo, Junfeng ; Shin, Kyung M. ; Lim, Jae-Kwang ; Kim, Ki B. ; Comellas, Alejandro P. ; Kaczka, David W. ; Peterson, Samuel ; Lee, Chang H. ; Hoffman, Eric A.: Quantitative CT Imaging and Advanced Visualization Methods: Potential Application in Novel Coronavirus Disease 2019 (COVID-19) Pneumonia. In: BJR | Open 3 (2021), Januar, Nr. 1, 20200043. <http://dx.doi.org/10.1259/bjro.20200043>. (Zugriff: 2024-08-13). – DOI 10.1259/bjro.20200043. – ISSN 2513–9878

- [11] Fang, Wei ; Li, Liang: Attenuation Image Referenced (AIR) Effective Atom Number Image Calculation for MeV Dual-Energy Container CT Using Image-Domain Deep Learning Framework. In: *Results in Physics* 35 (2022), 105406. <http://dx.doi.org/10.1016/j.rinp.2022.105406>. (Zugriff: 2022-11-08). – DOI 10.1016/j.rinp.2022.105406. – ISSN 2211-3797
- [12] Rahe, Christiane ; Kelly, Stephen T. ; Rad, Mansoureh N. ; Sauer, Dirk U. ; Mayer, Joachim ; Figgemeier, Egbert: Nanoscale X-ray Imaging of Ageing in Automotive Lithium Ion Battery Cells. In: *Journal of Power Sources* 433 (2019), 126631. <http://dx.doi.org/10.1016/j.jpowsour.2019.05.039>. (Zugriff: 2024-10-22). – DOI 10.1016/j.jpowsour.2019.05.039. – ISSN 0378-7753
- [13] Alvarez, R E. ; Macovski, A: Energy-Selective Reconstructions in X-ray Computerised Tomography. In: *Physics in Medicine and Biology* 21 (1976), September, Nr. 5, 733–744. <http://dx.doi.org/10.1088/0031-9155/21/5/002>. (Zugriff: 2022-02-17). – DOI 10.1088/0031-9155/21/5/002. – ISSN 00319155
- [14] Heismann, B. J. ; Leppert, J. ; Stierstorfer, K.: Density and Atomic Number Measurements with Spectral X-Ray Attenuation Method. In: *Journal of Applied Physics* 94 (2003), August, Nr. 3, 2073–2079. <http://dx.doi.org/10.1063/1.1586963>. (Zugriff: 2024-10-22). – DOI 10.1063/1.1586963. – ISSN 0021-8979, 1089-7550
- [15] Nachtrab, F. ; Weis, S. ; Keßling, P. ; Sukowski, F. ; Haßler, U. ; Fuchs, T. ; Uhlmann, N. ; Hanke, R.: Quantitative Material Analysis by Dual-Energy Computed Tomography for Industrial NDT Applications. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 633 (2011), S159-S162. <http://dx.doi.org/10.1016/j.nima.2010.06.154>. (Zugriff: 2022-08-31). – DOI 10.1016/j.nima.2010.06.154. – ISSN 0168-9002
- [16] Wang, Shaoyu ; Cai, Ailong ; Wu, Weiwen ; Zhang, Tao ; Liu, Fenglin ; Yu, Hengyong: IMD-MTFC: Image-Domain Material Decomposition via Material-Image Tensor Factorization and Clustering for Spectral CT. In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 7 (2023), April, Nr. 4, 382–393. <http://dx.doi.org/10.1109/TRPMS.2023.3234613>. (Zugriff: 2024-03-25). – DOI 10.1109/TRPMS.2023.3234613. – ISSN 2469-7311, 2469-7303
- [17] paperswithcode: *Object Detection on COCO Test-Dev*. <https://paperswithcode.com/sota/object-detection-on-coco>. (Zugriff: 2024-10-22). Version: 2024
- [18] paperswithcode: *Image Classification on ImageNet*. <https://paperswithcode.com/sota/image-classification-on-imagenet>. (Zugriff: 2024-10-22). Version: 2024
- [19] paperswithcode: *Semantic Segmentation on ADE20K*. <https://paperswithcode.com/sota/semantic-segmentation-on-ade20k>. (Zugriff: 2024-10-22). Version: 2024
- [20] paperswithcode: *Medical Image Segmentation on Kvasir-SEG*. <https://paperswithcode.com/sota/medical-image-segmentation-on-kvasir-seg>. (Zugriff: 2024-10-22). Version: 2024

- [21] paperswithcode: Medical Image Segmentation on CVC-ClinicDB. <https://paperswithcode.com/sota/medical-image-segmentation-on-cvc-clinicdb>. (Zugriff: 2024-10-22). Version: 2024
- [22] Weiss, Moritz ; Meisen, Tobias: Reviewing Material-Sensitive Computed Tomography: From Handcrafted Algorithms to Modern Deep Learning. In: NDT 2 (2024), Juli, Nr. 3, 286–310. <http://dx.doi.org/10.3390/ndt2030018>. (Zugriff: 2024-08-13). – DOI 10.3390/ndt2030018. – ISSN 2813-477X
- [23] Demtröder, Wolfgang: Experimentalphysik 3. Berlin, Heidelberg : Springer Berlin Heidelberg, 2010 (Springer-Lehrbuch). <http://link.springer.com/10.1007/978-3-642-03911-9>. (Zugriff: 2023-08-24). – ISBN 978-3-642-03910-2 978-3-642-03911-9
- [24] Bellon, C.: aRTist - Analytical RT Inspection Simulation Tool. BAM Federal Institute for Materials Research and Testing. <https://artist.bam.de/en/index.htm>. (Zugriff: 2024-05-08)
- [25] Becker, F. ; Köhnke, D. ; Reichardt, M. ; Budelmann, H.: Investigation of Various Concrete Compositions with Respect to Gamma-Radiation Transmission Properties of Cs-137. In: Radiation Physics and Chemistry 171 (2020), Juni, 108679. <http://dx.doi.org/10.1016/j.radphyschem.2020.108679>. (Zugriff: 2025-01-27). – DOI 10.1016/j.radphyschem.2020.108679. – ISSN 0969806X
- [26] Thompson, A. ; Attwood, D. ; Gullikson, E. ; Howells, M. ; Kim, KJ. ; Kirz, J. ; Kortright, J. ; Lindau, I. ; Liu, Y. ; Pianetta, P. ; Robinson, A. ; Scofield, J. ; Underwood, J. ; Williams, G. ; Winick, H.: X-RAY DATA BOOKLET. 3. Berkeley : Lawrence Berkeley National Laboratory, 2009
- [27] E07 Committee: Practice for Manufacturing Characterization of Digital Detector Arrays. <http://www.astm.org/cgi-bin/resolver.cgi?E2597E2597M-14>. (Zugriff: 2024-10-17)
- [28] Gonzalez, Rafael C. ; Woods, Richard E.: Digital Image Processing. Third Edition. New York, New York : Pearson Education, 2008. – ISBN 0-13-168728-X
- [29] Lee, Eunae ; Hong, Eunyeong ; Kim, Dong S.: Using Deep Learning for Pixel-Defect Corrections in Flat-Panel Radiography Imaging. In: Journal of Medical Imaging 8 (2021), März, Nr. 02. <http://dx.doi.org/10.1117/1.JMI.8.2.023501>. (Zugriff: 2024-10-16). – DOI 10.1117/1.JMI.8.2.023501. – ISSN 2329-4302
- [30] Maier, Andreas (Hrsg.) ; Steidl, Stefan (Hrsg.) ; Christlein, Vincent (Hrsg.) ; Hornegger, Joachim (Hrsg.): Lecture Notes in Computer Science. Bd. 11111: Medical Imaging Systems: An Introductory Guide. Cham : Springer International Publishing, 2018 <http://link.springer.com/10.1007/978-3-319-96520-8>. (Zugriff: 2022-02-15). – ISBN 978-3-319-96519-2 978-3-319-96520-8
- [31] Radon, Johann: Über Die Bestimmung von Funktionen Längs Gewisser Mannigfaltigkeiten. In: Berichte über die Verhandlungen der Königlich-Sächsischen

Gesellschaft der Wissenschaften zu Leipzig 69, Nr. 1917, S. 262–277. (Zugriff: 2022-08-30)

- [32] Kak, Avinash C. ; Slaney, Malcolm: Principles of Computerized Tomographic Imaging: "This SIAM Edition Is an Unabridged Republication of the Work First Published by IEEE Press, New York, 1988.". Philadelphia, Pa : Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 2001 (Classics in Applied Mathematics 33). – ISBN 978-0-89871-494-4 978-0-89871-927-7
- [33] Feldkamp, L. A. ; Davis, L. C. ; Kress, J. W.: Practical Cone-Beam Algorithm. In: Journal of the Optical Society of America A 1 (1984), Juni, Nr. 6, 612. <http://dx.doi.org/10.1364/JOSAA.1.000612>. (Zugriff: 2022-02-15). – DOI 10.1364/JOSAA.1.000612. – ISSN 1084-7529, 1520-8532
- [34] Perry Sprawls: Physical Principles of Medical Imaging (Web Version). 2nd. 1993 (Zugriff: 2024-06-24)
- [35] McCollough, Cynthia H. ; Boedeker, Kirsten ; Cody, Dianna ; Duan, Xinhui ; Flohr, Thomas ; Halliburton, Sandra S. ; Hsieh, Jiang ; Layman, Rick R. ; Pelc, Norbert J.: Principles and Applications of Multienergy CT: Report of AAPM Task Group 291. In: Medical Physics 47 (2020), Juli, Nr. 7. <http://dx.doi.org/10.1002/mp.14157>. (Zugriff: 2022-01-07). – DOI 10.1002/mp.14157. – ISSN 0094-2405, 2473-4209
- [36] Ronneberger, Olaf ; Fischer, Philipp ; Brox, Thomas: U-Net: Convolutional Networks for Biomedical Image Segmentation. <http://arxiv.org/abs/1505.04597>. (Zugriff: 2022-06-09). Version: Mai 2015
- [37] Liu, Ze ; Lin, Yutong ; Cao, Yue ; Hu, Han ; Wei, Yixuan ; Zhang, Zheng ; Lin, Stephen ; Guo, Baining: Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. <http://arxiv.org/abs/2103.14030>. (Zugriff: 2023-04-03). Version: August 2021
- [38] Liu, Ze ; Hu, Han ; Lin, Yutong ; Yao, Zhuliang ; Xie, Zhenda ; Wei, Yixuan ; Ning, Jia ; Cao, Yue ; Zhang, Zheng ; Dong, Li ; Wei, Furu ; Guo, Baining: Swin Transformer V2: Scaling Up Capacity and Resolution. <http://arxiv.org/abs/2111.09883>. (Zugriff: 2024-09-18). Version: April 2022
- [39] Hornik, Kurt ; Stinchcombe, Maxwell ; White, Halbert: Multilayer Feedforward Networks Are Universal Approximators. In: Neural Networks 2 (1989), Januar, Nr. 5, 359–366. [http://dx.doi.org/10.1016/0893-6080\(89\)90020-8](http://dx.doi.org/10.1016/0893-6080(89)90020-8). (Zugriff: 2025-02-21). – DOI 10.1016/0893-6080(89)90020-8. – ISSN 08936080
- [40] Goodfellow, I. ; Bengio, Y. ; Courville, A.: Deep Learning: Das Umfassende Handbuch : Grundlagen, Aktuelle Verfahren Und Algorithmen, Neue Forschungsansätze. 1. Frechen : mitp, 2018. – ISBN 978-3-95845-700-3

- [41] Ciampiconi, Lorenzo ; Elwood, Adam ; Leonardi, Marco ; Mohamed, Ashraf ; Rozza, Alessandro: A Survey and Taxonomy of Loss Functions in Machine Learning. (2024), November. <http://dx.doi.org/10.48550/arXiv.2301.05579>. (Zugriff: 2025-01-17). – DOI 10.48550/arXiv.2301.05579
- [42] Loshchilov, Ilya ; Hutter, Frank: Decoupled Weight Decay Regularization. (2019), Januar. <http://dx.doi.org/10.48550/arXiv.1711.05101>. (Zugriff: 2024-12-28). – DOI 10.48550/arXiv.1711.05101
- [43] Ioffe, Sergey ; Szegedy, Christian: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. (2015), März. <http://dx.doi.org/10.48550/arXiv.1502.03167>. (Zugriff: 2025-01-02). – DOI 10.48550/arXiv.1502.03167
- [44] Lin, Runze: Analysis on the Selection of the Appropriate Batch Size in CNN Neural Network. In: 2022 International Conference on Machine Learning and Knowledge Engineering (MLKE) (2022), Februar, 106–109. <http://dx.doi.org/10.1109/MLKE55170.2022.00026>. (Zugriff: 2025-01-19). – DOI 10.1109/MLKE55170.2022.00026
- [45] Srivastava, Nitish ; Hinton, Geoffrey ; Krizhevsky, Alex ; Sutskever, Ilya ; Salakhutdinov, Ruslan: Dropout: A Simple Way to Prevent Neural Networks from Overfitting. In: Journal of Machine Learning Research 15 (2014), Nr. 56, 1929–1958. <http://jmlr.org/papers/v15/srivastava14a.html>. (Zugriff: 2024-12-28)
- [46] Loshchilov, Ilya ; Hutter, Frank: SGDR: Stochastic Gradient Descent with Warm Restarts. (2017), Mai. <http://dx.doi.org/10.48550/arXiv.1608.03983>. (Zugriff: 2025-01-10). – DOI 10.48550/arXiv.1608.03983
- [47] Goyal, Priya ; Dollár, Piotr ; Girshick, Ross ; Noordhuis, Pieter ; Wesolowski, Łukasz ; Kyrola, Aapo ; Tulloch, Andrew ; Jia, Yangqing ; He, Kaiming: Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour. <http://dx.doi.org/10.48550/arXiv.1706.02677>. (Zugriff: 2025-01-20). Version: April 2018
- [48] Kalra, Dayal S. ; Barkeshli, Maissam: Why Warmup the Learning Rate? Underlying Mechanisms and Improvements. (2024), November. <http://dx.doi.org/10.48550/arXiv.2406.09405>. (Zugriff: 2025-01-20). – DOI 10.48550/arXiv.2406.09405
- [49] LeCun, Yann ; Bottou, Leon ; Bengio, Yoshua ; Ha, Patrick: Gradient-Based Learning Applied to Document Recognition. In: Proceedings of the IEEE 86 (1998), Nr. 11, S. 2278–2324. <http://dx.doi.org/10.1109/5.726791>. (Zugriff: 2024-10-24). – DOI 10.1109/5.726791
- [50] Sobel, I. ; Feldman, G.: A 3x3 Isotropic Gradient Operator for Image Processing. (1968). https://www.researchgate.net/publication/285159837_A_33_isotropic_gradient_operator_for_image_processing. (Zugriff: 2025-01-20)

- [51] He, Kaiming ; Zhang, Xiangyu ; Ren, Shaoqing ; Sun, Jian: Deep Residual Learning for Image Recognition. (2015), Dezember. <http://dx.doi.org/10.48550/arXiv.1512.03385>. (Zugriff: 2025-01-20). – DOI 10.48550/arXiv.1512.03385
- [52] Sugirtha, T ; Sridevi, M: Semantic Segmentation Using Modified U-Net for Autonomous Driving. In: 2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS) (2022), Juni, 1–7. <http://dx.doi.org/10.1109/IEMTRONICS55184.2022.9795710>. (Zugriff: 2024-12-04). – DOI 10.1109/IEMTRONICS55184.2022.9795710
- [53] Chhor, Guillaume ; Bougdal-Lambert, Ianis: Satellite Image Segmentation for Building Detection Using U-net. <https://cs229.stanford.edu/proj2017/final-reports/5243715.pdf>. (Zugriff: 2024-12-04). Version: 2017
- [54] Kramarczyk, P. ; Hejmanowska, B.: UNET NEURAL NETWORK IN AGRICULTURAL LAND COVER CLASSIFICATION USING SENTINEL-2. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLVIII-1/W3-2023 (2023), Oktober, 85–90. <http://dx.doi.org/10.5194/isprs-archives-XLVIII-1-W3-2023-85-2023>. (Zugriff: 2024-12-04). – DOI 10.5194/isprs-archives-XLVIII-1-W3-2023-85-2023. – ISSN 2194-9034
- [55] Baumann, Anton ; Roßberg, Thomas ; Schmitt, Michael: Probabilistic MIMO U-Net: Efficient and Accurate Uncertainty Estimation for Pixel-wise Regression. In: 2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW) (2023), Oktober, 4500–4508. <http://dx.doi.org/10.1109/ICCVW60793.2023.00484>. (Zugriff: 2024-12-04). – DOI 10.1109/ICCVW60793.2023.00484
- [56] Vaswani, Ashish ; Shazeer, Noam ; Parmar, Niki ; Uszkoreit, Jakob ; Jones, Llion ; Gomez, Aidan N. ; Kaiser, Lukasz ; Polosukhin, Illia: Attention Is All You Need. (2017), Dezember. <http://dx.doi.org/10.48550/arXiv.1706.03762>. (Zugriff: 2023-02-07). – DOI 10.48550/arXiv.1706.03762
- [57] Dosovitskiy, Alexey ; Beyer, Lucas ; Kolesnikov, Alexander ; Weissenborn, Dirk ; Zhai, Xiaohua ; Unterthiner, Thomas ; Dehghani, Mostafa ; Minderer, Matthias ; Heigold, Georg ; Gelly, Sylvain ; Uszkoreit, Jakob ; Houlsby, Neil: An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. (2021), Juni. <http://dx.doi.org/10.48550/arXiv.2010.11929>. (Zugriff: 2023-04-18). – DOI 10.48550/arXiv.2010.11929
- [58] Cao, Hu ; Wang, Yueyue ; Chen, Joy ; Jiang, Dongsheng ; Zhang, Xiaopeng ; Tian, Qi ; Wang, Manning: Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation. (2021), Mai. <http://dx.doi.org/10.48550/arXiv.2105.05537>. (Zugriff: 2024-10-02). – DOI 10.48550/arXiv.2105.05537
- [59] Vom Brocke, Jan ; Simons, Alexander ; Riemer, Kai ; Niehaves, Björn ; Plattfaut, Ralf ; Cleven, Anne: Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. In: Communications of

- the Association for Information Systems 37 (2015). <http://dx.doi.org/10.17705/1CAIS.03709>. (Zugriff: 2024-05-13). – DOI 10.17705/1CAIS.03709. – ISSN 15293181
- [60] Pranckutė, Raminta: Web of Science (WoS) and Scopus: The Titans of Bibliographic Information in Today's Academic World. In: Publications 9 (2021), März, Nr. 1, 12. <http://dx.doi.org/10.3390/publications9010012>. (Zugriff: 2025-01-27). – DOI 10.3390/publications9010012. – ISSN 2304-6775
- [61] Roessl, E ; Proksa, R: K-Edge Imaging in x-Ray Computed Tomography Using Multi-Bin Photon Counting Detectors. In: Physics in Medicine and Biology 52 (2007), August, Nr. 15, 4679–4696. <http://dx.doi.org/10.1088/0031-9155/52/15/020>. (Zugriff: 2024-03-25). – DOI 10.1088/0031-9155/52/15/020. – ISSN 0031-9155, 1361–6560
- [62] Son, Kihong ; Kim, Daehong ; Lee, Sooyeul: Improving the Accuracy of the Effective Atomic Number (EAN) and Relative Electron Density (RED) with Stoichiometric Calibration on PCD-CT Images. In: Sensors 22 (2022), November, Nr. 23, 9220. <http://dx.doi.org/10.3390/s22239220>. (Zugriff: 2024-03-25). – DOI 10.3390/s22239220. – ISSN 1424-8220
- [63] Jumanazarov, Doniyor ; Alimova, Asalkhon ; Abdikarimov, Azamat ; Koo, Jakeoung ; Poulsen, Henning F. ; Olsen, Ulrik L. ; Iovea, Mihai: Material Classification Using Basis Material Decomposition from Spectral X-ray CT. In: Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 1056 (2023), November, 168637. <http://dx.doi.org/10.1016/j.nima.2023.168637>. (Zugriff: 2024-03-06). – DOI 10.1016/j.nima.2023.168637. – ISSN 01689002
- [64] Firsching, Markus: Material Reconstruction in X-ray Imaging. Erlangen-Nürnberg, Friedrich-Alexander-Universität, Dissertation, 2009. https://ecap.nat.fau.de/wp-content/uploads/2017/05/2009_Firsching_Dissertation.pdf. (Zugriff: 2022-04-21)
- [65] Abascal, Juan F P J. ; Ducros, Nicolas ; Pronina, Valeriya ; Rit, Simon ; Rodesch, Pierre-Antoine ; Broussaud, Thomas ; Bussod, Suzanne ; Douek, Philippe ; Hauptmann, Andreas ; Arridge, Simon ; PEYRIN, Françoise: Material Decomposition in Spectral CT Using Deep Learning: A Sim2Real Transfer Approach. In: IEEE Access 9 (2021), 25632–25647. <http://dx.doi.org/10.1109/ACCESS.2021.3056150>. (Zugriff: 2022-02-08). – DOI 10.1109/ACCESS.2021.3056150
- [66] Bussod, Suzanne ; Abascal, Juan F. ; Arridge, Simon ; Hauptmann, Andreas ; Chappard, Christine ; Ducros, Nicolas ; Peyrin, Françoise: Convolutional Neural Network for Material Decomposition in Spectral CT Scans. In: 2020 28th European Signal Processing Conference (EUSIPCO) (2021), Januar, 1259–1263. <http://dx.doi.org/10.23919/Eusipco47968.2020.9287781>. (Zugriff: 2024-03-25). – DOI 10.23919/Eusipco47968.2020.9287781

- [67] Long, Zourong ; Feng, Peng ; He, Peng ; Wu, Xiaochuan ; Guo, Xiaodong ; Ren, Xuezhi ; Chen, Mianyi ; Gao, Jingxuan ; Wei, Biao ; Cong, Wenxiang: Fully Convolutional Pyramidal Residual Network for Material Discrimination of Spectral CT. In: *IEEE Access* 7 (2019), 167187–167194. <http://dx.doi.org/10.1109/ACCESS.2019.2953942>. (Zugriff: 2024-03-25). – DOI 10.1109/ACCESS.2019.2953942. – ISSN 2169–3536
- [68] Guo, Xiaodong ; He, Peng ; Lv, Xiaojie ; Ren, Xuezhi ; Li, Yonghui ; Liu, Yuanfeng ; Lei, Xiaohua ; Feng, Peng ; Shan, Hongming: Material Decomposition of Spectral CT Images via Attention-Based Global Convolutional Generative Adversarial Network. In: *Nuclear Science and Techniques* 34 (2023), März, Nr. 3, 45. <http://dx.doi.org/10.1007/s41365-023-01184-5>. (Zugriff: 2024-03-25). – DOI 10.1007/s41365-023-01184-5. – ISSN 1001–8042, 2210–3147
- [69] Shi, Zaifeng ; Kong, Fanning ; Cheng, Ming ; Cao, Huaisheng ; Ouyang, Shunxin ; Cao, Qingjie: Multi-Energy CT Material Decomposition Using Graph Model Improved CNN. In: *Medical & Biological Engineering & Computing* 62 (2024), April, Nr. 4, 1213–1228. <http://dx.doi.org/10.1007/s11517-023-02986-w>. (Zugriff: 2024-03-25). – DOI 10.1007/s11517-023-02986-w. – ISSN 0140–0118, 1741–0444
- [70] Shi, Zaifeng ; Li, Huilong ; Li, Jinzhuo ; Wang, Zhongqi ; Cao, Qingjie: Raw-Data-Based Material Decomposition Using Modified U-Net for Low-Dose Spectral CT. In: *2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)* (2019), Oktober, 1–5. <http://dx.doi.org/10.1109/CISP-BMEI48845.2019.8965998>. (Zugriff: 2024-03-25). – DOI 10.1109/CISP-BMEI48845.2019.8965998
- [71] Xing, Yuxiang ; Zhang, Li ; Duan, Xinhui ; Cheng, Jianping ; Chen, Zhiqiang: A Reconstruction Method for Dual High-Energy CT With MeV X-Rays. In: *IEEE Transactions on Nuclear Science* 58 (2011), April, Nr. 2, 537–546. <http://dx.doi.org/10.1109/TNS.2011.2112779>. (Zugriff: 2023-03-09). – DOI 10.1109/TNS.2011.2112779. – ISSN 0018–9499, 1558–1578
- [72] Wang, Guoshuai ; Liu, Zhou ; Huang, Zhengyong ; Zhang, Na ; Luo, Honghong ; Liu, Lijian ; Shen, Hao ; Che, Canwen ; Niu, Tianye ; Liang, Dong ; Luo, Dehong ; Hu, Zhanli: Improved GAN: Using a Transformer Module Generator Approach for Material Decomposition. In: *Computers in Biology and Medicine* 149 (2022), Oktober, 105952. <http://dx.doi.org/10.1016/j.combiomed.2022.105952>. (Zugriff: 2024-03-20). – DOI 10.1016/j.combiomed.2022.105952. – ISSN 00104825
- [73] Li, Yinsheng ; Tie, Xin ; Li, Ke ; Zhang, Ran ; Qi, Zhihua ; Budde, Adam ; Grist, Thomas M. ; Chen, Guang-Hong: A Quality-checked and Physics-constrained Deep Learning Method to Estimate Material Basis Images from single-kV Contrast-enhanced Chest CT Scans. In: *Medical Physics* 50 (2023), Juni, Nr. 6, 3368–3388. <http://dx.doi.org/10.1002/mp.16352>. (Zugriff: 2024-03-25). – DOI 10.1002/mp.16352. – ISSN 0094–2405, 2473–4209

- [74] Su, Ting ; Sun, Xindong ; Yang, Jiecheng ; Mi, Donghua ; Zhang, Yikun ; Wu, Haodi ; Fang, Shibo ; Chen, Yang ; Zheng, Hairong ; Liang, Dong ; Ge, Yongshuai: DIRECT-Net: A Unified Mutual-domain Material Decomposition Network for Quantitative Dual-energy CT Imaging. In: *Medical Physics* 49 (2022), Februar, Nr. 2, 917–934. <http://dx.doi.org/10.1002/mp.15413>. (Zugriff: 2024-03-25). – DOI 10.1002/mp.15413. – ISSN 0094–2405, 2473–4209
- [75] Cao, Wenchao ; Shapira, Nadav ; Maidment, Andrew ; Daerr, Heiner ; Noël, Peter B.: Hepatic Dual-Contrast CT Imaging: Slow Triple kVp Switching CT with CNN-based Sinogram Completion and Material Decomposition. In: *Journal of Medical Imaging* 9 (2022), Januar, Nr. 01. <http://dx.doi.org/10.1117/1.JMI.9.1.014003>. (Zugriff: 2024-03-25). – DOI 10.1117/1.JMI.9.1.014003. – ISSN 2329–4302
- [76] Weiss, Moritz ; Brierley, Nick ; Von Schmid, Mirko ; Meisen, Tobias: End-To-End Deep Learning Material Discrimination Using Dual-Energy LINAC-CT. In: *e-Journal of Nondestructive Testing* 29 (2024), März, Nr. 3. <http://dx.doi.org/10.58286/29244>. (Zugriff: 2024-04-19). – DOI 10.58286/29244. – ISSN 1435–4934
- [77] Krebbers, Leonard T. ; Grozmani, Natalia ; Lottermoser, Bernd G. ; Schmitt, Robert H.: Application of Multispectral Computed Tomography for the Characterisation of Natural Graphite. In: *e-Journal of Nondestructive Testing* 28 (2023), März, Nr. 3. <http://dx.doi.org/10.58286/27710>. (Zugriff: 2024-04-19). – DOI 10.58286/27710. – ISSN 14354934
- [78] Nadkarni, Rohan ; Allphin, Alex ; Clark, Darin ; Badea, Cristian: Material Decomposition from Photon-Counting CT Using a Convolutional Neural Network and Energy-Integrating CT Training Labels. In: *7th International Conference on Image Formation in X-Ray Computed Tomography* (2022), Oktober, 14. <http://dx.doi.org/10.1117/12.2646405>. (Zugriff: 2024-03-20). – DOI 10.1117/12.2646405
- [79] Gong, Hao ; Tao, Shengzhen ; Rajendran, Kishore ; Zhou, Wei ; McCollough, Cynthia H. ; Leng, Shuai: Deep-learning-based Direct Inversion for Material Decomposition. In: *Medical Physics* 47 (2020), Dezember, Nr. 12, 6294–6309. <http://dx.doi.org/10.1002/mp.14523>. (Zugriff: 2024-03-25). – DOI 10.1002/mp.14523. – ISSN 0094–2405, 2473–4209
- [80] Geng, Mufeng ; Tian, Zifeng ; Jiang, Zhe ; You, Yunfei ; Feng, Ximeng ; Xia, Yan ; Yang, Kun ; Ren, Qiushi ; Meng, Xiangxi ; Maier, Andreas ; Lu, Yanye: PMS-GAN: Parallel Multi-Stream Generative Adversarial Network for Multi-Material Decomposition in Spectral Computed Tomography. In: *IEEE Transactions on Medical Imaging* 40 (2021), Februar, Nr. 2, 571–584. <http://dx.doi.org/10.1109/TMI.2020.3031617>. (Zugriff: 2024-03-25). – DOI 10.1109/TMI.2020.3031617. – ISSN 0278–0062, 1558–254X
- [81] Jost, Gregor ; McDermott, Michael ; Gutjahr, Ralf ; Nowak, Tristan ; Schmidt, Bernhard ; Pietsch, Hubertus: New Contrast Media for K-Edge Imaging With Photon-Counting Detector CT. In: *Investigative Radiology* 58 (2023), Juli, Nr. 7,

- 515–522. <http://dx.doi.org/10.1097/RLI.0000000000000978>. (Zugriff: 2024-05-06). – DOI 10.1097/RLI.0000000000000978. – ISSN 1536–0210, 0020–9996
- [82] Busi, Matteo ; Mohan, K. A. ; Dooraghi, Alex A. ; Champley, Kyle M. ; Martz, Harry E. ; Olsen, Ulrik L.: Method for System-Independent Material Characterization from Spectral X-ray CT. In: NDT & E International 107 (2019), Oktober, 102136. <http://dx.doi.org/10.1016/j.ndteint.2019.102136>. (Zugriff: 2024-02-07). – DOI 10.1016/j.ndteint.2019.102136. – ISSN 09638695
- [83] Schneider, Uwe ; Pedroni, Eros ; Lomax, Antony: The Calibration of CT Hounsfield Units for Radiotherapy Treatment Planning. In: Physics in Medicine and Biology 41 (1996), Januar, Nr. 1, 111–124. <http://dx.doi.org/10.1088/0031-9155/41/1/009>. (Zugriff: 2025-01-03). – DOI 10.1088/0031-9155/41/1/009. – ISSN 0031–9155, 1361–6560
- [84] Segars, W. P. ; Sturgeon, G. ; Mendonca, S. ; Grimes, Jason ; Tsui, B. M. W.: 4D XCAT Phantom for Multimodality Imaging Research. In: Medical Physics 37 (2010), September, Nr. 9, 4902–4915. <http://dx.doi.org/10.1118/1.3480985>. (Zugriff: 2024-05-14). – DOI 10.1118/1.3480985. – ISSN 0094–2405, 2473–4209
- [85] Segars, W. P. ; Mahesh, M. ; Beck, T. J. ; Frey, E. C. ; Tsui, B. M. W.: Realistic CT Simulation Using the 4D XCAT Phantom: Realistic CT Simulation Using the 4D XCAT Phantom. In: Medical Physics 35 (2008), Juli, Nr. 8, 3800–3808. <http://dx.doi.org/10.1118/1.2955743>. (Zugriff: 2024-04-09). – DOI 10.1118/1.2955743. – ISSN 00942405
- [86] Yu, Zhicong ; Noo, Frédéric ; Dennerlein, Frank ; Wunderlich, Adam ; Lauritsch, Günter ; Hornegger, Joachim: Simulation Tools for Two-Dimensional Experiments in x-Ray Computed Tomography Using the FORBILD Head Phantom. In: Physics in Medicine and Biology 57 (2012), Juli, Nr. 13, N237-N252. <http://dx.doi.org/10.1088/0031-9155/57/13/N237>. (Zugriff: 2024-04-10). – DOI 10.1088/0031-9155/57/13/N237. – ISSN 0031–9155, 1361–6560
- [87] Emil Sidky ; Xiaochuan Pan: Report on the AAPM Deep-Learning Spectral CT Grand Challenge. In: Medical Physics 51 (2024), S. 772–785. <http://dx.doi.org/10.1002/mp.16363>. (Zugriff: 2024-04-29). – DOI 10.1002/mp.16363
- [88] Maier, Andreas ; Schebesch, Frank ; Syben, Christopher ; Würfl, Tobias ; Steidl, Stefan ; Choi, Jang-Hwan ; Fahrig, Rebecca: Precision Learning: Towards Use of Known Operators in Neural Networks. (2018), Oktober. <http://dx.doi.org/10.48550/arXiv.1712.00374>. (Zugriff: 2024-05-14). – DOI 10.48550/arXiv.1712.00374
- [89] Goodfellow, Ian J. ; Pouget-Abadie, Jean ; Mirza, Mehdi ; Xu, Bing ; Warde-Farley, David ; Ozair, Sherjil ; Courville, Aaron ; Bengio, Yoshua: Generative Adversarial Networks. (2014), Juni. <http://dx.doi.org/10.48550/arXiv.1406.2661>. (Zugriff: 2025-01-05). – DOI 10.48550/arXiv.1406.2661

- [90] Tzeng, Eric ; Hoffman, Judy ; Saenko, Kate ; Darrell, Trevor: Adversarial Discriminative Domain Adaptation. (2017), Februar. <http://dx.doi.org/10.48550/arXiv.1702.05464>. (Zugriff: 2025-02-15). – DOI 10.48550/arXiv.1702.05464
- [91] Johnson, Justin ; Alahi, Alexandre ; Fei-Fei, Li: Perceptual Losses for Real-Time Style Transfer and Super-Resolution. (2016), März. <http://dx.doi.org/10.48550/arXiv.1603.08155>. (Zugriff: 2024-04-29). – DOI 10.48550/arXiv.1603.08155
- [92] Chen, Hu ; Zhang, Yi ; Kalra, Mannudeep K. ; Lin, Feng ; Chen, Yang ; Liao, Peixi ; Zhou, Jiliu ; Wang, Ge: Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network. In: *IEEE Transactions on Medical Imaging* 36 (2017), Dezember, Nr. 12, 2524–2535. <http://dx.doi.org/10.1109/TMI.2017.2715284>. (Zugriff: 2024-04-30). – DOI 10.1109/TMI.2017.2715284. – ISSN 0278-0062, 1558–254X
- [93] Henke, B.L. ; Gullikson, E.M. ; Davis, J.C.: X-Ray Interactions: Photoabsorption, Scattering, Transmission, and Reflection at $E = 50\text{--}30,000 \text{ eV}$, $Z = 1\text{--}92$. In: *Atomic Data and Nuclear Data Tables* 54 (1993), Nr. 2, 181–342. <http://dx.doi.org/10.1006/adnd.1993.1013>. (Zugriff: 2024-10-27). – DOI 10.1006/adnd.1993.1013. – ISSN 0092–640X
- [94] Deng, Jia ; Dong, Wei ; Socher, Richard ; Li, Li-Jia ; Kai Li ; Li Fei-Fei: ImageNet: A Large-Scale Hierarchical Image Database. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition* (2009), Juni, 248–255. <http://dx.doi.org/10.1109/CVPR.2009.5206848>. (Zugriff: 2025-01-11). – DOI 10.1109/CVPR.2009.5206848
- [95] Vidal, Franck P. ; Garnier, Manuel ; Freud, Nicolas ; Létang, Jean M. ; John, Nigel W.: Simulation of X-ray Attenuation on the GPU. In: *Theory and Practice of Computer Graphics* (2009), 8 pages. <http://dx.doi.org/10.2312/LOCALCHAPTEREVENTS/TPCG/TPCG09/025-032>. (Zugriff: 2025-01-17). – DOI 10.2312/LOCALCHAPTEREVENTS/TPCG/TPCG09/025–032. ISBN 9783905673715
- [96] Pointon, Jamie L. ; Wen, Tianci ; Tugwell-Allsup, Jenna ; Sújar, Aaron ; Létang, Jean M. ; Vidal, Franck P.: Simulation of X-ray Projections on GPU: Benchmarking gVirtualXray with Clinically Realistic Phantoms. In: *Computer Methods and Programs in Biomedicine* 234 (2023), Juni, 107500. <http://dx.doi.org/10.1016/j.cmpb.2023.107500>. (Zugriff: 2025-01-18). – DOI 10.1016/j.cmpb.2023.107500. – ISSN 01692607
- [97] Badal, Andreu ; Badano, Aldo: Accelerating Monte Carlo Simulations of Photon Transport in a Voxelized Geometry Using a Massively Parallel Graphics Processing Unit. In: *Medical Physics* 36 (2009), November, Nr. 11, 4878–4880. <http://dx.doi.org/10.1118/1.3231824>. (Zugriff: 2025-01-19). – DOI 10.1118/1.3231824. – ISSN 0094–2405, 2473–4209
- [98] Reuter, Tamara ; Borges De Oliveira, Fabrício ; Abt, Christian ; Ballach, Frederic ; Bartscher, Markus ; Bellon, Carsten ; Dennerlein, Frank ; Fuchs, Patrick ; Günnewig,

- Olaf ; Hausotte, Tino ; Hess, Jonathan ; Kasperl, Stefan ; Maass, Nicole ; Kimmig, Wolfgang ; Schielein, Richard ; Von Schmid, Mirko ; Suppes, Alexander ; Wagner, Georg ; Watzl, Christia ; Wohlgemuth, Florian: Introduction to “Realistic Simulation of Real CT Systems with a Basic-Qualified Simulation Software - CTSimU2”. In: e-Journal of Nondestructive Testing 28 (2023), März, Nr. 3. <http://dx.doi.org/10.58286/27715>. (Zugriff: 2025-01-18). – DOI 10.58286/27715. – ISSN 14354934
- [99] Plotzki, David ; Hartlaub, Bendix ; Borges De Oliveira, Fabrício ; Reuter, Tamara ; Wohlgemuth, Florian ; Bellon, Carsten ; Hausotte, Tino: The CTSimU Software Toolbox for CT-related Image Processing and Quality Assessment. In: e-Journal of Nondestructive Testing 27 (2022), März, Nr. 3. <http://dx.doi.org/10.58286/26564>. (Zugriff: 2025-01-18). – DOI 10.58286/26564. – ISSN 14354934
- [100] Library of Congress: STL (STereoLithography) File Format Family. <https://www.loc.gov/preservation/digital/formats/fdd/fdd000504.shtml>. (Zugriff: 2024-11-04). Version: September 2019
- [101] Sujar, Aaron ; Meuleman, Andreas ; Villard, Pierre-Frederic ; García, Marcos ; Vidal, Franck P.: gVirtualXRay: Virtual X-Ray Imaging Library on GPU. In: Computer Graphics and Visual Computing (CGVC) (2017), 8 pages. <http://dx.doi.org/10.2312/CGVC.20171279>. (Zugriff: 2025-01-18). – DOI 10.2312/CGVC.20171279. ISBN 9783038680505
- [102] Poludniowski, Gavin ; Omar, Artur ; Bujila, Robert ; Andreo, Pedro: Technical Note: SpekPy v2.0—a Software Toolkit for Modeling X-ray Tube Spectra. In: Medical Physics 48 (2021), Juli, Nr. 7, 3630–3637. <http://dx.doi.org/10.1002/mp.14945>. (Zugriff: 2025-01-18). – DOI 10.1002/mp.14945. – ISSN 0094–2405, 2473–4209
- [103] Schoonjans, Tom ; Brunetti, Antonio ; Golosio, Bruno ; Sanchez Del Rio, Manuel ; Solé, Vicente A. ; Ferrero, Claudio ; Vincze, Laszlo: The Xraylib Library for X-ray–Matter Interactions. Recent Developments. In: Spectrochimica Acta Part B: Atomic Spectroscopy 66 (2011), November, Nr. 11-12, 776–784. <http://dx.doi.org/10.1016/j.sab.2011.09.011>. (Zugriff: 2025-01-18). – DOI 10.1016/j.sab.2011.09.011. – ISSN 05848547
- [104] Van Aarle, Wim ; Palenstijn, Willem J. ; Cant, Jeroen ; Janssens, Eline ; Bleichrodt, Folkert ; Dabrowski, Andrei ; De Beenhouwer, Jan ; Joost Batenburg, K. ; Sijbers, Jan: Fast and Flexible X-ray Tomography Using the ASTRA Toolbox. In: Optics Express 24 (2016), Oktober, Nr. 22, 25129. <http://dx.doi.org/10.1364/OE.24.025129>. (Zugriff: 2024-11-18). – DOI 10.1364/OE.24.025129. – ISSN 1094–4087
- [105] Dask: Dask. <https://docs.dask.org/en/stable/>. (Zugriff: 2025-02-28)
- [106] Canny, John: A Computational Approach to Edge Detection. In: IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-8 (1986), November, Nr. 6, 679–698. <http://dx.doi.org/10.1109/TPAMI.1986.4767851>. (Zugriff: 2025-01-30). – DOI 10.1109/TPAMI.1986.4767851. – ISSN 0162–8828

- [107] Krizhevsky, Alex ; Sutskever, Ilya ; Hinton, Geoffrey E.: ImageNet Classification with Deep Convolutional Neural Networks. In: Communications of the ACM 60 (2017), Mai, Nr. 6, 84–90. <http://dx.doi.org/10.1145/3065386>. (Zugriff: 2025-02-10). – DOI 10.1145/3065386. – ISSN 0001-0782, 1557-7317
- [108] Kong, Fanjie ; Henao, Ricardo: Efficient Classification of Very Large Images with Tiny Objects. (2021), Dezember. <http://dx.doi.org/10.48550/arXiv.2106.02694>. (Zugriff: 2025-02-10). – DOI 10.48550/arXiv.2106.02694
- [109] Dreier, Till ; Nilsson, Daniel ; Espes, Emil: In-Line and at-Line Battery CT Enabled by MetalJet Sources. In: e-Journal of Nondestructive Testing 29 (2024), März, Nr. 3. <http://dx.doi.org/10.58286/29240>. (Zugriff: 2024-07-10). – DOI 10.58286/29240. – ISSN 1435-4934
- [110] Steuwe, Andrea ; Valentin, Birte ; Bethge, Oliver T. ; Ljimani, Alexandra ; Niegisch, Günter ; Antoch, Gerald ; Aissa, Joel: Influence of a Deep Learning Noise Reduction on the CT Values, Image Noise and Characterization of Kidney and Ureter Stones. In: Diagnostics 12 (2022), Juli, Nr. 7, 1627. <http://dx.doi.org/10.3390/diagnostics12071627>. (Zugriff: 2025-01-15). – DOI 10.3390/diagnostics12071627. – ISSN 2075-4418
- [111] Han, Kai ; Wang, Yunhe ; Chen, Hanting ; Chen, Xinghao ; Guo, Jianyuan ; Liu, Zhenhua ; Tang, Yehui ; Xiao, An ; Xu, Chunjing ; Xu, Yixing ; Yang, Zhaojun ; Zhang, Yiman ; Tao, Dacheng: A Survey on Visual Transformer. In: IEEE Transactions on Pattern Analysis and Machine Intelligence 45 (2023), Januar, Nr. 1, 87–110. <http://dx.doi.org/10.1109/TPAMI.2022.3152247>. (Zugriff: 2025-02-03). – DOI 10.1109/TPAMI.2022.3152247. – ISSN 0162-8828, 2160-9292, 1939-3539
- [112] Touvron, Hugo ; Vedaldi, Andrea ; Douze, Matthijs ; Jégou, Hervé: Fixing the Train-Test Resolution Discrepancy. (2019). <http://dx.doi.org/10.48550/ARXIV.1906.06423>. (Zugriff: 2025-03-02). – DOI 10.48550/ARXIV.1906.06423
- [113] Buchmann, F.: Handbuch Radiographie. <https://download.sstmed.com/Tools/Aufnahmehelfer.pdf>. (Zugriff: 2024-12-20). Version: 1996
- [114] Winston, Roland ; Miñano, Juan C. ; Benítez, Pablo: IMAGING APPLICATIONS OF NONIMAGING CONCENTRATORS. Version: 2005. <http://dx.doi.org/10.1016/B978-012759751-5/50009-9>. In: Nonimaging Optics. Elsevier, 2005. – DOI 10.1016/B978-012759751-5/50009-9. – ISBN 978-0-12-759751-5, 219–234
- [115] Mathieu, Michael ; Couprie, Camille ; LeCun, Yann: Deep Multi-Scale Video Prediction beyond Mean Square Error. (2016), Februar. <http://dx.doi.org/10.48550/arXiv.1511.05440>. (Zugriff: 2025-02-12). – DOI 10.48550/arXiv.1511.05440
- [116] Sun, Yukun ; Yuan, Yuebo ; Lu, Languang ; Han, Xuebing ; Kong, Xiangdong ; Wang, Hewu ; Ouyang, Minggao ; Gao, Panlong ; Zheng, Hexing ; Wang, Keming: A Comprehensive Research on Internal Short Circuits Caused by Copper Particle Contaminants on Cathode in Lithium-Ion Batteries. In: eTransportation 13

- (2022), August, 100183. <http://dx.doi.org/10.1016/j.etran.2022.100183>. (Zugriff: 2025-01-26). – DOI 10.1016/j.etran.2022.100183. – ISSN 25901168
- [117] Chen, Wei ; Han, Xuebin ; Pan, Yue ; Yuan, Yuebo ; Kong, Xiangdong ; Liu, Lishuo ; Sun, Yukun ; Shen, Weixiang ; Xiong, Rui: Defects in Lithium-Ion Batteries: From Origins to Safety Risks. In: Green Energy and Intelligent Transportation (2024), November, 100235. <http://dx.doi.org/10.1016/j.geits.2024.100235>. (Zugriff: 2025-01-26). – DOI 10.1016/j.geits.2024.100235. – ISSN 27731537
- [118] Selles, Mark ; Van Osch, Jochen A. ; Maas, Mario ; Boomsma, Martijn F. ; Wellenberg, Ruud H.: Advances in Metal Artifact Reduction in CT Images: A Review of Traditional and Novel Metal Artifact Reduction Techniques. In: European Journal of Radiology 170 (2024), Januar, 111276. <http://dx.doi.org/10.1016/j.ejrad.2023.111276>. (Zugriff: 2025-02-02). – DOI 10.1016/j.ejrad.2023.111276. – ISSN 0720048X
- [119] Sadia, Rabeya T. ; Chen, Jin ; Zhang, Jie: CT Image Denoising Methods for Image Quality Improvement and Radiation Dose Reduction. In: Journal of Applied Clinical Medical Physics 25 (2024), Februar, Nr. 2, e14270. <http://dx.doi.org/10.1002/acm2.14270>. (Zugriff: 2025-02-14). – DOI 10.1002/acm2.14270. – ISSN 1526-9914, 1526-9914
- [120] Joscha Maier: Artifact Correction and Real-Time Scatter Estimation for X-Ray Computed Tomography in Industrial Metrology. Heidelberg, Universität Heidelberg, Dissertation, Juni 2019. https://archiv.ub.uni-heidelberg.de/volltextserver/26701/1/Dissertation_JoschaMaier.pdf. (Zugriff: 2025-02-05)
- [121] Hasan Tercan: Machine Learning-Based Predictive Quality in Manufacturing Processes. Wuppertal, Bergische Universität Wuppertal, Dissertation, 2023. (Zugriff: 2025-02-06)
- [122] Lin, Tsung-Yi ; Maire, Michael ; Belongie, Serge ; Bourdev, Lubomir ; Girshick, Ross ; Hays, James ; Perona, Pietro ; Ramanan, Deva ; Zitnick, C. L. ; Dollár, Piotr: Microsoft COCO: Common Objects in Context. (2015), Februar. <http://dx.doi.org/10.48550/arXiv.1405.0312>. (Zugriff: 2023-06-06). – DOI 10.48550/arXiv.1405.0312
- [123] Condon, Amariah ; Buscarino, Bailey ; Moch, Eric ; Sehnert, William J. ; Miles, Owen ; Herring, Patrick K. ; Attia, Peter M.: A Dataset of over One Thousand Computed Tomography Scans of Battery Cells. In: Data in Brief 55 (2024), August, 110614. <http://dx.doi.org/10.1016/j.dib.2024.110614>. (Zugriff: 2025-02-20). – DOI 10.1016/j.dib.2024.110614. – ISSN 23523409