

Aus dem Lehrstuhl für Computerunterstützte Klinische Medizin
der Medizinischen Fakultät Mannheim
(Direktor: Prof. Dr. rer. nat. Lothar R. Schad)

Digital and Physical Phantoms for Image-guided Interventions

Inauguraldissertation
zur Erlangung des Doctor scientiarum humanarum (Dr. sc. hum.)
der
Medizinischen Fakultät Mannheim
der Ruprecht-Karls-Universität zu Heidelberg

vorgelegt von

Dominik Fabian Bauer, M.Sc.

aus

Schwäbisch Hall

2022

Dekan: Prof. Dr. med. Sergij Goerd
Referent: Prof. Dr. rer. nat. Lothar R. Schad

Digital and Physical Phantoms for Image-guided Interventions

With recent advances in Deep Learning, the need for labeled training data in medical imaging has increased. However, due to privacy laws and because annotation by medical experts is time-consuming, labeled training data are scarce. In this work, digital and physical phantom data is introduced to overcome this data shortage. The phantoms are used to develop image processing algorithms and to validate interventional workflows with a focus on liver interventions. Digital phantom data is used to develop image registration algorithms and a deep learning computed tomography (CT) reconstruction for the mitigation of metal artifacts. Furthermore, physical phantoms are manufactured for the validation of robotic needle guidance systems and to optimize interventional imaging protocols.

In the first part of this thesis, a synthesis framework for multimodal abdominal image data is presented. The generated CT, cone beam CT (CBCT), and magnetic resonance imaging (MRI) dataset is inherently registered and was used to optimize registration algorithms. Compared to real patient data, the synthetic data showed good agreement regarding the image voxel intensity distribution and the noise characteristics.

In the second part, an end-to-end deep learning CT reconstruction technique called iCTU-Net is developed for metal artifact reduction. The network was trained with simulated metal artifact data obtained from a data generation system that was developed in this work. The iCTU-Net was the only investigated method that was able to eliminate metal artifacts. For projection data exhibiting severe artifacts, the iCTU-Net achieved reconstructions with $\text{SSIM} = 0.970 \pm 0.009$ and $\text{PSNR} = 40.7 \pm 1.6$. The best reference method, an image based post-processing network, only achieved $\text{SSIM} = 0.944 \pm 0.024$ and $\text{PSNR} = 39.8 \pm 1.9$.

The third and fourth part focus on the manufacturing of physical phantoms for the validation of interventional workflows. An abdominal phantom incorporating a liver and six liver lesions with varying visibility in CT and MRI was manufactured to validate a standardized oligometastatic disease (OMD) diagnosis workflow. The workflow includes multimodal image acquisition, image segmentation and registration, and robotically assisted liver biopsy. Using similar materials and a similar manufacturing process, a pelvis phantom with a prostate and four prostate lesions was manufactured. The pelvis phantom was used to perform an MRI-guided prostate biopsy using an MRI-compatible robotic system for needle guidance.

The presented work enables the creation of phantom data for the development and validation of a plethora of medical imaging applications. Algorithms for multimodal image registration and CT image reconstruction were successfully developed and physical abdomen and pelvis phantoms for image-guided interventions were manufactured. While the focus of this work was on liver interventions, the presented frameworks can readily be adapted to other body regions.

Digitale und Physische Phantome für Bildgestützte Interventionen

Mit den jüngsten Fortschritten im Bereich Deep Learning ist der Bedarf an annotierten Trainingsdaten in der medizinischen Bildgebung gestiegen. Aufgrund von Datenschutzgesetzen und weil die Annotation durch medizinische Experten zeitaufwändig ist, sind annotierte Trainingsdaten jedoch rar. In dieser Arbeit werden digitale und physische Phantomdaten eingeführt, um diesen Datenmangel zu überwinden. Die Phantome werden zur Entwicklung von Bildverarbeitungsalgorithmen und zur Validierung interventioneller Arbeitsabläufe verwendet, wobei der Schwerpunkt auf Eingriffen an der Leber liegt. Digitale Phantomdaten werden zur Entwicklung von Bildregistrierungsalgorithmen und einer Deep Learning Computertomographie (CT) Rekonstruktion zur Abschwächung von Metallartefakten verwendet. Darüber hinaus werden physische Phantome zur Validierung von robotischen Nadelführungssystemen und zur Optimierung von interventionellen Bildgebungsprotokollen hergestellt.

Im ersten Teil dieser Arbeit wird ein Framework für die Synthese von abdominalen Bilddaten vorgestellt. Der generierte CT, Cone Beam CT (CBCT) und Magnetresonanztomographie (MRI) Datensatz ist inhärent registriert und wurde zur Optimierung von Registrierungsalgorithmen verwendet. Im Vergleich zu echten Patientendaten zeigten die synthetischen Daten eine gute Übereinstimmung hinsichtlich der Intensitätsverteilung der Bildvoxel und der Rauschcharakteristik.

Im zweiten Teil wird eine End-to-End Deep Learning CT-Rekonstruktionstechnik namens iCTU-Net zur Reduzierung von Metallartefakten entwickelt. Das Netzwerk wurde mit simulierten Metallartefaktdaten trainiert, die von einem in dieser Arbeit entwickelten Datenerstellungssystem stammen. Das iCTU-Net war die einzige untersuchte Methode, die in der Lage war, Metallartefakte zu eliminieren. Bei Projektionsdaten mit erheblichen Artefakten erzielte das iCTU-Netz Rekonstruktionen mit $SSIM = 0,970 \pm 0,009$ und $PSNR = 40,7 \pm 1,6$. Die beste Referenzmethode, ein bildbasiertes Post-Processing-Netzwerk, erreichte nur $SSIM = 0,944 \pm 0,024$ und $PSNR = 39,8 \pm 1,9$.

Der dritte und vierte Teil befasst sich mit der Herstellung physischer Phantome für die Validierung von interventionellen Arbeitsabläufen. Zur Validierung eines standardisierten Arbeitsablaufs für die Diagnose oligometastatischer Erkrankungen (OMD) wurde ein abdominales Phantom hergestellt, das eine Leber und sechs Leberläsionen mit unterschiedlicher Sichtbarkeit in CT und MRI enthält. Der Arbeitsablauf umfasst eine multimodale Bildaufnahme, Bildsegmentierung und -registrierung sowie eine robotergestützte Leberbiopsie. Unter Verwendung ähnlicher Materialien und eines ähnlichen Herstellungsprozesses wurde ein Beckenphantom mit einer Prostata und vier Prostataläsionen hergestellt. Das Beckenphantom wurde zur Durchführung einer MRI-gestützten Prostatabiopsie unter Verwendung eines MRI-kompatiblen Robotersystems zur Nadelführung verwendet.

Die vorgestellte Arbeit ermöglicht die Erstellung von Phantomdaten für die Entwicklung und Validierung einer Vielzahl von medizinischen Bildgebungsanwendungen. Es wurden erfolgreich Algorithmen für multimodale Bildregistrierung und CT-Bildrekonstruktion entwickelt und physische Abdomen- und Beckenphantome für

bildgesteuerte Eingriffe hergestellt. Auch wenn der Schwerpunkt dieser Arbeit auf Eingriffen an der Leber lag, lassen sich die vorgestellten Konzepte problemlos auf andere Körperregionen übertragen.

Contents

List of Figures	xiii
List of Tables	xv
1 Introduction and Outline	1
1.1 Motivation	1
1.2 Outline	2
1.3 Citation of Previous Publications	3
2 Theoretical Background	5
2.1 Medical Imaging	5
2.1.1 Magnetic Resonance Imaging	5
2.1.2 Computed Tomography	6
2.2 Deep Learning	9
2.2.1 Convolutional Neural Networks	10
2.2.2 Generative Adversarial Networks	12
2.3 Additive Manufacturing	14
3 Generation of registration ground truth	17
3.1 Introduction	17
3.1.1 Background	17
3.1.2 Related Work	17
3.1.3 Contribution	18
3.2 Materials and Methods	19
3.2.1 Image Synthesis Framework	19
3.2.2 Training and Loss Functions	20
3.2.3 Data	21
3.2.4 Evaluation Metrics	23
3.2.5 Proof of Principle Registration Evaluation	24
3.3 Results	25
3.3.1 Synthetic Images	25
3.3.2 Proof of Principle Registration	26
3.4 Discussion	28
3.5 Conclusions	29
4 Deep Learning CT Reconstruction	31
4.1 Introduction	31
4.2 Materials and Methods	32

4.2.1	iCTU-Net	32
4.2.2	Reference MAR Networks	33
4.2.3	Data Generation	35
4.2.4	Training	36
4.2.5	Evaluation	37
4.2.6	Experiments	38
4.3	Results	39
4.3.1	Ablation Study	39
4.3.2	Input Study	41
4.3.3	Comparison Study	43
4.4	Discussion	45
4.5	Conclusions	46
5	Anthropomorphic Abdomen Phantom	47
5.1	Introduction	47
5.2	Materials and Methods	49
5.2.1	OMD diagnosis workflow	49
5.2.2	Phantom design and construction	49
5.2.3	Image acquisition	53
5.2.4	Image processing	53
5.2.5	Lesion visibility	54
5.2.6	Interventional biopsy experiment	55
5.3	Results	55
5.3.1	Multimodal imaging	55
5.3.2	Interventional biopsy experiment	57
5.4	Discussion	57
5.5	Conclusions	59
6	Anthropomorphic Pelvis Phantom	61
6.1	Introduction	61
6.2	Materials and Methods	62
6.2.1	Phantom design	62
6.2.2	Phantom construction	62
6.2.3	Multimodal imaging	64
6.2.4	MRI-guided biopsy experiment	64
6.3	Results	65
6.3.1	Multimodal imaging	65
6.3.2	MRI-guided biopsy experiment	67
6.4	Discussion	67
6.5	Conclusions	68
7	Summary	69
8	Outlook	73
9	Bibliography	75
10	Appendix	91

10.1 Supplementary Material Chapter 3	91
10.1.1 ASSD and DSC Registration Evaluation Metrics	91
10.1.2 Res-Net network architecture	94
10.2 Supplementary Material Chapter 6	95
11 Publications	97
12 Curriculum Vitae	101
13 Acknowledgements	103

List of Figures

2.1	Generation of X-rays	6
2.2	Filtered back projection	8
2.3	Convolutional and max pooling layers	10
2.4	Deep learning activation functions	12
2.5	GAN network architecture	13
2.6	Additive manufacturing techniques	15
3.1	Schematic of the data simulation framework.	19
3.2	CycleGAN network architecture	20
3.3	Intensity histograms of patient and synthetic images	26
3.4	Evaluation of the proof of principle registrations (DSC)	27
3.5	Results of the proof of principle registrations	28
4.1	Deep learning architectures for metal artifact reduction	34
4.2	Generation of the metal masks	36
4.3	Generation of the metal training dataset	37
4.4	Results of the ablation study	40
4.5	Results of the input study	42
4.6	Results of the comparison study	44
5.1	The envisioned OMD diagnosis workflow	48
5.2	Visualization of the abdomen phantom	51
5.3	Body hull of the abdomen phantom	52
5.4	Components of the abdomen phantom	52
5.5	Imaging of the abdomen phantom	56
5.6	Control scans of the biopsy experiment	58

6.1	Manufacturing process of the phantom	63
6.2	In-bore MRGB experiment	65
6.3	CT, MRI and US scans of the phantom	66
10.1	Evaluation of the proof of principle registrations (ASSD)	92
10.2	Res-Net architecture used for the CycleGAN generators	94
10.3	Dehydrated agarose prostate	95
10.4	Unedited ultrasound scan of the pelvis phantom	95
10.5	DWI measurement of the pelvis phantom	96

List of Tables

3.1	Training data statistics	22
3.2	Evaluation of the synthetic images	25
4.1	SSIM and PSNR evaluation metrics for the ablation study	39
4.2	SSIM and PSNR evaluation metrics for the input study	41
4.3	SSIM and PSNR evaluation metrics for the comparison study	43
5.1	Contrast agent concentration and hardness of the hepatic lesions	53
5.2	Summary of image acquisition parameters	53
5.3	CNR of the liver and the liver lesion types	57
6.1	Summary of the CT and MRI scan parameters	64
6.2	T1-, T2-relaxation times and CT numbers of prostate and lesions	66
10.1	Registration evaluation DSC metric	93
10.2	Registration evaluation ASSD metric	94

1. Introduction and Outline

1.1 Motivation

Metastasis is the spread of tumor cells from a primary tumor to a distant site via the circulatory system [1]. It is the main difficulty in treating cancer and is responsible for up to 90% of cancer-related deaths [2]. The liver is a frequent site for metastases, and liver metastases are more common than primary hepatic tumors [3]. Liver metastases originate from cancers of the gastrointestinal tract, breast, and prostate, as well as neuroendocrine tumors, and sarcomas [4]. Metastatic liver disease is a life-threatening condition with poor life expectancy and prognosis [5]. With palliative treatment methods, local tumor control is possible and survival can be substantially improved [6]. When the number of lesion foci and lesions is small, the disease is referred to as oligometastatic disease (OMD). In this case, local ablative treatment approaches promise improved disease control and clinical outcome [7]. Interventional image guidance using computed tomography (CT), cone beam CT (CBCT), magnetic resonance imaging (MRI), or ultrasound (US) allows to target metastasis in percutaneous ablations. Improvements in imaging techniques have enabled comprehensive detection and reliable delineation of metastases, which is critical for the application of local treatment [8].

This work was conducted as a part of the research campus “Mannheim Molecular Intervention Environment” (M²OLIE). M²OLIE aims to improve the clinical outcome for patients with OMD. Morphological, molecular, and functional information about the metastases is collected for diagnosis, allowing patient-specific treatment planning and minimally invasive local treatment. This is achieved by deploying novel image processing techniques and robotic systems to support biopsies. The focus of M²OLIE is on the treatment of liver metastases, but focal therapy for prostate cancer is also of interest.

The development of techniques for image processing often depends on the availability of suitable data. With the recent advances in deep learning, the need for labeled training data increased [9]. However, labeled training data is scarce in medical image analysis [10]. Reasons include privacy laws and the time- and labor-intensive labeling of data that requires medical experts. This lack of data can be addressed with digital phantom data. They can be used to develop and validate image processing methods, for example, for image registration or CT metal artifact reduction, for which obtaining suitable patient data is particularly difficult. Registration algorithms are needed to fuse multimodal image information. To assess the quality of registration algorithms, multimodal ground truth data is necessary. Due to internal patient motion and non-reproducible patient positioning, this ground truth is usually not available [11]. The visibility of metastases during surgery is often affected by metal artifacts caused by metallic instruments or contrast agents [12]. To train deep learning networks for the reduction of CT metal artifacts, pairs of metal-free and corresponding metal-affected data are needed. This can only be achieved with simulated data or with phantom measurements.

Physical phantoms, on the other hand, are useful for the development of interventional workflows. In addition to the intervention itself, these workflows also include

patient positioning, image acquisition, and image processing. The phantoms are especially suited for needle-based interventions such as biopsies or ablation techniques [13]. Physical phantoms are useful to develop imaging protocols and to validate robotic navigation systems before moving on to clinical studies. This allows to rule out safety concerns, and to fine tune the systems. Users of the navigation systems additionally gain experience and thus further increase safety and decrease procedure time. However, the results of phantom measurements can only be transferred to patients to a limited extent, since the similarity between phantoms and patients is often not sufficient.

The aim of this thesis is to develop phantoms for image-guided interventional applications. Digital phantom data are used for the development of image processing techniques. To overcome data shortage in the development of image registration algorithms, a CycleGAN network architecture is used to generate a multimodal synthetic dataset. Similarly, metal artifacts are simulated in CT phantom data to serve as training data for an end-to-end deep learning CT reconstruction for the mitigation of metal artifacts. Physical phantoms are used to validate robotic needle guidance systems and to optimize interventional imaging protocols. A liver and a pelvis phantom are designed and manufactured and robotically assisted biopsies are performed under CBCT- and MRI-guidance, respectively.

1.2 Outline

This thesis is written cumulatively. In [Chapter 3](#) and [Chapter 4](#), digital phantoms are used to develop image processing and image reconstruction algorithms. The manufacturing process of physical phantoms for interventional imaging is described in [Chapter 5](#) and [Chapter 6](#). Each of these four chapters is a self-contained scientific study.

In [Chapter 2](#), the theoretical background relevant for this thesis is introduced. First, an overview of the basics of MRI and CT is provided. Next, the deep learning fundamentals that are used in the studies in [Chapter 3](#) and [Chapter 4](#) are presented. Finally, the additive manufacturing techniques used to produce the phantoms in [Chapter 5](#) and [Chapter 6](#) are introduced.

In [Chapter 3](#), a framework for the synthesis of a realistic abdominal image dataset is presented. This multimodal CT, CBCT, and MRI dataset is inherently registered and thus serves as a registration ground truth. Because organ masks are automatically provided in the framework, it additionally serves as a ground truth for image segmentation.

In [Chapter 4](#), an end-to-end deep learning CT reconstruction technique for the reduction of metal artifacts is developed. The network is trained with simulated interventional phantom data of the abdomen. The data generation pipeline enables targeted placement of metal objects in specific anatomies. Beam hardening and noise are accurately modelled, resulting in the generation of realistic metal artifacts.

In [Chapter 5](#), a standardized workflow for the diagnosis of OMD is proposed. An abdominal phantom is manufactured to validate the whole workflow, including multimodal image acquisition, image segmentation and registration, and robotically

assisted liver biopsy. The phantom incorporates a liver and six liver lesions with varying visibility in CT and MRI.

In [Chapter 6](#), the manufacturing process of a pelvis phantom is described. The phantom has a hollow rectum that allows transrectal access to a prostate with four prostate lesions. An in-bore MRI-guided biopsy (MRGB) is performed aided by an MRI-compatible remote controlled manipulator (RCM).

[Chapter 7](#) provides a general summary of this thesis and summaries of the four presented studies from [Chapter 3](#) to [Chapter 6](#).

In [Chapter 8](#), future plans of the presented work are discussed.

1.3 Citation of Previous Publications

Several chapters of this thesis have already been published or are currently submitted for publication. The citations for these chapters are:

Chapter 3: D. F. Bauer, T. Russ, B. I. Waldkirch, C. Tönnies, W. P. Segars, L. R. Schad, F. G. Zöllner and A.-K. Golla. Generation of annotated multimodal ground truth datasets for abdominal medical image registration. *International Journal of Computer Assisted Radiology and Surgery*, 16, pp.1277–1285, doi: 10.1007/s11548-021-02372-7, 2021.

Chapter 4: D. F. Bauer, C. Ulrich, T. Russ, A.-K. Golla, L. R. Schad and F. G. Zöllner. End-to-End Deep Learning CT Image Reconstruction for Metal Artifact Reduction. *Applied Sciences*, 12(1), 404, doi: 10.3390/app12010404, 2022.

Chapter 5: D. F. Bauer, J. Rosenkranz, A.-K. Golla, C. Tönnies, I. Hermann, T. Russ, G. Kabelitz, A. J. Rothfuss, L. R. Schad, J. L. Stallkamp and F. G. Zöllner. Development of an abdominal phantom for the validation of an oligometastatic disease diagnosis workflow. Under review (submitted 19.05.2021).

Chapter 6: D. F. Bauer, A. Adlung, I. Brumer, A.-K. Golla, T. Russ, E. Oelschlegel, F. Tollens, S. Clausen, P. Aumüller, L. R. Schad, D. Nörenberg and F. G. Zöllner. An anthropomorphic pelvis phantom for MR-guided prostate interventions. *Magnetic Resonance in Medicine*, 87(3), pp.1605-1612, doi: 10.1002/mrm.29043, 2021.

2. Theoretical Background

2.1 Medical Imaging

Medical imaging methods can be used to obtain morphological and functional information for the diagnosis of diseases. They can also be used for image-guidance during treatment. In this section, the basics of CT and MR imaging are introduced. These tomographic methods provide superposition free cross sections of an imaged object.

2.1.1 Magnetic Resonance Imaging

Certain atomic nuclei have a property called spin, which induces a magnetic moment [14]. Hydrogen MRI is the clinical standard because hydrogen is abundant in the human body. The nucleus of a hydrogen nuclei consists of a single proton and thus acts like a dipole magnet with a spin quantum number of $1/2$. If the nuclei are placed inside a strong external magnetic field B_0 , they will precess around an axis along the magnetic field at the Larmor frequency ω_0 :

$$\omega_0 = \gamma B_0, \quad (2.1)$$

with the gyromagnetic ratio γ . The magnetic moments of the protons statistically align themselves either parallel or anti-parallel to the direction of the external magnetic field. At room temperature, the number protons in the lower-energy (parallel) state slightly exceed the number of protons in the anti-parallel state. For a large number of hydrogen atoms, this results in a considerable net magnetization parallel to the field. This longitudinal magnetization can be perturbed by application of a radio frequency pulse and the resultant signal can be measured with a receiving coil. The resulting transverse magnetization is rotating with the Larmor frequency and induces a measurement signal in a receiving coil. After excitation, the magnetization returns back to its equilibrium state through various relaxation processes. The T_1 relaxation is the exponential recovery of the longitudinal magnetization M_z :

$$M_z(t) = M_{z,0} - (M_{z,0} - M_z(0)) e^{-\frac{t}{T_1}}, \quad (2.2)$$

where $M_{z,0}$ is the initial longitudinal magnetization before the pulse and T_1 is a tissue-specific time constant that characterizes the relaxation. The T_2 relaxation is the loss of phase coherence in the transverse plane. After the application of the radio frequency pulse, the transversal magnetization $M_{x,y}$ decays according to:

$$M_{x,y}(t) = M_{x,y}(0) e^{-\frac{t}{T_2}}, \quad (2.3)$$

where $M_{x,y}(0)$ is the transversal magnetization after the pulse was applied and all the spins are still in phase. T_2 is a tissue-specific time constant that characterizes the relaxation without taking field inhomogeneities into account. The contrast

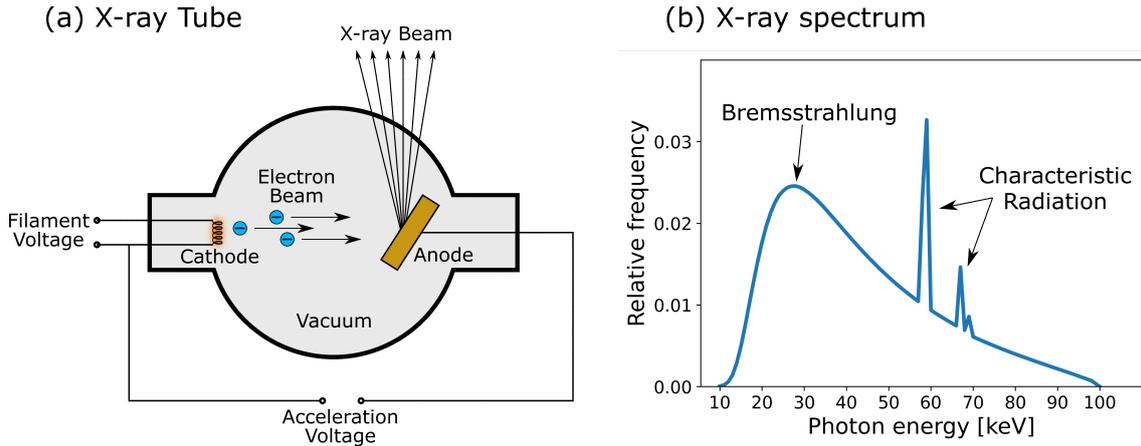


Figure 2.1: Generation of X-rays. (a) Schematic drawing of an X-ray tube. (b) X-ray spectrum with a tube peak voltage of 100 kVp and 1 mm aluminium filtration.

between tissues in MRI is determined by the signal intensities, which in turn are mainly dependent on the T_1 and T_2 relaxation times. An image where the tissue contrast is mainly due to differences in T_1 is called a T_1 -weighted (T1w) image. The same applies for T_2 and T_2 -weighted imaging. For the spatial localization of the MRI signal, magnetic field gradients are required. Linear gradients are applied in three orthogonal directions and the three dimensions are localized via slice-selection, phase-encoding, and frequency-encoding.

2.1.2 Computed Tomography

A CT scanner consists of an X-ray tube and a detector array placed in a gantry. A schematic of an X-ray tube is shown in Figure 2.1 (a). A small voltage of about 10 V is applied to a filament inside the vacuum tube. The resulting current heats the filament, which then emits electrons through thermionic emission. The filament acts as a cathode and the electrons are accelerated towards an anode inside the tube. In CT imaging, typical tube voltages to accelerate the electrons range from about 70 kV to 140 kV [15]. The electrons are decelerated when they collide with the anode material and the lost kinetic energy is emitted again as photons, the so-called bremsstrahlung. About 99% of the energy is converted to heat when the electrons collide with the anode material. Only about 1% of the energy is emitted as bremsstrahlung and characteristic X-rays. Bremsstrahlung is produced by the deceleration of electrons scattered from positively charged atomic nuclei of the anode. Characteristic X-ray radiation is emitted when a bound electron gets ejected from the inner shell of an anode atom after getting struck by an incident electron. The vacancy in the inner shell is filled with an outer shell electron and a photon with an energy equivalent to the energy difference of the low and high states is emitted. The resulting X-ray energy spectrum consists of a continuous bremsstrahlung component and discrete lines of characteristic radiation. Low-energy photons in the spectrum are undesirable because they are mostly absorbed in the patient. Therefore, they would only increase the radiation exposure to patients without contributing to the measured signal. To filter out the low-energy photons, materials such as aluminum or tin are placed in the path of the beam at the front of the tube. The attenuation of

X-ray beams is generally higher for low energies, therefore, proportionally more low-energy photons are absorbed in the filter than high-energy photons. In Figure 2.1 (b) a spectrum obtained with a tube peak voltage of 100 kVp and 1 mm aluminium filtration is shown. The maximum energy of an X-ray spectrum is determined by the used tube voltage according to

$$E = e \cdot U, \quad (2.4)$$

where e is the elementary charge and U the tube voltage. After the X-rays leave the tube, some of the radiation passes through the patient and some is attenuated.

The dominant photon-matter interactions at photon energies used in CT are the photoelectric effect and Compton scattering. The photoelectric effect describes the complete absorption of a photon when it interacts with an electron bound in an atom. The photon transfers all of its energy to the electron and vanishes. The electron is ejected from the shell of the atom. In Compton scattering, a photon collides with a loosely bound electron and loses some of its energy. The photon is not absorbed in this process, but it changes its direction. Both interactions lead to a reduction of the X-ray beam intensity. The initial intensity I_0 of a monoenergetic X-ray beam is exponentially attenuated according to Lambert-Beer's law:

$$I = I_0 e^{-\int \mu(x,y) ds}, \quad (2.5)$$

where the spatial dependent attenuation coefficient $\mu(x, y)$ is integrated along a beam path s . The extension of the Lambert-Beer's law for polyenergetic beams can be seen in equation Equation 4.1. In CT imaging, a set of line integrals are measured with a detector array behind the patient:

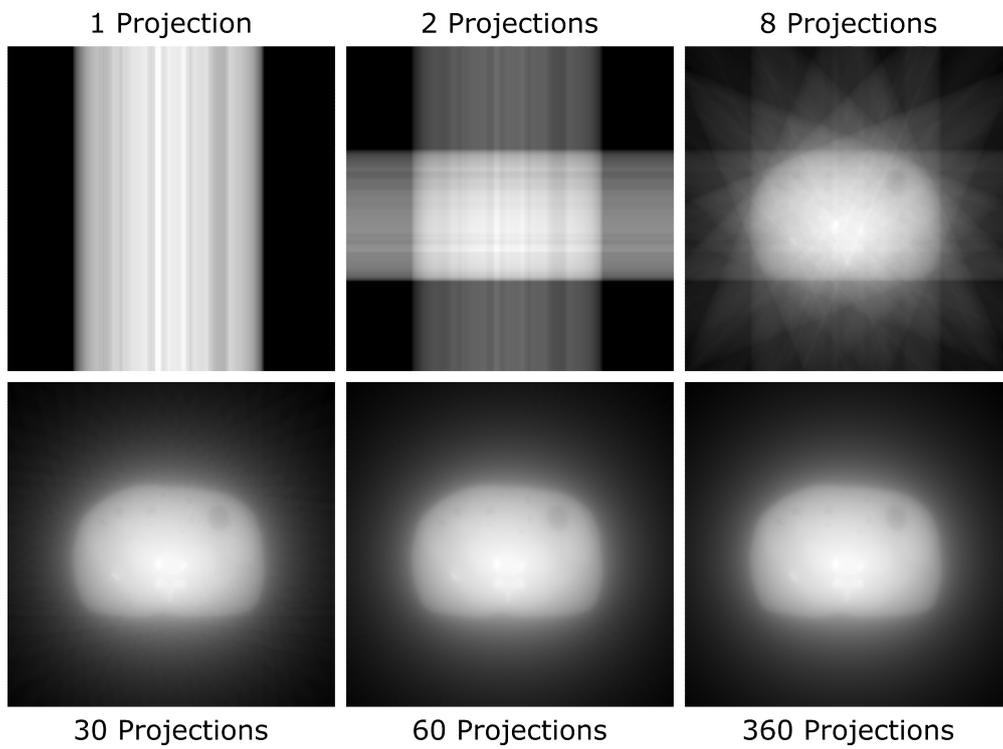
$$p = -\ln\left(\frac{I}{I_0}\right) = \int \mu(x, y) ds. \quad (2.6)$$

For a CT scan, the X-ray tube and detector are rotated around the patient, and multiple X-ray projection measurements from different angles are acquired. At least 180° of data are needed to reconstruct a CT image. The line integrals from different angles are gathered in so-called sinograms. Examples of the sinogram representation of projection data are shown in Figure 4.3. The goal of a CT reconstruction is to recover the spatial distribution of attenuation coefficients $\mu(x, y)$ from the sinogram data. In practice, the radiodensity of CT voxels is given in the Hounsfield scale. Following linear transformation is used to transform attenuation coefficients into CT-numbers:

$$\text{CT-number}(\mu) = \frac{\mu - \mu_{\text{Water}}}{\mu_{\text{Water}}} \cdot 1000 \text{ HU}, \quad (2.7)$$

where μ_{Water} is the absorption coefficient for water. The CT-number is given in Hounsfield units (HU) and water by definition has 0 HU.

(a) Unfiltered Back Projection



(b) Filtered Back Projection

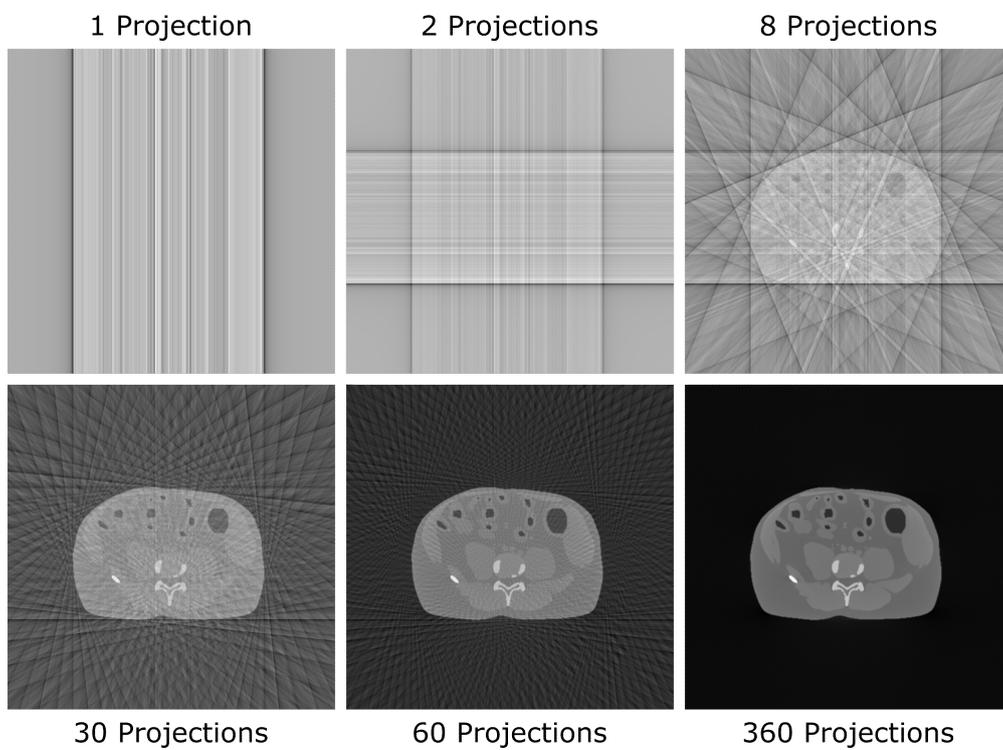


Figure 2.2: CT reconstruction for different numbers of projections. (a) Unfiltered back projection. (b) Filtered back projection with a ramp filter.

Filtered Backprojection

The reconstruction of the CT image from the projection data is an inverse problem. A simple approach to reconstruct CT images is back projection, where the projections for each acquisition angle are smeared back into the image domain at the acquisition angle. Back projections of a digital CT phantom for different numbers of projections are shown in [Figure 2.2 \(a\)](#). The reconstruction of the phantom improves the more projections are used. However, the back projection with one projection already reveals the problem of this method. As the projection is smeared across the whole image, signal is falsely attributed to background pixels. The final image is the sum of all back projections. The signal buildup outside of the phantom leads to a glow around the phantom. Additionally, the reconstruction of the phantom is very blurry.

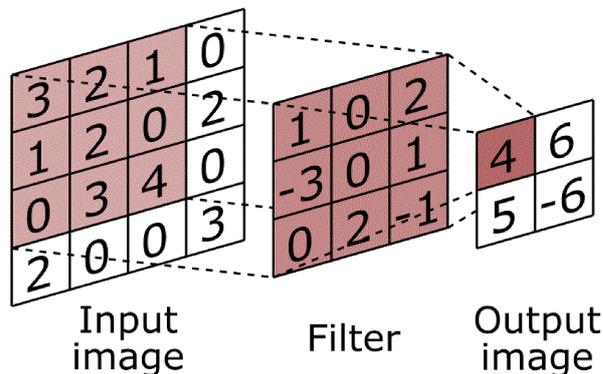
To counteract this blur, a ramp filter has to be applied to the projection data before back projection. The results of this filtered back projection (FBP) are shown in [Figure 2.2 \(b\)](#). The ramp filter suppresses low frequencies, which are responsible for the blurring, and enhances high-frequency edge information. The FBP does not create a glow around the phantom because the ramp filter creates some negative pixels that compensate for the extra smearing caused by the back projection.

However, real CT data contains noise, which primarily consists of high-frequency components. Therefore, the ramp filter also leads to increasing noise in the reconstructed image. Other filters can also be used, such as the Shepp-Logan filter, which has a roll off at higher frequencies. Compared to the ramp filter, this leads to a reduction in noise, but also to a reduction in spatial resolution. In general, the FBP does not perform well when the raw data is noisy or incomplete, which can lead to many kinds of artifacts [16]. For example, if too few projections are acquired, the reconstruction will contain streak artifacts. These can be seen in [Figure 2.2 \(b\)](#) when 60 or fewer projections are used. Metal artifacts are another kind of artifacts that occur when high attenuation objects are present in the scanning field. As shown in [Figure 4.3](#), metal artifacts lead to streak and extinction artifacts that can obscure clinically relevant structures. Noise and many other types of CT artifacts can be reduced to some degree by using iterative reconstruction methods. In iterative methods, the FBP is often used as an initial estimate that is refined over several iterations. Consequently, iterative reconstructions are much more computationally intensive than FBP.

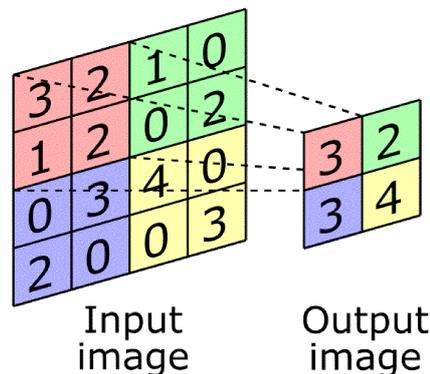
2.2 Deep Learning

Deep learning is a subcategory of machine learning, which itself is a subcategory of artificial intelligence. Deep learning is relevant for a plethora of applications in medical imaging. Prominent examples are image classification, object detection, image reconstruction, image segmentation, image registration, and image generation. Machine learning methods can be divided into supervised learning, unsupervised learning, and reinforced learning. The main difference between supervised and unsupervised learning is the availability of labeled output data. In supervised learning, an objective function, also known as a loss function, is defined between input and

(a) Convolutional layer



(b) Max pooling layer

**Figure 2.3:** A convolutional and a max pooling layer applied to a 4×4 image.(a) 3×3 convolutional layer with a stride of one and without image padding.(b) 2×2 max pooling layer with a stride of two.

output data. Supervised learning algorithms are trained by minimizing the loss function. In unsupervised learning, labeled output data is not available. Unsupervised learning algorithms learn to describe structures and patterns in the input data. In reinforcement learning, an agent interacts with an environment and obtains feedback in form of rewards if certain goals are achieved. The rewards are a measure of the success and failure of the agent. The agent is trying to maximize the reward and improves its performance. An example of reinforcement learning in medical imaging is the development of catheter navigation systems based on phantom measurements [17].

2.2.1 Convolutional Neural Networks

Convolutional neural networks (CNNs) are deep learning algorithms that have become dominant in the field of computer vision. CNNs vary greatly in architecture and complexity as can be seen for the networks used in this work in Figure 4.1 and Figure 10.2. However, their basic components are the same: convolutional layers, pooling layers, and activation functions.

Convolutional Layers

The essential components of CNNs are the convolutional layers. A convolution with a filter kernel is performed with the aim to extract features from an input image. The filter kernels of a network are initialized randomly before training. During training, the entries of the kernels (weights) are updated using backpropagation in order to minimize the loss function [18]. In Figure 2.3 (a) a convolution using a filter kernel with size 3×3 and a stride of 1 is illustrated. The filter kernel slides over the pixels of the input image and the output value is calculated via a dot product. The dot product calculates the sum of the products of the corresponding entries in the filter and the filter-sized input patch. The resulting outputs of a convolutional layer are called feature maps. The stride specifies the number of pixels the filter moves over the input image after calculating the dot product. A stride larger than 1 results in a downsampling of feature map size. If the filter kernel is larger than 1×1 , the

image size shrinks after every convolution operation. Padding of the input image is necessary in order to maintain the image size after convolution. A convolutional layer can have several filters, which each process the input on their own. The number of filters in a convolutional layer determines the number of calculated feature maps, which are also referred to as output channels. The input to a convolutional layer can also have multiple channels. A filter kernel always has the same amount of channels as the input and acts as a three-dimensional convolution over the input channels. Nevertheless, the resulting output for each kernel is a single two-dimensional feature map.

Pooling Layers

Pooling layers are used to downsample the feature maps. Similar to the convolution operation, the pooling operation performs over a sliding window. The two most commonly used types of pooling are max pooling and average pooling. For max pooling, the maximum value in the sliding window is computed, while the average is calculated for average pooling. Figure 2.3 (b) shows a 2×2 max pooling operation with a stride of two. Stride is often used for pooling operations to avoid overlapping windows.

Activation Functions

Networks with nonlinearities are needed to solve nontrivial problems, but convolutions are linear operations. Activation functions are used after convolutional layers to introduce nonlinearities in the network. Below, four activation functions are defined and visualized in Figure 2.4. The most commonly used activation function in CNNs is the rectified linear unit (ReLU), which is a linear function for positive values and otherwise outputs zero.

$$\text{ReLU}(x) = \begin{cases} x & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (2.8)$$

A potential concern with the ReLU activation function is the dying ReLU problem. It occurs when most of the inputs to the ReLU are negative. Since the slope of the ReLU is zero in the negative input range, the gradients do not flow during backpropagation which is why the weights are not updated. The leaky ReLU, which has a small slope for negative values, can be used to mitigate the dying ReLU problem.

$$\text{LeakyReLU}(x) = \begin{cases} x & \text{if } x > 0, \\ \alpha x & \text{otherwise, with } 0 \leq \alpha \leq 1. \end{cases} \quad (2.9)$$

The hard shrinkage activation function sets values between $-\lambda$ and λ to zero and is a linear function otherwise. The threshold value λ is usually very small and the shrinkage function therefore only affects values close to zero.

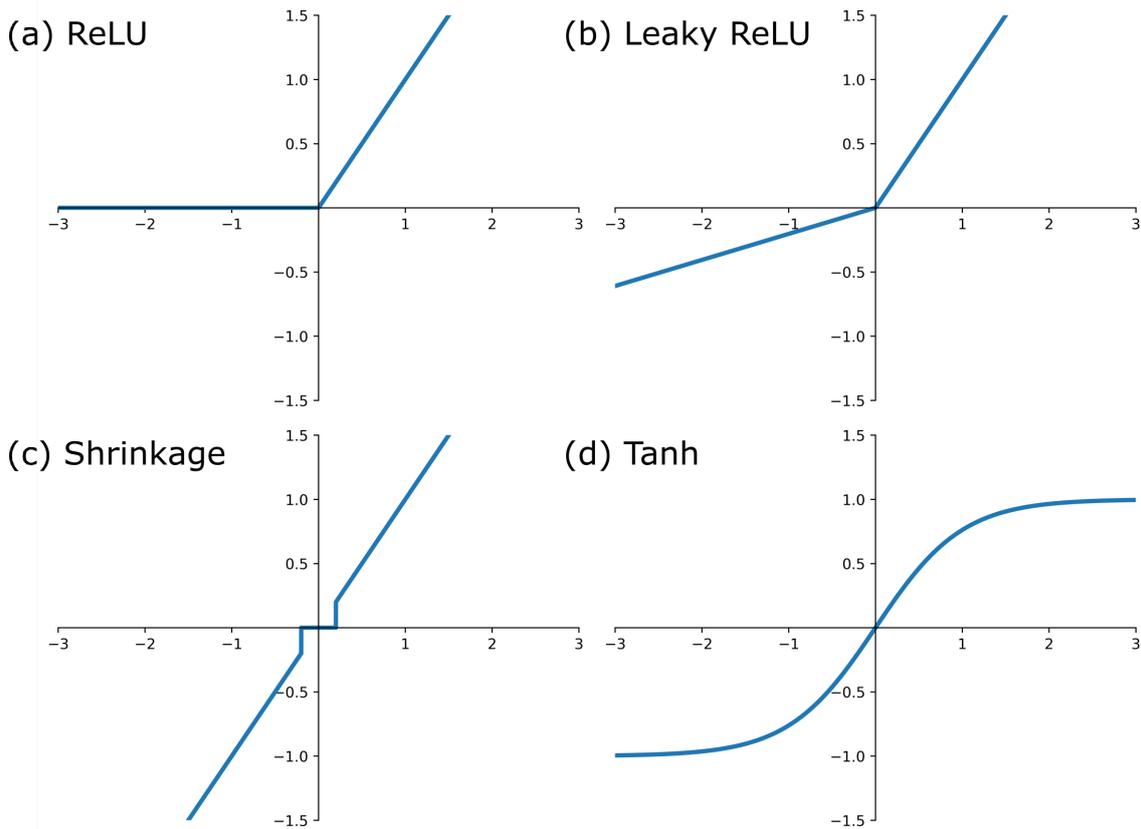


Figure 2.4: Commonly used activation functions in deep learning. (a) ReLU. (b) Leaky ReLU with $\alpha = 0.2$. (c) Shrinkage function with threshold $\lambda = 0.2$. (d) Tanh.

$$\text{Shrinkage}(x) = \begin{cases} x & \text{if } |x| > \lambda, \\ 0 & \text{otherwise.} \end{cases} \quad (2.10)$$

The hyperbolic tangent function (\tanh) is a sigmoid function that maps the input to a range from -1 to 1. Sigmoid activation functions are often used in the last layer of the network to normalize the output. This is particularly useful for classification or segmentation tasks, as the results can then be interpreted as probabilities.

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (2.11)$$

2.2.2 Generative Adversarial Networks

In image-to-image translation, generative models learn to map between an input image domain X to an output image domain Y . In [Chapter 3](#), GANs are used to map XCAT phantom images to CT, CBCT, or MRI images. Since paired training data is not available, the generated images cannot be compared to a label image.

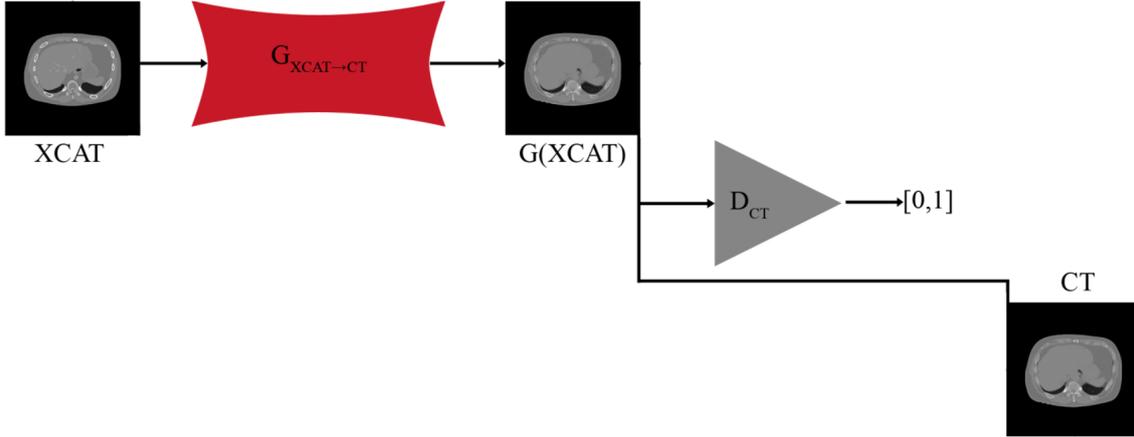


Figure 2.5: GAN network architecture. The generator G generates a fake CT image $G(\text{XCAT})$ from an XCAT phantom input. The discriminator D gets a real CT image and the fake CT image as input.

Therefore, a supervised training approach is not applicable because no distance measure between input and label can be defined as a loss function.

Generative adversarial networks (GANs) are generative models that consist of two CNNs. The architecture of a GAN is shown in Figure 2.5. A generator model generates new images via a mapping function $G : X \rightarrow Y$ and a discriminator model D_Y classifies the generated images as real or fake. The output range of the discriminator is the interval $[0, 1]$, where 0 means that the image is classified as fake, and 1 means that it is classified as real. Instead of using a distance measure for training, the generator is trained to fool the discriminator. This allows to train the network in an unsupervised manner with unpaired training data. The generative adversarial loss function is defined as follows [19]:

$$L_{\text{GAN}}(G, D_Y) = \log D_Y(y) + \log (1 - D_Y(G(x))), \quad (2.12)$$

where $G(x)$ is the output of the generator network given an input image $x \in X$. $D_Y(G(x))$ is the output by the discriminator for the fake image $G(x)$ and $D_Y(y)$ is the discriminator output for a real image $y \in Y$. The generator and discriminator are trained simultaneously in a two-player minimax game:

$$G^*, F^* = \arg \min_G \max_{D_Y} L_{\text{GAN}}(G, D_Y). \quad (2.13)$$

If the generator fails to fool the discriminator, the weights of the generator network get updated such that $D_Y(G(x))$ gets maximized. The generator does not affect the first term in Equation 2.13, thus maximizing $D_Y(G(x))$ means minimizing L_{GAN} . The goal is for the generator to produce images that are more similar to the real images and thus better fool the discriminator. Similarly, the discriminator network gets improved by feeding it real and fake data with the goal to maximizing L_{GAN}

which leads to a minimization of $D_Y(G(x))$ and a maximization $D_Y(y)$. In this way, the discriminator learns to distinguish real images from fake ones. Convergence is reached, when the discriminator is fooled by the generator about half the time, i. e. $D_Y(G(x)) = 0.5$.

CycleGAN

GANs often suffer from what is known as mode collapse, where the generator maps all input images to the same output image. The CycleGAN network addresses this problem with the introduction of a second generator network $F : Y \rightarrow X$ that transforms the generated image $G(x)$ back into the original image domain. The goal is to enforce a cycle consistency, meaning F and G should be inverse to each other: $F(G(x)) \approx x$ and $G(F(y)) \approx y$. This is achieved by introducing the cycle consistency loss:

$$L_{\text{cyc}}(G, F) = \|F(G(x)) - x\|_1 + \|G(F(y)) - y\|_1. \quad (2.14)$$

To train the CycleGAN network, a second discriminator D_X is introduced which aims to distinguish between images x and $F(y)$, just like D_Y aims to discriminate between y and $G(x)$. This effectively means that the CycleGAN consists of two separate GANs with the addition of the cycle consistency. Figure 3.2 shows the complete CycleGAN network architecture for the XCAT phantom and CT image domain. Both generators and both discriminators are trained simultaneously. The network of interest is usually the generator G , while F , D_X and D_Y are only necessary to train G . The loss function of the CycleGAN is a combination of two adversarial losses and the cycle consistency loss:

$$\begin{aligned} L_{\text{CycleGAN}}(G, F, D_X, D_Y) &= L_{\text{adv}}(G, D_Y) \\ &+ L_{\text{adv}}(F, D_X) \\ &+ \lambda_{\text{cyc}} L_{\text{cyc}}(G, F). \end{aligned} \quad (2.15)$$

where λ_{cyc} is a weighting factor that is often set to 10. Typical architectures used for generators are networks with encoding and decoding stages that have equal input and output image sizes. Examples are the U-Net in Figure 4.1 (d) and the Res-Net in Figure 10.2.

2.3 Additive Manufacturing

Additive manufacturing (AM), also known as 3D printing, allows a 3D object to be constructed directly from computer-aided design CAD models [20]. The STL file format is most commonly used to describe a CAD model. In the STL format, the surfaces of the model are approximated by a triangular mesh. In contrast to computer numerical control (CNC), a subtractive process that removes material from a block of material, parts in AM are produced by adding material layer by

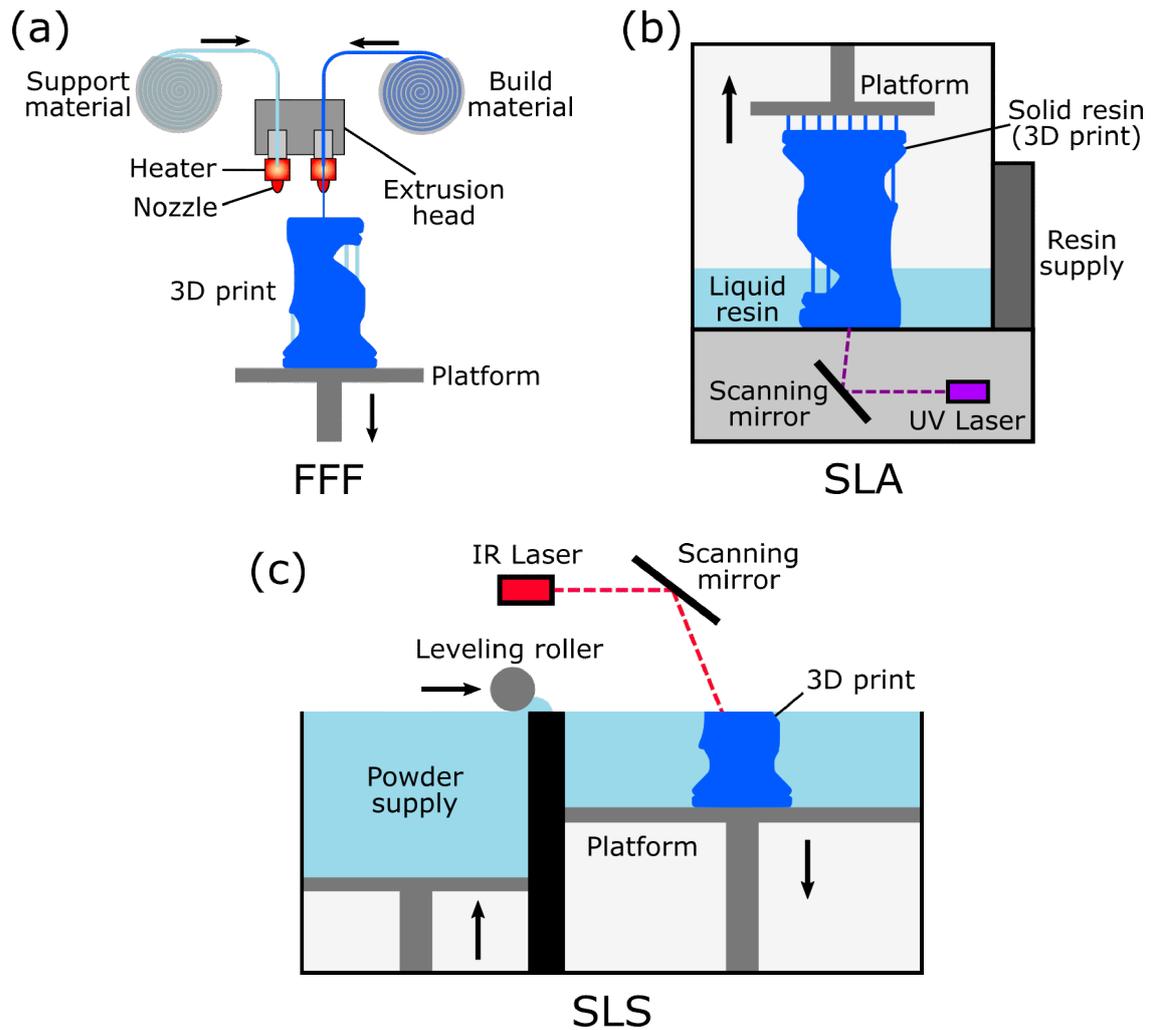


Figure 2.6: Additive manufacturing techniques. (a) Fused Filament Fabrication. (b) Stereolithography. (c) Selective laser sintering.

layer. This approach enables the production of hollow bodies and other complex geometries. The conversion of the STL file into individual layers is done by a slicing software, which also calculates the path within each layer. Originally, AM was used for custom manufacturing (rapid prototyping), but recent improvements in precision and material versatility now make the technology viable for end-use applications [21]. In medicine, AM is used in a wide range of applications, including dentistry and tissue engineering [22]. In medical physics, phantoms for quality assurance and radiation dosimetry measurements are built with AM [23]. Printing patient-specific anthropomorphic phantoms is possible by obtaining 3D models from the patient's own CT or MRI scan.

There are a plethora of AM technologies, but the following sections will only discuss those used in Chapter 5 and Chapter 6. All these technologies are based on the principle of layering, but they differ in the materials used and the way the layers are created.

Fused filament fabrication

Fused Filament Fabrication (FFF), also known as Fused Deposition Modeling (FDM), is an AM process in which the layers of an object are printed by depositing molten material. The configuration of FFF is shown in [Figure 2.6 \(a\)](#). A filament of thermoplastic polymers is fed through an extrusion head, where the material is heated. The molten polymers are then extruded through a nozzle and the layer is built by moving the extrusion head in a predefined path. The different layers of the printed object are created either by moving the extrusion head upwards or by lowering the build platform. FFF requires support structures to create structures with overhangs. Supports are sometimes printed with the same material as the object, but usually support materials with special properties (e.g. water solubility) are used. A second nozzle is integrated in the extrusion head to extrude the support material. Due to the low cost of FFF machines and filaments, FFF is the most commonly used AM technology [24]. Disadvantages of FFF are thick layer sizes and rough surfaces of the printed structures. Smooth surfaces can nevertheless be achieved by grinding and polishing.

Stereolithography

Stereolithography (SLA) was invented in the 1980s and is considered the first AM technology [25]. The concept of SLA is displayed in [Figure 2.6 \(b\)](#). An ultraviolet (UV) laser is used to harden liquid photopolymer resin. The liquid resin is stored in a tank inside the printer which is connected to a resin supply. A layer is cured by scanning the laser with a mirror system. After a layer is completed, the solidified resin is gradually pulled upwards by the build platform. Support structures are required in SLA and are made out of the same resin as the printed object. The support structures are very thin and hardly touch the object. They are easy to remove manually, and after sanding, a smooth surface is obtained. If no support structures are used, sanding is usually not necessary because the prints are already very smooth. Compared to FFF, SLA offers better resolution, but the cost of both the machine and the material is considerably higher.

Selective laser sintering

The concept of selective laser sintering (SLS) illustrated in [Figure 2.6 \(c\)](#) is very similar to SLA. Instead of a liquid resin, polyamide powder is used, which gets solidified by an infrared (IR) laser [26]. This process, in which the powder is transformed into a solid mass by heating without the powder particles melting, is known as sintering. Layers are created by scanning the laser with mirrors, and when a layer is complete, the build platform is lowered. The powder in the construction chamber is replenished by conveying powder from a powder supply with a leveling roller. The resolution of SLS is slightly inferior to that of SLA and the surfaces are somewhat rougher. A major advantage of SLS is that no support structures are needed, as the unsintered powder supports the object at all times. This makes SLS ideal for printing complex geometries.

3. "Generation of annotated multimodal ground truth datasets for abdominal medical image registration", *Int J Comput Assist Radiol Surg, doi: 10.1007/s11548-021-02372-7*

3.1 Introduction

3.1.1 Background

Multimodal imaging plays an important part in the diagnosis of cancers, such as liver cancer [27]. A variety of treatment options are available for hepatocellular carcinoma (HCC), the sixth most common malignancy worldwide and the third leading cause of cancer-related deaths [28]. These include interventional procedures such as transarterial chemoembolizations (TACE) [29] or radioembolization [30]. The treatment planning benefits from using multimodal registration to combine pre- and intrainterventional data. Each imaging modality has strengths and weaknesses. Image registration enables the fusion of complementary information of each modality.

The lack of convenient ground truth data is a major limitation in the field of medical image segmentation and registration [11]. The generation of organ masks for segmentation requires labor-intensive manual annotation. For the development of image registration algorithms (especially for non-rigid image registration methods) and validation of registration accuracy, the ground truth is generally not available [11]. This is because the patient positioning in-between scans usually cannot be reproduced, particularly in the case of multimodal imaging. In the abdomen, the variable content of the bladder and bowel and additional patient motion like respiration and heartbeat further exacerbate the problem.

3.1.2 Related Work

To evaluate registration results or to train deep learning registration approaches, either anatomical multi-label organ masks or landmarks are required. However, generating these labels is labor-intensive, subjective or even impractical for large datasets. Plum *et al.* introduced a semi-automatic framework that allows the creation of a large number of high-quality landmarks with minimal user effort [31]. Established ground truth datasets for the validation of image registration are usually only available for brain imaging. For instance, the simulated brain database (BrainWeb) provides simulated MRI imaging sequences (T1-weighted, T2-weighted,

and proton density) [32]. The images are perfectly aligned, since they are calculated from the same model.

Image synthesis is able to reduce multimodal registration problems to monomodal problems by first converting one modality into the other. Modality reduction has shown improvements in registration accuracy for the brain [33] and the pelvis [34].

For MRI-only radiotherapy planning Wolterink *et al.* demonstrated feasible results using a CycleGAN approach for MRI-to-CT translation and showed that training with unpaired images is superior to training with paired images [35].

Using the digital 4D extended cardiac-torso (XCAT) phantom [36] instead of patient images for image synthesis is beneficial, because organ masks and motion displacement fields are provided by the phantom. Tmenova *et al.* presented a CycleGAN to synthesize X-ray angiograms from the XCAT phantom, which proved to be useful as a data augmentation strategy [37]. Russ *et al.* synthesized abdominal CT images using a CycleGAN and the XCAT phantom [38]. They showed that a vessel segmentation network trained on a combination of real CT and synthetic CT images achieved a superior performance compared to a network trained only on real data. Analytical models [39, 40] and a GAN approach [41] to transform the CT XCAT phantom into cardiac or abdominal MRI images have been developed. To our knowledge, no multimodal registration ground truth dataset of the abdomen created from the same XCAT or digital phantom has been reported in the literature.

3.1.3 Contribution

Our approach to bypass the lack of ground truth data in image registration and segmentation is the generation of a multimodal synthetic dataset from the XCAT phantom. In this work, we focus on multimodal image registration. The synthesis is performed via CycleGAN networks, which were separately trained for each modality. To improve the preservation of high-contrast structures, we extend the CycleGAN generator loss with an intensity loss and a gradient difference loss.

Interventions are often monitored via CBCT, whereas CT and MRI images are taken for diagnosis beforehand to assist the navigation during intervention [29]. Thus, our multimodal dataset consists of T1-weighted MRI, CT, and CBCT images. We use XCAT data in the inhaled and exhaled motion state to synthesize images. Since the same XCAT phantom is used as the starting point for all modalities, the resulting multimodal synthetic data is perfectly co-registered. Displacement fields for respiratory movements and segmentation masks for all organs are provided by the XCAT phantom. Therefore, it serves as a ground truth dataset for registration.

To demonstrate the utility of the multimodal dataset for the optimization of registration algorithms, we evaluate a multimodal non-rigid registration for varying parameter settings. We focus on the registration of the liver, however, the registration quality can be assessed for any other organ.

3.2 Materials and Methods

3.2.1 Image Synthesis Framework

A schematic of our simulation framework is shown in Figure 3.1. Starting from the CT XCAT phantom, CBCT and MRI XCAT versions are generated by applying a FOV mask or by simulating the volume interpolated breathhold exam (VIBE) signal equation [39], respectively. Organ masks for each modality are extracted from the phantoms. Images are synthesized via CycleGAN networks using the XCAT phantom as input. CycleGANs learn the mapping between two domains X and Y given unpaired training samples $x \in X$ and $y \in Y$ [42]. The mapping functions $G : X \rightarrow Y$ and $F : Y \rightarrow X$ are called generators. Two discriminators D_X and D_Y aim to distinguish between real images and generated images. Figure 3.2 shows the complete CycleGAN network architecture for the XCAT and CT image domain. Synthetic CT, CBCT, and MRI images are created via separately trained CycleGAN networks. The cycle consistency loss $L_{cyc}(G, F)$ enforces forward and backward consistency for the generators, i.e. $F(G(x)) \approx x$ and $G(F(y)) \approx y$. With a least squares generative adversarial loss $L_{adv}(G, F, D_X, D_Y)$, the generators are trained to generate images that cannot be distinguished from real images by the discriminator. The discriminators are 70×70 PatchGANs, which are trained with a least squares generative adversarial loss function. For the generators we use a Res-Net architecture with an encoding stage, 9 residual blocks and a decoding stage.

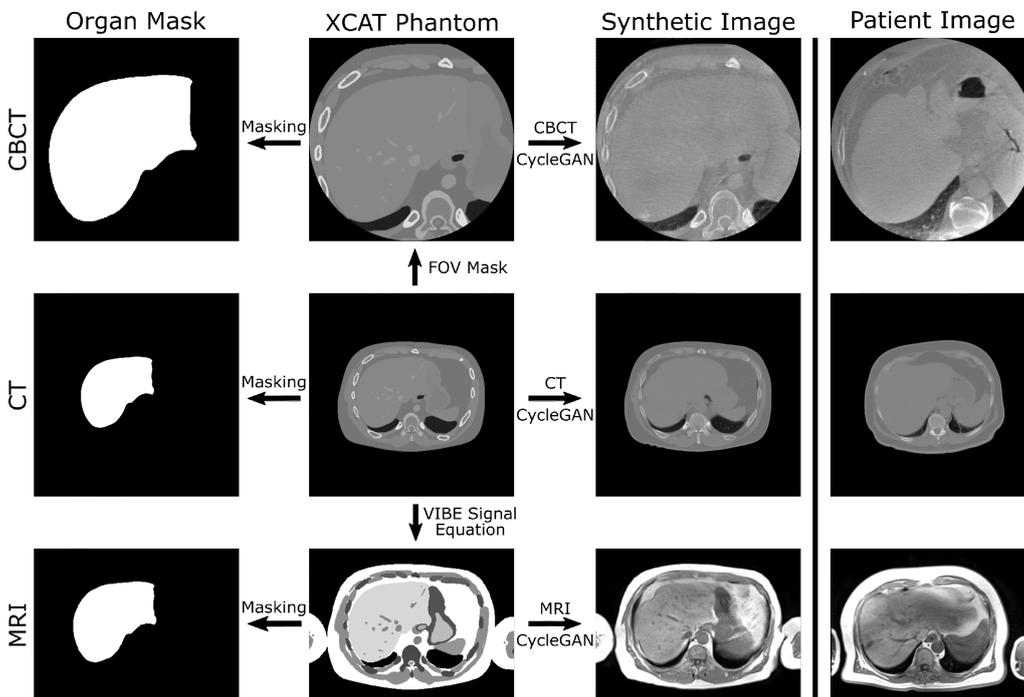


Figure 3.1: Schematic of the simulation framework. Starting point is the CT XCAT, from which CBCT and MRI versions are derived. Synthetic CT, CBCT, and MRI images are created via separately trained CycleGAN networks. Organ masks can be obtained from the XCAT phantoms. Patient images that are used to train the CycleGANs are shown on the right hand side.

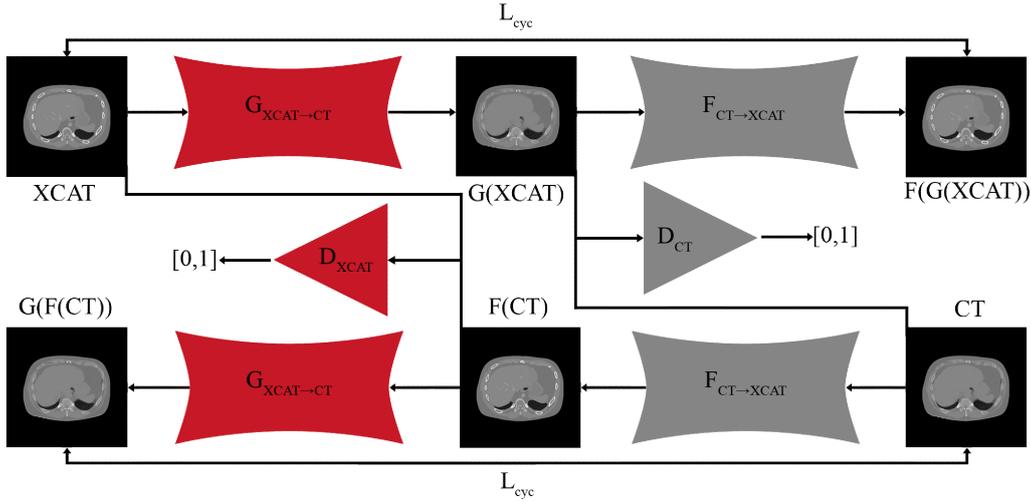


Figure 3.2: CycleGAN network architecture: The generators $G_{\text{XCAT} \rightarrow \text{CT}}$ and $F_{\text{CT} \rightarrow \text{XCAT}}$ map images from the XCAT domain to the CT domain and vice versa. CycleGAN networks for MRI and CBCT images are trained analogously.

3.2.2 Training and Loss Functions

Training is performed using the Adam optimizer with a learning rate of 0.0002. The network is trained with 256x256 pixel image patches and a batch size of 4. For each axial slice one random patch is extracted. We train the networks for 150.000 steps.

To enhance the preservation of high-contrast structures, we extend the generator loss with an intensity loss and a gradient difference loss:

$$L_{\text{int}}(G, F) = \|(G(x) - x)\|_1 + \|(F(y) - y)\|_1, \quad (3.1)$$

$$L_{\text{gdl}}(G, x) = \sum_{i,j} \left| |x_{i,j} - x_{i-1,j}| - |G(x)_{i,j} - G(x)_{i-1,j}| \right|^2 + \left| |x_{i,j} - x_{i,j-1}| - |G(x)_{i,j} - G(x)_{i,j-1}| \right|^2. \quad (3.2)$$

The intensity loss preserves the signal intensity of the organs provided by the XCAT phantom. As shown by Nie *et al.*, the gradient difference loss prevents blurring and therefore sharpens the synthesized images [43]. The total generator loss is a combination of the previously defined losses with different weights:

$$L_{\text{gen}}(G, F, D_X, D_Y) = L_{\text{adv}}(G, F, D_X, D_Y) + \lambda_{\text{cyc}} L_{\text{cyc}}(G, F) + \lambda_{\text{int}} L_{\text{int}}(G, F) + \lambda_{\text{gdl}} (L_{\text{gdl}}(G, x) + L_{\text{gdl}}(F, y)). \quad (3.3)$$

We train three CycleGAN networks for CBCT, CT, and MRI with empirically chosen combinations of weights $\lambda_{\text{cyc}}/\lambda_{\text{int}}/\lambda_{\text{gdl}}$, of 10/10/5, 10/10/5, and 10/0.4/0.4,

respectively. As shown in our previous study, a combination of the gradient loss and the intensity loss yields the best results [44]. Further increasing the weighting factors leads to excessive regularization and thus the networks learns an identity mapping. For the MRI networks lower over-regularization thresholds are found.

3.2.3 Data

We train our CycleGAN network to map between XCAT phantom data and real patient data. The goal is to obtain networks that generate realistic looking synthetic data using the XCAT phantoms as input. In the following paragraphs we will address the real patient and the XCAT training data separately.

3.2.3.1 Patient Data

The patient training data is retrospectively extracted from our Picture Archiving and Communication System (PACS). CT and T1-weighted MRI scans are acquired as part of routine clinical practice before TACE patients undergo a CBCT-guided TACE intervention. All scans are acquired on whole body clinical devices (Siemens Healthineers, Forchheim, Germany; CT: Somatom Emotion 16; CBCT: Artis Zeego; MRI: Magnetom Tim Trio). The MRI images are acquired at 3 Tesla with the VIBE sequence. For each modality the patient images are resampled to a unified voxel spacing given in Table 3.1. All scans include the whole liver and the narrow field of view of the CBCT scan is focused on the liver. The MRI images include arms, the CT and CBCT images do not. As the XCAT phantom does not include the patient couch, we removed the patient couch from the CT patient volumes. For CBCT and MRI no couch is visible in the patient images.

The image intensities are windowed to the ranges given in Table 3.1. For CT and CBCT a fixed window was used. Since MRI intensities vary widely from image to image, the 10th and 90th percentile of each volume (whole 3D matrix) was used for windowing. For training, a linear intensity transformation is applied to transform the intensities from the windowing interval to $[-1,1]$. Normalization of training data is a crucial step in improving training performance, regardless of the normalization method used [45].

Table 3.1: Training data statistics. For CT and MRI the number of slices per image vary in the given interval.

	Parameter	Resolution (x/y/z) [mm]	Windowing	Volumes	Volume size	Arms	Age
CT Patient	100-130 kVp	1/1/2	[-1024, 1500] HU	22	$512 \times 512 \times [52, 151]$	no	66 ± 9
CT Phantom	90-120 kVp	1/1/2	[-1024, 1500] HU	56	$512 \times 512 \times [80, 124]$	no	51 ± 14
CBCT Patient	93-124.7 kVp	0.486/0.486/0.486	[-1024, 2000] HU	24	$512 \times 512 \times 386$	no	67 ± 10
CBCT Phantom	90-120 kVp	0.486/0.486/0.486	[-1024, 2000] HU	56	$512 \times 512 \times 386$	no	51 ± 14
MRI Patient	3 T	1/1/3	[10th , 90th] percentile	24	$330 \times 450 \times [48, 93]$	yes	67 ± 10
MRI Phantom	3 T	1/1/3	[10th , 90th] percentile	56	$330 \times 450 \times [59, 88]$	yes	51 ± 14

3.2.3.2 XCAT Phantom Data

The XCAT model provides highly detailed whole-body anatomies. Organ masks can be easily obtained within the XCAT framework. Since CycleGANs maintain the geometry provided by the XCAT, the organ masks can be used as segmentation masks in the synthesized images. The phantom includes female and male models for varying ages. The heart beat and respiratory motions can be simulated and displacement fields of these motions can be generated. The anatomy and motion can be adapted by various parameters. This allows the creation of highly individual patient geometries. For the XCAT training data we generate one XCAT volume per XCAT model for each modality with 56 different models of varying ages. The XCATs include the whole liver and are generated with the same voxel spacing, windowing and normalization as the resampled patient data. Arms are included only in the MRI XCATs.

The XCAT phantom provides attenuation coefficients for all organs. We vary the simulated tube energy of the CBCT and CT phantoms from 90-120 keV in steps of 5 keV. This leads to a variation of attenuation coefficients in the phantoms. Afterwards, those are transformed into Hounsfield Units. To obtain CBCT and MRI XCAT data, we need to convert the CT XCAT. For the CBCT XCAT we apply a field of view mask obtained from the patient CBCTs, which is centered on the liver. For the MRI phantoms we replace the attenuation coefficients for each organ with simulated MRI values using the signal equation for the VIBE sequence. It ensures that the MRI signal is initialized with realistic values matching the MRI training data. This enables us to use the aforementioned intensity and gradient loss for the generation of synthetic MRI images, since the transformation with the CycleGAN is now monomodal. The signal intensities (SI) for the VIBE sequence in terms of acquisition parameters repetition time TR , echo time TE , and flip angle α and tissue-specific $T1$, $T2$ relaxation times, and proton density ρ is given by:

$$SI = \frac{\rho \sin \alpha (1 - \exp^{-\frac{TR}{T1}})}{(1 - \cos \alpha \exp^{-\frac{TR}{T1}})} \exp \frac{-TE}{T2}. \quad (3.4)$$

We calculate the MRI intensity for all 44 abdominal organs present in the XCAT. The imaging parameters $TE = 4.54$ ms, $TR = 7.25$ ms, and $\alpha = 10^\circ$ are obtained from the patient VIBE scans. The values for the proton density ρ are taken from [40]. $T1$ and $T2$ relaxation times for 3 T for blood and the spinal cord are obtained from [46] and the rest from [47]. For organs with no available $T1$, $T2$ or ρ we use values of similar organs. To simulate some organ variability, we randomly vary $T1$, $T2$, and ρ by $\pm 5\%$ using a uniform distribution.

3.2.4 Evaluation Metrics

Quantification of the synthetic image quality is difficult, since there are no corresponding real images for comparison [37]. Therefore, metrics that require a one-to-one correspondence like the mean absolute error (MAE) cannot be calculated between synthetic and real images. Instead, we calculate one-to-one corresponding

metrics between the synthetic images and the XCATs, to investigate the magnitude of change from the XCAT phantoms. Real patient images and synthetic images are then compared by assessing their noise characteristics and voxel intensity distributions.

3.2.4.1 Synthetic vs. XCAT

The axial slices of the synthetic CT volumes are compared to the corresponding axial slices of the XCAT volumes with respect to anatomical accuracy. The MAE is calculated to assess the change of the intensity values. We exclude the background for the calculation of the MAE. The similarity of structure and features is evaluated using structural similarity index measure (SSIM) and feature similarity index measure (FSIM) [48, 49]. Additionally, we calculate the edge preservation ratio (EPR) and edge generation ratio (EGR) [38, 50].

3.2.4.2 Synthetic vs. Real Patient

Regarding realistic noise characteristics and intensity distribution, the 3D synthetic volumes are compared to the 3D patient volumes. For the noise characteristics, only liver voxels are considered. Limiting the noise considerations to the liver is reasonable, since the liver is a large and mostly homogeneous organ. We manually segmented the liver in 4 patients for each modality. The liver segmentations for the 56 synthetic images are provided by the XCAT phantom. The noise texture is evaluated using an estimation of the radial noise power spectrum (NPS). The radial NPS of the synthetic and patient images is compared by calculating the Pearson correlation coefficient, further called the NPS correlation coefficient (NCC) [38]. In addition to noise texture, we calculate the noise magnitude (NM), i.e. the standard deviation of the liver voxel intensities.

Furthermore, intensity distribution histograms of patient and synthetic images are calculated. To quantify their similarity, the Pearson correlation coefficient between them is calculated (HistCC).

3.2.5 Proof of Principle Registration Evaluation

We perform a proof of principle image registration to demonstrate the feasibility of the multimodal dataset for evaluation and thus development of registration algorithms. Our goal is to investigate different parameter settings to optimize the registration result. We implement the registration in Python 3.5 with SimpleITK 1.2.4. A non-rigid B-spline transform with a gradient descent optimizer, a learning rate of 1 and a maximum of 300 iterations is used. Three different registration metrics are considered, namely Mattes Mutual Information (MMI), Normalized Correlation (NC), and Mean Squares (MS). For the MMI, 50 histogram bins are used. The MS metric is only used for the monomodal CT to CT registration, since it is not suited for multimodal images. Additionally, we vary the spacing of the B-spline control points from 50 mm to 150 mm in steps of 20 mm. For the multimodal and monomodal registrations this results in 12 and 18 different parameter settings, respectively.

The registration is performed on the synthetic data from all 56 XCAT models. We registered the CT, MRI, and CBCT images in the inhaled state to the CT image in

the exhaled state. To evaluate the registration, we take advantage of the liver organ masks obtained from the XCAT phantoms. The veins and arteries inside the liver are included in the liver mask by using a morphological closing operation. We apply the registration transform to the liver masks in the inhaled state and compare the result to the CT liver mask in the exhaled state. The overlap of the two masks is assessed by calculating the Dice similarity coefficient (DSC).

3.3 Results

3.3.1 Synthetic Images

First, we consider the metrics that compare the synthetic images with the XCAT phantoms shown in the upper half of Table 3.2. The FSIM and SSIM indicate that image structures and features were well preserved in the CT and CBCT images, whereas the synthetic MRIs showed little structural and feature similarity to the XCATs. Regarding edges, the EPR is similar for all modalities, whereas the EGR is largest for the CBCT images. The MAE is slightly larger than the NM (synthetic) for every modality. The MAE for CBCT is more than twice as high as the MAE for CT.

Table 3.2: Image quality metrics for the evaluation of the synthetic images.

	CBCT	CT	MRI
SSIM	0.85 ± 0.05	0.94 ± 0.02	0.59 ± 0.04
FSIM	0.82 ± 0.03	0.82 ± 0.02	0.51 ± 0.02
EPR	0.47 ± 0.06	0.43 ± 0.04	0.40 ± 0.03
EGR	3.0 ± 0.7	1.9 ± 0.3	1.7 ± 0.2
MAE	109 ± 14	51 ± 16	37 ± 6
NCC	0.997 ± 0.001	0.980 ± 0.010	0.86 ± 0.04
NM (Synthetic)	52 ± 13	39 ± 5	25 ± 3
NM (Patient)	60 ± 16	39 ± 19	22 ± 5
HistCC	0.994 ± 0.003	0.999 ± 0.002	0.94 ± 0.03

Secondly, we compare the synthetic images to the patient images. The two right columns of Figure 3.1 show axial synthetic and patient slices of each modality. Qualitatively, the style of the synthesized images is in good agreement with the real patient images. To quantify this observation we compared the noise characteristics and voxel intensity distribution of the synthetic images to the patient images, the results are listed in the lower half of Table 3.2. A high NCC for all modalities indicates that the noise texture was emulated realistically, albeit the NCC is slightly smaller for the synthetic MRI images. For all modalities, the NM (synthetic) is in excellent agreement with the NM (patient). In Figure 3.3 the intensity histograms are shown. In general, the synthetic intensity distributions match the patient intensity distributions nicely. This is underlined by the overall high HistCC values in Table 3.2. However, for CT and CBCT the soft tissue peaks are modeled a bit too narrowly. The lung tissue peak is shifted towards higher CT numbers for the CT. In the MRI, the soft tissues is slightly underrepresented.

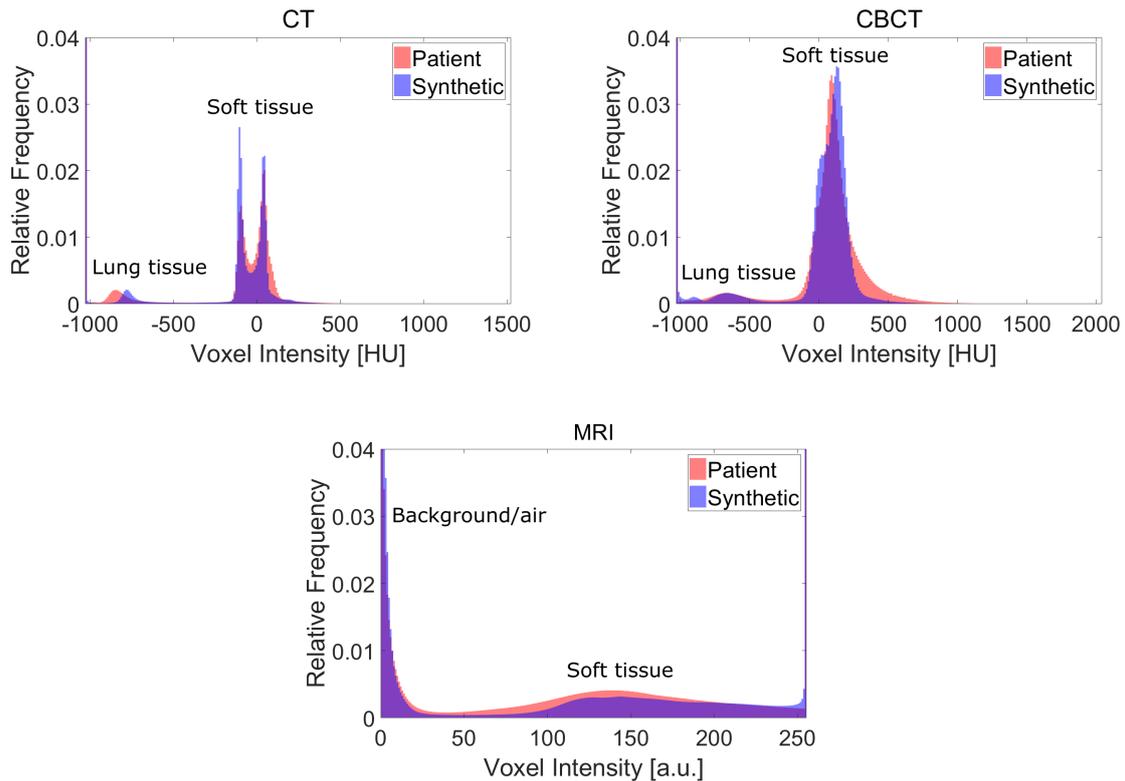


Figure 3.3: Intensity histograms of the patient and synthetic images averaged over all volumes. Note that the background peaks are cropped.

3.3.2 Proof of Principle Registration

The DSC for the evaluation of the proof of principle registration is shown in Figure 3.4. The monomodal CT to CT registration yielded good results for all registration metrics and grid point spacings, with the best result for MMI with 50 mm grid point spacing. For CBCT, the MMI again worked well, whereas the registrations using the NC mostly failed. The best results were again obtained with MMI and a grid point spacing of 50 mm. For MRI, the registrations with MMI and NC yielded similar results with the best result obtained for NC with a grid point spacing of 150 mm. Overall, the monomodal CT to CT registration achieved the best results.

Coronal views of the registration results for the best settings of each modality are visualized in Figure 3.5. The registered images in the middle row show a large similarity to the ground truth. This observation is further supported by the overlaid liver contours. The post-registration liver contour (yellow) is in high agreement with the ground truth liver contour (red).

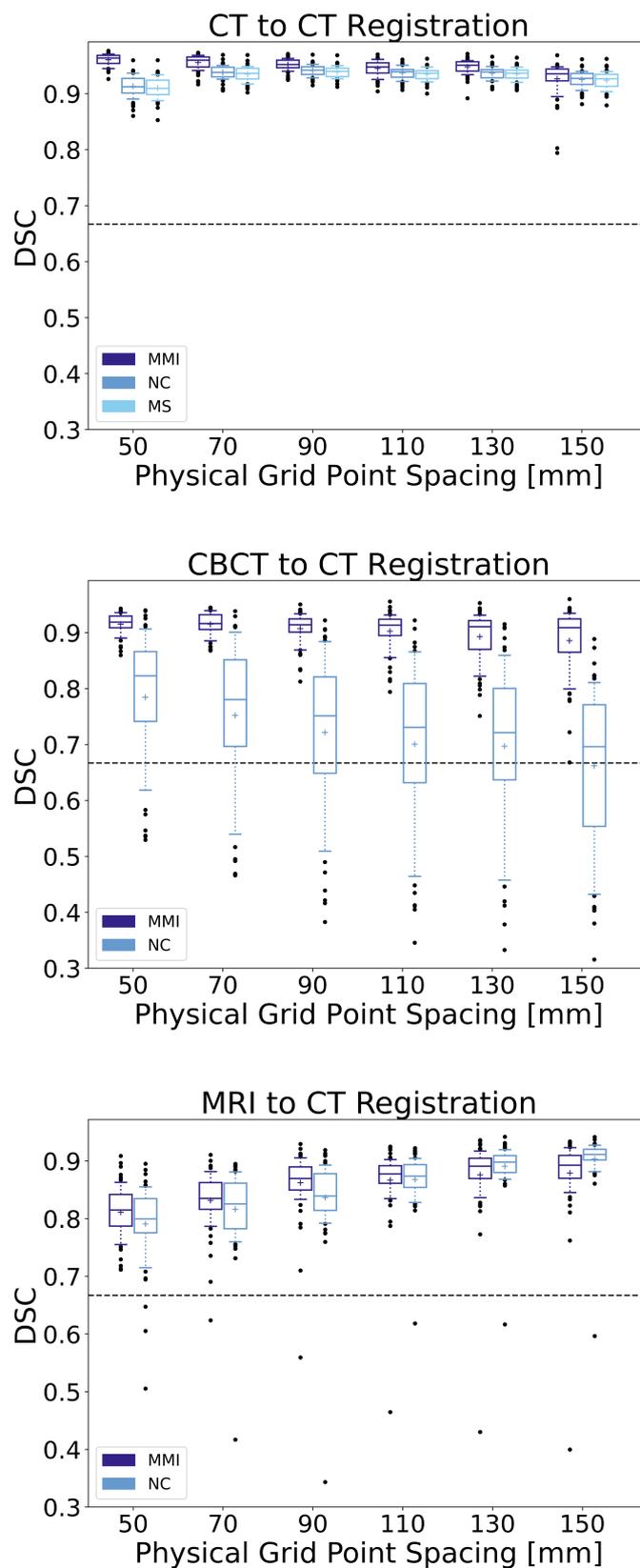


Figure 3.4: DSC for the proof of principle registrations with 56 data points each. The mean is marked as a "+" and the whiskers indicate the 10th and 90th percentile. The dashed horizontal line shows the mean pre-registration DSC.

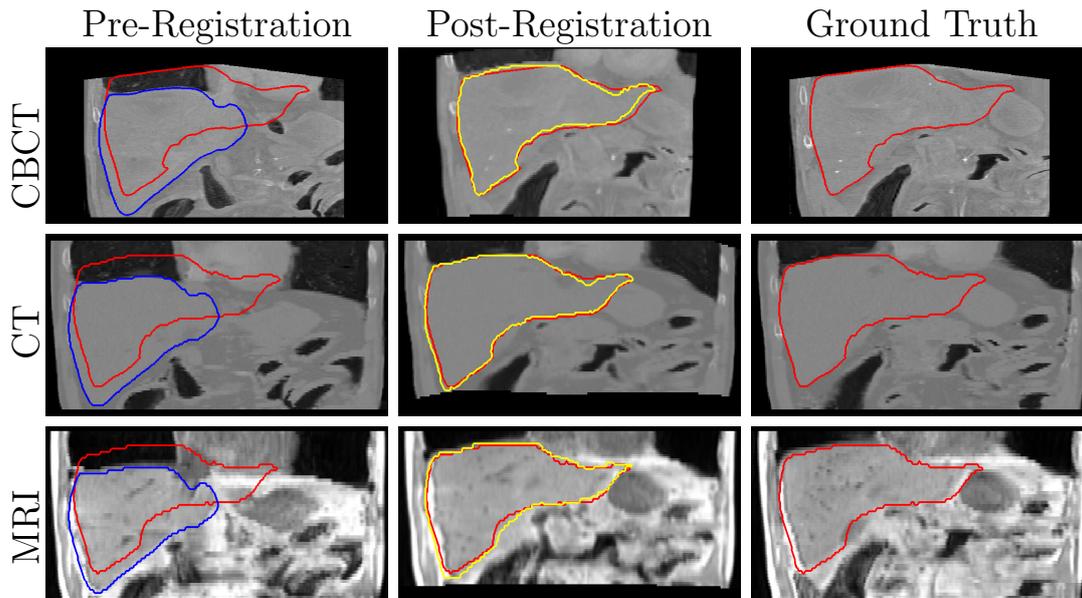


Figure 3.5: Pre- and Post-registered images and the corresponding ground truth. Red contours indicate the ground-truth boundaries of the liver (target). Blue and yellow contours represent the boundaries of the liver before and after deformation respectively.

3.4 Discussion

The image quality metrics in Table 3.2 demonstrate that our framework provides realistic multimodal image data. Low SSIM and FSIM for the MRI images indicate, that image values of the homogeneous organs in the MRI XCAT phantoms needed to be altered more strongly by the networks in comparison to the CT XCATs. The lower SSIM, FSIM, and EPR for MRI are likely a result of lower weighting of the gradient loss and intensity loss.

The ratio of MAE to NM (synthetic) is 2.1, 1.3, and 1.5 for CBCT, CT, and MRI, respectively. Assuming normally distributed noise, the ratio of MAE to NM is only about 0.8 [51]. This means that the MAE cannot be attributed to noise alone. The large MAE for the CBCT images compared to the CT images could be due to the introduction of metal artifacts, as the patient CBCTs showed metal artifacts in the liver caused by medical instruments. This is supported by a large EGR for CBCT.

For all modalities realistic noise texture and magnitude was achieved. Additionally, the voxel intensity distribution was modeled adequately. Most of the discrepancies between the patient and synthetic histograms in Figure 3.3 can be explained by inspecting the XCAT phantoms. The deviation of the CT lung peaks (synthetic -780 HU, patient -835 HU) can be explained by an overestimated initial lung value of -760 HU given by the XCAT. The narrow soft tissue peaks for CT and CBCT could be due to insufficient variation in organ attenuation coefficients. The under representation of soft tissue in the synthetic MRI is due to the body size of the patients and XCATs. We found that in the MRI patient dataset 66.5% of the image voxels show the body, whereas for the MRI XCAT dataset, it is only 46.5%. A rather large HistCC of 0.94 ± 0.03 was still achieved, since this under representation

has only a minor effect on the correlation between the histograms. We prepared the XCAT data such that it matches the patient dataset as good as possible, see Table 3.1. In the future we will consider the patient body size beforehand and adjust the XCAT body size accordingly.

An important requirement for using synthetic data to evaluate or train registration algorithms is a realistic respiratory deformation model. The XCAT framework allows the respiratory rate, the amount of diaphragmatic motion, the amount of chest expansion, and the amount of cardiac motion due to respiration to be varied. In addition, curves controlling the diaphragm motion and chest expansion can be individualized. However, as the breathing patterns of patients are complex and diverse, the XCAT model may still lack generalizability.

The results of the proof of principle registration demonstrate that the synthetic dataset can be used to evaluate different registration algorithms. We were able to evaluate the performance of different registration algorithms and to fine tune parameter settings. It is noticeable that the registration for CT and CBCT works better for small B-spline grid point spacings, while it is the opposite for MRI. Further studies are needed to assess whether this is due to the smaller image size of the MRI data and thus the total number of B-spline control points or to other imaging characteristics of MRI and CT. Computation time can be measured and taken into consideration. For example, registrations with smaller grid point spacings take much longer. Thus, choosing MMI with 150 mm for CT and CBCT registrations might be a reasonable trade-off, as the registration quality is only slightly lower, while the registration time is substantially reduced. The availability of organ masks enabled a rather simple registration evaluation.

3.5 Conclusions

The presented simulation framework can be used to extend small datasets by transferring the style of the dataset onto the geometry given by the XCAT phantom. The obtained datasets can serve as a ground truth for image registration. A multimodal dataset consisting of T1-weighted MRI, CT, and CBCT images was created and used to demonstrate the refinement and evaluation of multimodal image registration algorithms. In the future, the framework will be extended to other modalities, such as T2-weighted MRI or PET, which can further boost the performance of multimodal methods. An extension to other body regions, such as the thorax or pelvis, is also possible. Synthetic images over larger body regions are especially interesting for whole body segmentation. Expansion of datasets using this method provides a promising tool to overcome the dearth of medical training data.

4. ”End-to-End Deep Learning CT Image Reconstruction for Metal Artifact Reduction”, *Appl Sci*, doi: *10.3390/app12010404*

4.1 Introduction

The presence of high attenuation objects in the scanning field leads to artifacts in CT imaging, which substantially decrease the image quality. The generic term for these kinds of artifacts is metal artifacts, which are a combination of beam hardening, scattering, photon starvation, and edge effects [16]. Metal artifacts are common in CT-guided interventions due to the presence of metallic instruments such as biopsy needles [12, 52, 53] or catheters [54]. In many interventions iodine contrast agent is used, leading to additional beam hardening [55]. These artifacts often obscure clinically relevant structures, which can complicate the intervention. For example, the visibility of liver lesions is significantly reduced during liver biopsy [12] or during transarterial chemoembolization (TACE) [56, 57], where catheters are used in combination with contrast agents.

Several CT reconstruction methods have been developed to improve image quality in the presence of metal objects. Statistical iterative reconstruction techniques can be used to correct beam hardening and thus mitigate metal artifacts [58]. Furthermore, dual-energy CT allows to reconstruct virtual monoenergetic images at high kiloelectron volt levels, which substantially reduces metal artifacts [54, 59]. The most common type of metal artifact reduction (MAR) methods is based on inpainting projection data that has been affected by metal. In these approaches, the metal objects are first automatically detected (e.g. via thresholding) in the uncorrected CT image. The metal objects are then forward projected into sinogram domain to obtain a metal trace. The projection data in this metal trace are treated as missing data and are interpolated, e.g. via linear interpolation (LIMAR) [60]. Meyer *et al.* proposed a modification of the LIMAR approach called normalized MAR (NMAR) [61]. NMAR uses a forward projection of an image prior to flatten the uncorrected sinogram before interpolation. This additional step smoothes the sinogram, which reduces the streak artifacts caused by interpolation. In NMAR, the image prior is obtained by identifying air, soft tissue, and bone in either the uncorrected CT or pre-corrected LIMAR image.

With the rapidly increasing popularity of deep learning in medical imaging in recent years [62], a plethora of novel MAR methods have emerged. Deep learning networks are mostly trained in a supervised manner and thus require a metal-free and a corresponding metal-affected dataset. These metal-affected data are commonly synthesized by inserting metallic objects into the metal-free data. Zhang *et al.* presented a CNN called CNN-MAR, that outputs an improved image prior [63]. This

image prior is forward projected and the resulting sinogram data are used to fill in the metal trace in the original sinogram. Several CNN approaches that operate in the sinogram domain have been introduced [64–66]. Lossau *et al.* developed a sophisticated sinogram inpainting approach that works in the presence of motion. A segmentation network identifies the metal trace in the projection domain, a second network fills in the missing sinogram data, and after reconstruction, a third network reinserts the metal objects in the corrected image [67]. A popular class of deep learning MAR techniques are image-based CNNs. They take the uncorrected images as input and either learn a direct mapping to the artifact-free images [63, 68, 69] or to the artifact residuals [70]. These image-based methods often rely on input data that has already been pre-corrected to produce reasonable results [63, 68]. Another option for MAR in the image domain is unsupervised image-to-image translation, which has the advantage that no synthesized metal artifacts are necessary and thus training can be conducted with unaltered clinical data [71–73]. Compared to supervised models, unsupervised models can achieve similar performance on synthetic data [71]. Lin *et al.* recently proposed an end-to-end trainable network called Dual Domain Network (DuDoNet) [74]. It consists of a sinogram enhancement network and an image enhancement network, which are connected by a Radon inversion layer (RIL). The RIL reconstructs the CT images using the filtered back projection (FBP) and allows gradient propagation during training.

In this work, we present an end-to-end deep learning CT reconstruction called iCTU-Net for the correction of metal artifacts. The network learns the mapping from the metal-affected sinograms to the artifact-free images. It consists of three parts which are trained simultaneously: sinogram refinement, back projection, and image refinement. To our knowledge, we are the first to train a single end-to-end deep learning network for the task of reducing metal artifacts with a learnable back projection operation. Since the whole reconstruction process, including the back projection, is learned, the network is able to freely adapt the reconstruction to the imperfections of the sinogram data. The reconstruction is trained in a supervised manner with simulated interventional training data. We focus on liver interventions, thus we generate abdominal liver data including metal objects. We compare our iCTU-Net to the classical NMAR algorithm and to a sinogram refinement and an image refinement deep learning network. Both of these networks employ the same U-Net architecture that is used in our network, which allows a fair comparison. These reference networks were selected to investigate the performance of deep learning MAR approaches in three different domains: sinogram pre-processing, image post-processing, and reconstruction.

4.2 Materials and Methods

4.2.1 iCTU-Net

The design of our iCTU-Net displayed in Figure 4.1 (a) is based on the iCT-Net by Li *et al.* [75], which in turn is inspired by the classical FBP. The reconstruction is trained end-to-end, i. e. the inputs of the iCTU-Net are sinograms and the outputs are reconstructed images. The network includes pre-processing layers and aims to emulate the filtration of the sinograms and the back projection into the image

domain. Post-processing layers were used to further refine the reconstruction. The network performs the complete CT image reconstruction and does not require a predefined back projection operator or the exact X-ray beam geometry.

In a first step, disturbances in the raw measurement data, such as excessive noise, are supposed to be suppressed using 3×3 convolutions (refining layers). The corrected sinogram is then filtered via 10×1 convolutions (filtering layers). By using 1×1 convolutions after the refining and filtering layers and by applying padding in all convolutions, the refined and filtered sinogram maintains the same size of the input sinogram. The convolutions in the refining layers employ a shrinkage activation function with a threshold of 0.0001 [75]. For the filtering layers a tanh activation function is used. Afterwards, the refined and filtered sinogram is projected into the image space in a back projection step. This is realized by a $d \times 1$ convolution with N^2 output channels without padding, where d is the number of detector elements and N is the output image size. This convolution connects every detector element with every pixel in the image space. Since the back projection is learned, sinograms acquired with different beam geometries can be used to train the network, such as parallel beam and fan beam. Then, the results for each view angle v are reshaped to images of size $N \times N$ and rotated according to the acquisition angle. The acquisition angle of the projections is the only geometrical information provided to the network. The rotated images are linearly interpolated and cropped to maintain an image size of $N \times N$. The back projected image is then obtained by combining all views with a 1×1 convolution using a leaky ReLU activation function [76]. Finally, the image output is further refined by a U-Net. The U-Net is a popular choice for post-processing to reduce artifact in CT imaging [77].

4.2.2 Reference MAR Networks

To compare our iCTU-Net to other methods, we implement two deep learning MAR algorithms similar to those of Gjestebj *et al.*. Both networks use pre-corrected NMAR inputs [66, 68]. One is based in the projection domain (U-Net Sino) and the other one in the image domain (U-Net Image). To ensure comparability, we use the same U-Net architecture in the iCTU-Net, U-Net Sino, and U-Net Image. In the U-Net Sino, the sinograms are first refined by a U-Net and the result is then reconstructed using the FBP [66]. In the U-Net Image, the sinograms are first reconstructed with the conventional FBP and then refined with a U-Net [68]. These reference networks were chosen to allow a comparison of sinogram pre-processing, image post-processing, and reconstruction deep learning MAR techniques.

The U-Net architecture is shown in Figure 4.1 (d) and is similar to the original U-Net by Ronneberger *et al.* [78]. It has four en- and decoding blocks consisting of 3×3 convolutions, which are connected via skip connections. Zero-padding is used in the convolutions to ensure that the network output is the same size as the network input. The blocks of the top level have 32 channels, which are doubled with each encoding block until the lowest block has 512 channels. Downsampling in the contracting path is performed via 2×2 max-pooling with stride 2, while upsampling in the expansive path is accomplished using 3×3 transposed convolutions with stride 2. All convolutional layers are followed by a ReLU activation function.

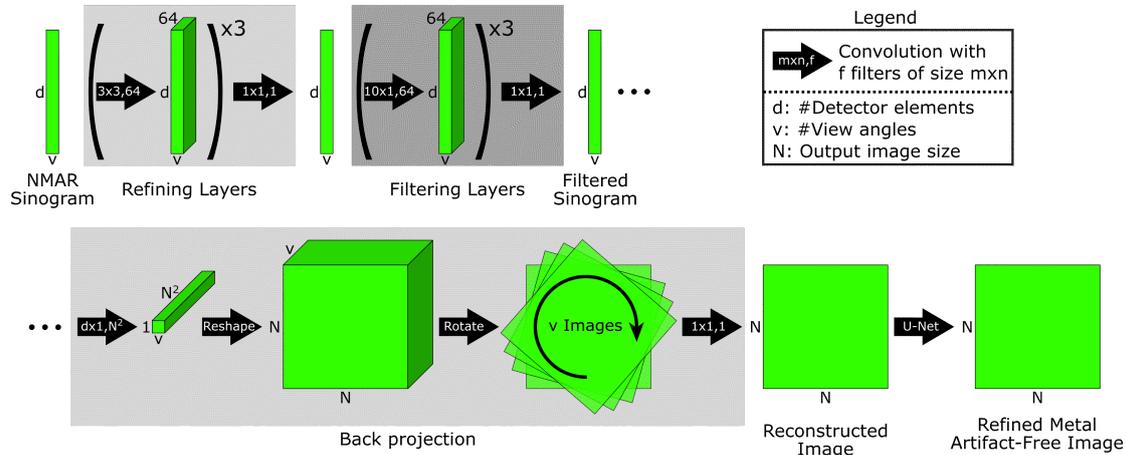
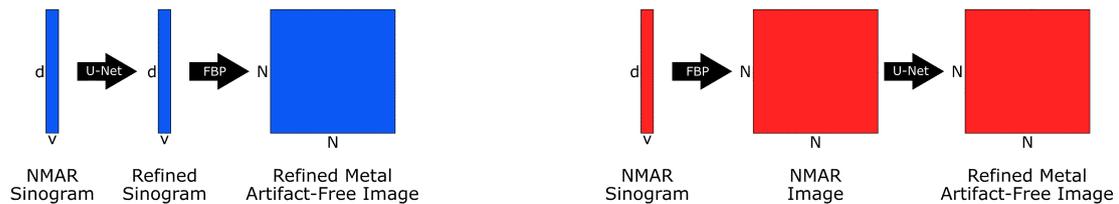
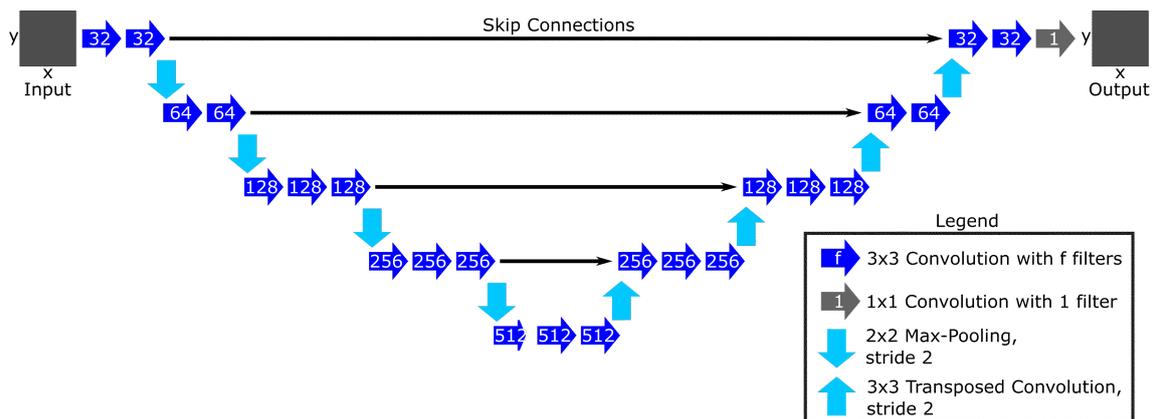
(a) iCTU-Net (End-to-End)**(b) U-Net Sino (Sinogram Domain) (c) U-Net Image (Image Domain)****(d) U-Net Architecture**

Figure 4.1: Deep learning architectures for metal artifact reduction. (a) iCTU-Net end-to-end CT reconstruction network architecture. (b) Sinogram domain U-Net. (c) Image domain U-Net. (d) U-Net network architecture used in the three networks in (a)-(c).

4.2.3 Data Generation

To simulate the training data, we use the XCAT phantom, which provides highly detailed whole-body anatomies [36]. The phantom includes female and male models of different ages, providing a wide variety of patient geometries. Further customization of anatomies by changing organ sizes is possible. We create 40 different XCAT models for training and 10 additional models for testing, resulting in 3964 and 991 slices of size 512×512 pixel with an in-plane resolution of $1 \times 1 \text{ mm}^2$ and a slice thickness of 2 mm, respectively. Because we choose liver interventions as a use case, we generate abdominal XCATs that include the whole liver.

Organ masks can be easily obtained within the XCAT framework. Utilizing these organ masks, we insert metal structures inside the veins of the XCAT phantoms, emulating contrast agents or interventional instruments, such as catheters. Metal objects are only placed inside thicker blood vessels and have a uniform size, independent of the blood vessel size. This is realized by first eroding the blood vessel masks of the XCAT phantom, using a disk with a radius of 3 pixels as a structuring element. The erosion is performed to exclude the smallest blood vessels. To obtain the final metal mask, we skeletonize the mask and then increase the thickness via dilation using a disk with a radius of 3 pixels. An example is shown in Figure 4.2, with the metal mask in red, the initial blood vessels in white, and the liver in green. Most of the metal structures are placed inside the liver or in the portal vein beneath the liver.

Our data generation pipeline is shown in Figure 4.3, which starts with the generation of the ground truth data in the first row. First, we create sinograms by forward projecting the XCAT image data using a parallel beam geometry with 736 projection beams and 360 projection angles. A polychromatic x-ray spectrum and the energy-dependence of the absorption coefficients are considered in the forward projection:

$$I = \sum_{i=1}^N I_0 \cdot \eta(E_i) \cdot e^{-\int_L \mu(x, E_i) dx}, \quad (4.1)$$

with weights of the energy spectrum $\eta(E_i)$. An incident flux of $I_0 = 4 \cdot 10^6$ photons is used, which is slightly increased compared to clinical levels [79], to combat photon starvation due to the presence of the metal objects. The x-ray energy spectrum illustrated in Figure 2.1 (b) is generated using the SpekCalc software with a tube peak voltage of 100 kVp and 1 mm aluminium filter [80]. We use 91 energy bins from 10 keV to 100 keV with a uniform size of 1 keV. The organ masks provided by the XCAT framework make it possible to assign an energy-dependent attenuation coefficient $\mu(x, E_i)$ to each organ. The sinograms are then reconstructed via a FBP. Since the energy dependence of the attenuation coefficients is accounted in the forward projection, beam hardening is present in the ground truth data.

To simulate data affected by metal, we utilize the previously mentioned metal mask to insert the attenuation coefficient of iron. Afterwards metal sinograms are created via forward projection using Equation 4.1. Noise is then added and the projection data is normalized and the negative logarithm is applied:

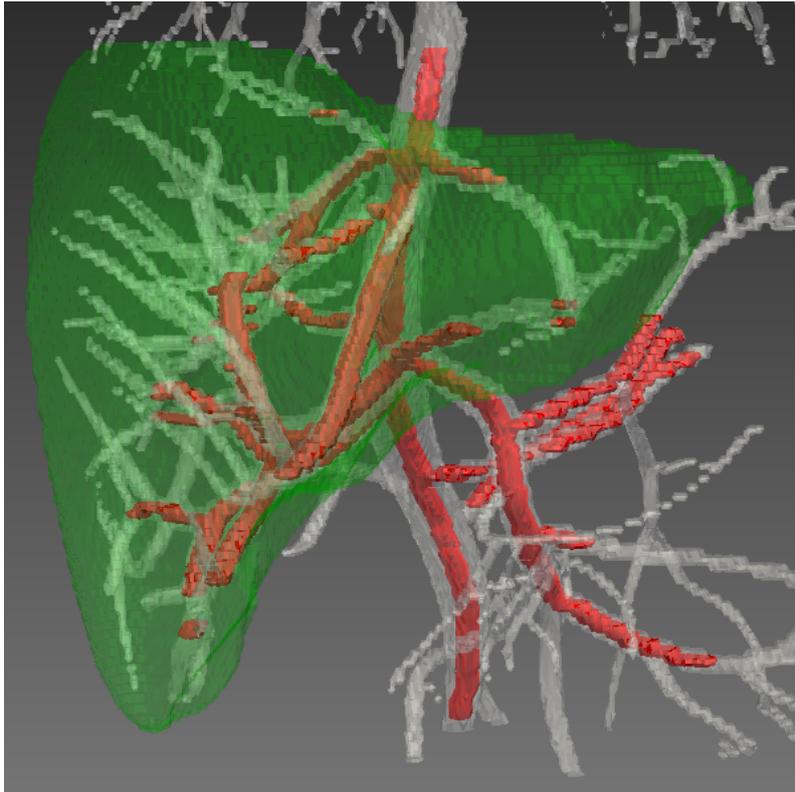


Figure 4.2: Generation of the metal masks inside blood vessels. The metal masks are shown in red, the initial blood vessel segmentation in white, and the liver in green.

$$p_n = -\ln \left(\frac{\text{Poisson}(I) + \mathcal{N}(0, \sigma^2)}{\sum_{i=1}^N I_0 \cdot \eta(E_i)} \right). \quad (4.2)$$

The photon production, attenuation, and detection is described by a Poisson distribution. Electronic noise of the detector is simulated with a Gaussian distribution \mathcal{N} with a mean value of zero and $\sigma^2 = 40$ [81, 82]. A subsequent FBP results in an image containing metal artifacts. As input for the training of our networks we do not use this artifact image, but data pre-corrected with the NMAR algorithm as shown in the third row of Figure 4.3. The prior image used for the normalization in NMAR is obtained by segmentation of soft tissue and bone in a LIMAR image [60, 61].

4.2.4 Training

The networks are trained with the SSIM loss function using the Adam optimizer with a learning rate of 0.001 [83]. We apply L2 regularization on the network weights, with a weighting factor of 10^{-6} . Each network is trained for 25 epochs. The training data in image domain is windowed to $[-1000, 1000]$ HU and then mapped to the interval $[-1, 1]$. The whole image slices are used for training, no patches are extracted. The sinogram training data is neither windowed nor normalized. The input and label

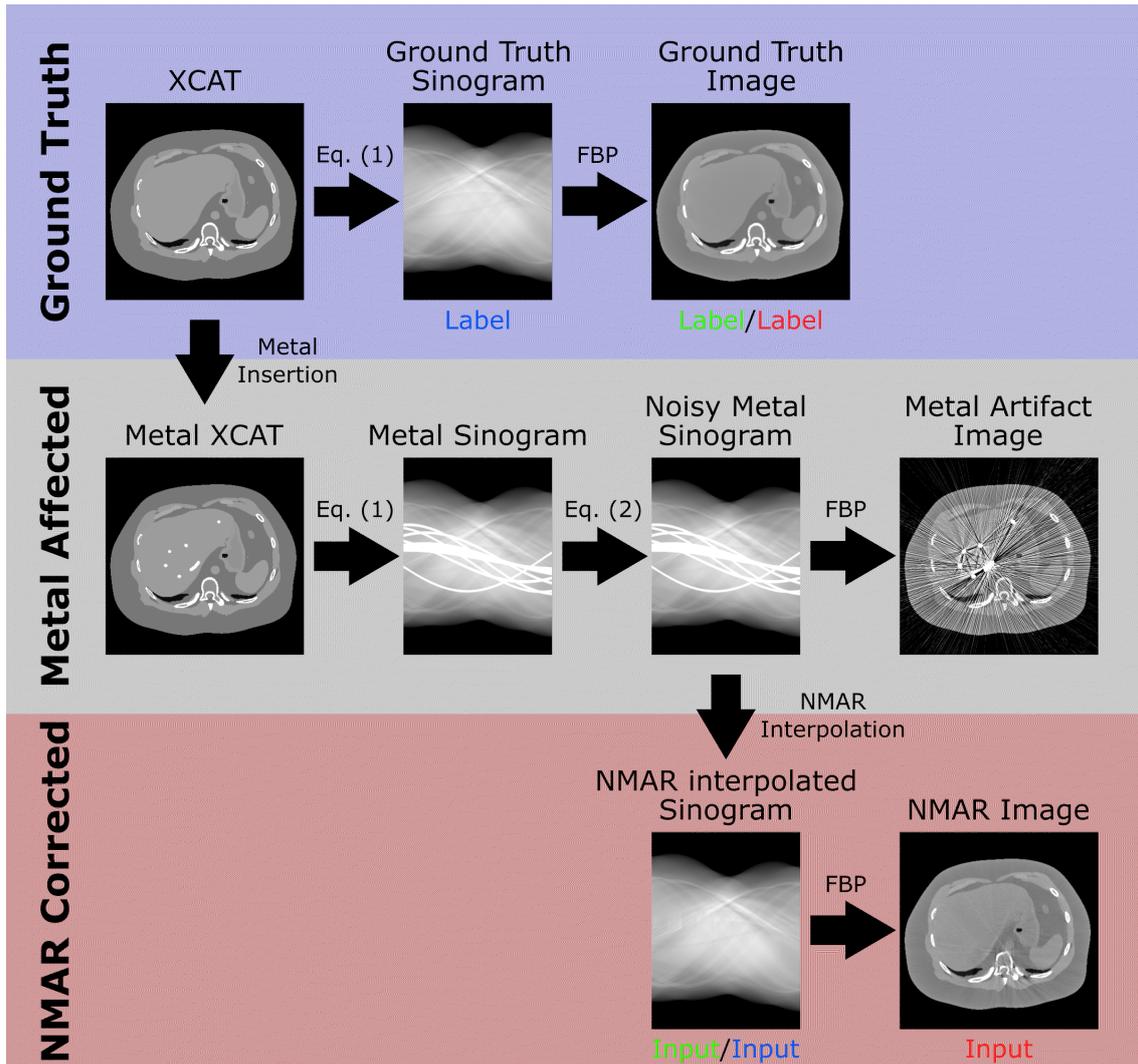


Figure 4.3: Generation of the metal training dataset. The input and labels for the three used deep learning networks are indicated by the colored words "Input" and "Label" beneath the corresponding images. The color coding corresponds to the colors used for the network architectures in Figure 4.1 (green: iCTU-Net, blue: U-Net Sino, and red: U-Net Image).

images for the iCTU-Net (green), U-Net Sino (blue), and U-Net Image (red) are noted in Figure 4.3.

4.2.5 Evaluation

The reconstructions are evaluated by calculating the peak signal-to-noise ratio (PSNR) and SSIM for the test data. We set the background values of ground truth and reconstructions to -1000 HU to focus the analysis on the body region where the relevant anatomy is located. For the evaluation, the slices of the test dataset are divided into three categories: no metal, moderate metal artifacts, and severe metal artifacts, with 106, 748, and 137 slices, respectively. This separation allows to evaluate the reconstructions when no metal is present. A slice is assigned to the severe metal artifact category if the FBP yields an SSIM value of less than 0.7, and to the moderate metal artifact category if the SSIM value of the FBP is greater than or equal

to 0.7. The SSIM threshold is chosen such, that the number of slices with severe metal artifacts is similar to the number of slices without metal.

4.2.6 Experiments

We conduct three experiments. First, we configure our iCTU-Net in an ablation study. Then, we investigate the impact of different sinogram input data for training in an input study. Finally, we compare our best network configuration with state-of-the-art MAR algorithms.

In the ablation study, we investigate different post-processing layers and loss functions. The purpose of the ablation study is to find settings for the iCTU-Net that yield the best reconstructions. The resulting network configuration will be used in following studies. We train three networks with different post-processing layers after back projection: no post-processing, three convolution layers, and a U-Net. All of these networks are trained with the SSIM loss and with pre-corrected NMAR sinograms as input. To investigate the influence of the loss function, we additionally train the U-Net post-processing network with the MSE loss. Both SSIM and MSE are commonly used loss functions in CT artifact reduction and CT reconstruction [77].

In the input study we train the network with different sets of training input data in addition to the previously used pre-corrected NMAR sinograms. The idea behind the input study is to find out how the network behaves for different kinds of sinogram input data. We use sinograms without metal (ground truth sinogram in [Figure 4.3](#) but with additional noise added via [Equation 4.2](#)) to investigate the network’s performance if no metal is present. In this way, the reconstruction performance and the ability to mitigate metal artifacts can be evaluated separately. We calculate the evaluation metrics for different categories of artifact severity, even though none of the test data contain any metal. Nevertheless, the categories are used to allow fair comparisons to the other networks. We also train a network with uncorrected metal sinograms (noisy metal sinogram in [Figure 4.3](#)), to see if an NMAR pre-correction is necessary.

Finally, in the comparison study, we compare our iCTU-Net with the NMAR sinogram inpainting algorithm and the U-Net Sino and U-Net Image networks described earlier.

Table 4.1: SSIM and PSNR evaluation metrics for the ablation study. All networks are trained with pre-corrected NMAR sinograms and the SSIM loss, except for the MSE iCTU-Net, which is trained with the MSE loss. The best result for each metric is marked **bold**.

		FBP	No Post- processing	3 Conv Layers	iCTU-Net	MSE iCTU-Net
No	SSIM	0.988 ± 0.020	0.865 ± 0.038	0.903 ± 0.028	0.969 ± 0.008	0.962 ± 0.010
Artifacts	PSNR	50.3 ± 4.9	31.0 ± 2.9	30.5 ± 4.3	37.7 ± 1.8	37.3 ± 1.3
Moderate	SSIM	0.869 ± 0.087	0.859 ± 0.036	0.897 ± 0.038	0.976 ± 0.007	0.967 ± 0.009
Artifacts	PSNR	15.2 ± 3.1	33.1 ± 2.1	34.3 ± 2.8	39.5 ± 1.8	39.3 ± 1.5
Severe	SSIM	0.625 ± 0.056	0.715 ± 0.06	0.832 ± 0.024	0.970 ± 0.009	0.946 ± 0.013
Artifacts	PSNR	10.6 ± 1.4	24.9 ± 5.6	34.2 ± 1.3	40.7 ± 1.6	39.6 ± 1.3
All	SSIM	0.848 ± 0.125	0.840 ± 0.06	0.889 ± 0.04	0.975 ± 0.008	0.964 ± 0.012
Images	PSNR	18.3 ± 11.6	31.7 ± 4.0	33.9 ± 3.1	39.5 ± 1.9	39.1 ± 1.6

4.3 Results

4.3.1 Ablation Study

The results of the evaluation metrics for the ablation study are shown in Table 4.1 and reconstructed images are shown in Figure 4.4. We first investigate the impact of the different post-processing layers. Compared to using the U-Net for post-processing, using no post-processing and three convolutional layers performs generally worse, especially for severe artifacts. When using no post-processing, clear streak and extinction artifacts are present. Using the three convolutional layers as post-processing improves SSIM and PSNR and the streak and extinction artifacts disappear. However, the geometry of some soft tissue organs such as the liver is not reconstructed correctly, which is particularly evident for severe metal artifacts. Using the U-Net as the final layers of the network substantially improves the evaluation metrics, completely eliminates artifacts and reconstructs organs more accurately. For no artifacts, the iCTU-Net underperforms compared to the FBP, especially in terms of PSNR. As shown by the arrows in the zoomed regions in Figure 4.4, the iCTU-Net is not capable of resolving small structures of only a few millimeters in size. From now on we will only use the U-Net for post-processing, as it yields the best results.

Finally, we train the iCTU-Net with the MSE loss. For no artifacts and moderate artifacts, the SSIM and PSNR evaluation metrics for the SSIM and MSE losses are similar. However, the SSIM metric for the MSE loss is considerably worse for severe artifacts and the reconstructions of the MSE iCTU-Net in Figure 4.4 look grainy. Thus, the network with U-net post-processing layers combined with SSIM loss performs best. Only this network configuration is referred to as iCTU-Net in this work.

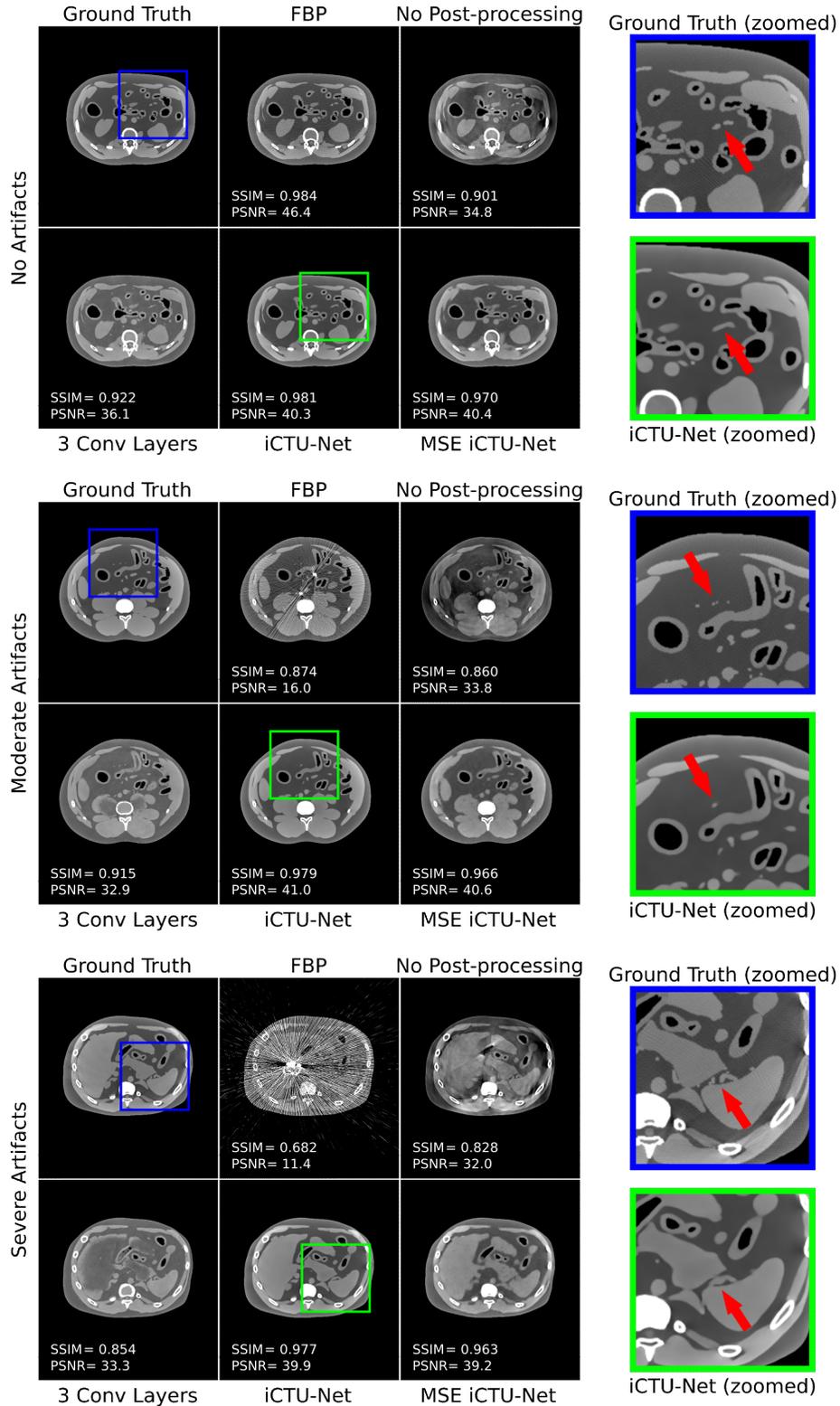


Figure 4.4: Results of the ablation study, where the ground truth and FBP are compared to different iCTU-Net settings. All networks are trained with pre-corrected NMAR sinograms and the SSIM loss, except for the MSE iCTU-Net, which is trained with the MSE loss. A slice without metal artifacts, with moderate metal artifacts, and with severe metal artifacts is shown. The scans are windowed to $[-300 \text{ HU}, 300 \text{ HU}]$ to increase the visibility of the artifacts. The arrows in the zoomed regions indicate small structures that the iCTU-Net cannot resolve accurately.

Table 4.2: SSIM and PSNR evaluation metrics for the input study. The differentiation of artifact severity is not meaningful for No Metal Input because none of the test data contain metal. Since this network is not trained with any metal data, it is not suitable for artifact reduction. However, to allow a reasonable comparison to the other methods, we keep the categories, meaning the same slices are used for evaluation. The best result for each metric is marked **bold**.

		FBP	No Metal Input	Metal Input	iCTU-Net
No Artifacts	SSIM	0.988 ± 0.020	0.967 ± 0.006	0.968 ± 0.011	0.969 ± 0.008
	PSNR	50.3 ± 4.9	37.8 ± 1.5	37.4 ± 1.9	37.7 ± 1.8
Moderate Artifacts	SSIM	0.869 ± 0.087	0.974 ± 0.004	0.975 ± 0.009	0.976 ± 0.007
	PSNR	15.2 ± 3.1	40.2 ± 1.5	39.6 ± 1.9	39.5 ± 1.8
Severe Artifacts	SSIM	0.625 ± 0.056	0.968 ± 0.005	0.962 ± 0.012	0.970 ± 0.009
	PSNR	10.6 ± 1.4	40.7 ± 0.8	40.4 ± 1.7	40.7 ± 1.6
All Images	SSIM	0.848 ± 0.125	0.972 ± 0.005	0.972 ± 0.011	0.975 ± 0.008
	PSNR	18.3 ± 11.6	40.0 ± 1.6	39.5 ± 2.0	39.5 ± 1.9

4.3.2 Input Study

In the input study, we investigate different sinogram inputs for the iCTU-Net. The results of the evaluation metrics for the input study are shown in Table 4.2 and reconstructed images are shown in Figure 4.5. The SSIM and PSNR in Table 4.2 show that the network performs similarly independent of the input. The network trained without metal in the input sinogram achieves the best PSNR, and the network trained with the pre-corrected NMAR sinograms achieved the best SSIM. However, these differences in SSIM and PSNR are not significant. For the metal input, some reconstruction inaccuracies close to metal objects can be observed, as indicated by the arrows in the zoomed images in Figure 4.5. Apart from this, the reconstructions in Figure 4.5 show no noticeable differences in image quality. Therefore, we continue to use the pre-corrected NMAR sinograms for the iCTU-Net. This allows for a fairer comparison with the deep learning reference methods, since they also use NMAR inputs.

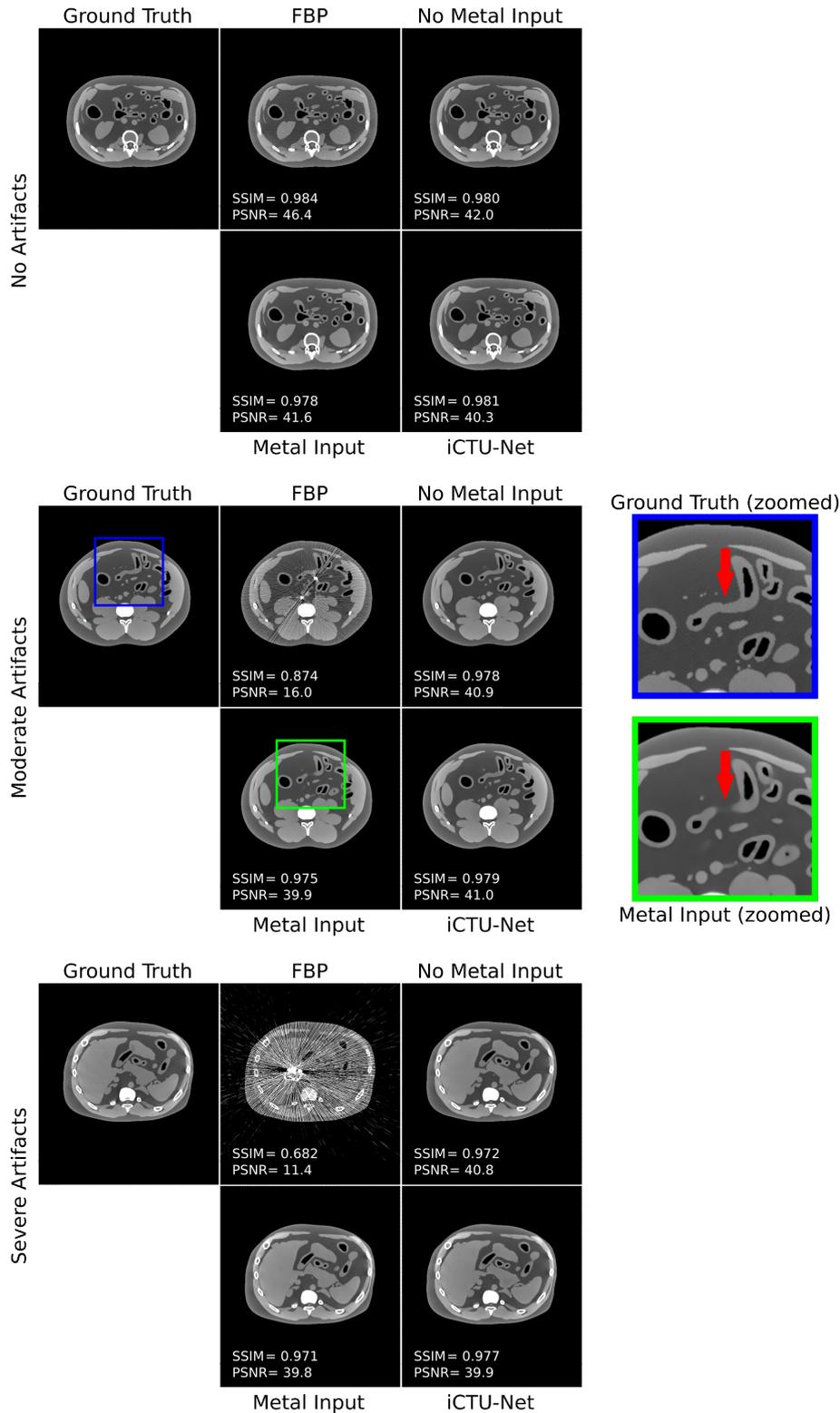


Figure 4.5: Results of the input study, where the ground truth and FBP are compared to iCTU-Nets trained with different input sinograms. All networks are trained with the U-Net post-processing layers and the SSIM loss, which yields the best results in the ablation study. No Metal Input, Metal Input, and iCTU-Net are respectively trained with metal-free, metal, and NMAR pre-corrected sinograms. A slice without metal artifacts, with moderate metal artifacts, and with severe metal artifacts is shown. The scans are windowed to $[-300 \text{ HU}, 300 \text{ HU}]$ to increase the visibility of the artifacts. The arrows in the zoomed images indicate an anatomy that the Metal Input network cannot resolve accurately.

Table 4.3: SSIM and PSNR evaluation metrics for the comparison study. The best result for each metric is marked **bold**.

		FBP	NMAR	U-Net Sino	U-Net Image	iCTU-Net
No	SSIM	0.988 ± 0.020	0.988 ± 0.020	0.990 ± 0.011	0.993 ± 0.010	0.969 ± 0.008
Artifacts	PSNR	50.3 ± 4.9	50.3 ± 4.9	46.5 ± 2.1	45.3 ± 3.2	37.7 ± 1.8
Moderate	SSIM	0.869 ± 0.087	0.976 ± 0.016	0.979 ± 0.012	0.983 ± 0.011	0.976 ± 0.007
Artifacts	PSNR	15.2 ± 3.1	44.3 ± 3.1	43.9 ± 2.6	42.8 ± 2.1	39.5 ± 1.8
Severe	SSIM	0.625 ± 0.056	0.911 ± 0.042	0.936 ± 0.025	0.944 ± 0.024	0.970 ± 0.009
Artifacts	PSNR	10.6 ± 1.4	38.2 ± 2.4	39.3 ± 1.9	39.8 ± 1.9	40.7 ± 1.6
All	SSIM	0.848 ± 0.125	0.968 ± 0.032	0.974 ± 0.021	0.978 ± 0.019	0.975 ± 0.008
Images	PSNR	18.3 ± 11.6	44.1 ± 4.4	43.5 ± 3.1	42.7 ± 2.6	39.5 ± 1.9

4.3.3 Comparison Study

The results of the evaluation metrics for the comparison study are shown in Table 4.3 and reconstructed images are shown in Figure 4.6. The deep learning reference methods U-Net Sino and U-Net Image both perform better than NMAR in terms of SSIM, especially for severe artifacts. In terms of PSNR, they perform worse when artifacts are not present and similarly when artifacts are present. The U-Net Image achieves a slightly higher SSIM than the U-Net Sino, but the performance of both methods is very similar. In Figure 4.6 no substantial removal of metal artifacts can be observed for the U-Net Sino and U-Net Image, only a smoothing of the streak artifacts is observed for the U-Net Image.

Without artifacts, the iCTU-Net is outperformed by all methods in terms of PSNR and SSIM, as they are all FBP-based, which already outperformed the iCTU-Net in the ablation study. For moderate artifacts, the iCTU-Net achieves competitive SSIM values compared to the reference methods, but performs worse in terms of PSNR. Nevertheless, the iCTU-Net is the only method capable of completely removing moderate metal artifacts, as shown in Figure 4.6. As indicated by the arrows in the zoomed images in Figure 4.6, the iCTU-net is also the only method that can restore a blood vessel into which a metal object has been inserted. For severe artifacts, the iCTU-Net performs better than all reference methods with $SSIM = 0.970 \pm 0.009$ and $PSNR = 40.7 \pm 1.6$. The second best method, the U-Net Image, only achieved $SSIM = 0.944 \pm 0.024$ and $PSNR = 39.8 \pm 1.9$. Averaged over all images, the SSIM of the iCTU-Net is competitive with the U-Net Image, but a worse PSNR is achieved. The iCTU-Net is able to remove severe metal artifacts completely, whereas for the other methods strong streak artifacts are still present over the whole image. The iCTU-net can not only efficiently remove severe artifacts, but also reliably restore the anatomy that is obstructed by these artifacts. This is especially evident inside the circles shown in the zoomed images in Figure 4.6. All other methods fail to restore the anatomy in this region.

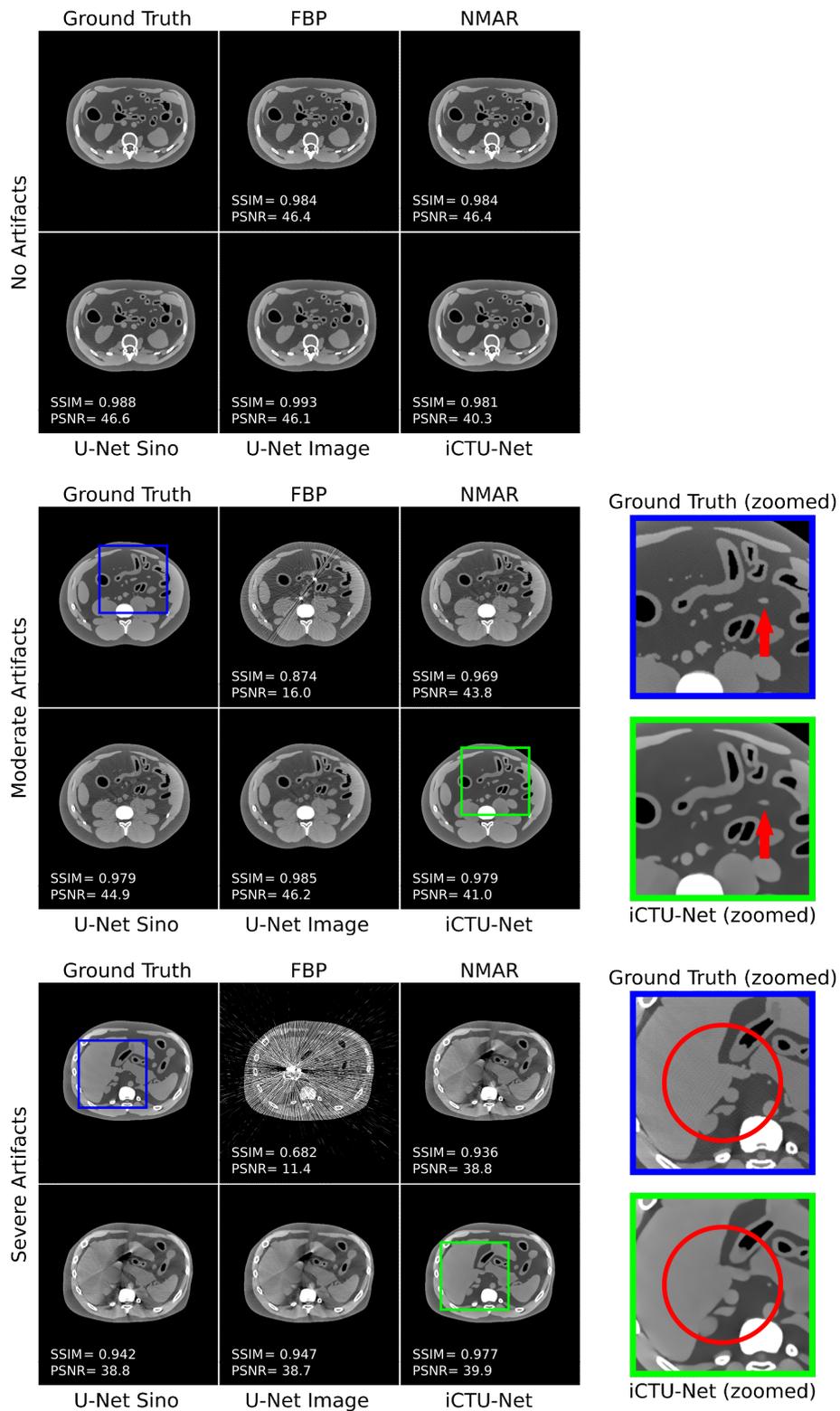


Figure 4.6: Results of the comparison study, where the ground truth and FBP are compared to NMAR, U-Net Sino, U-Net Image, and iCTU-Net. A slice without metal artifacts, with moderate metal artifacts, and with severe metal artifacts is shown. The scans are windowed to $[-300 \text{ HU}, 300 \text{ HU}]$ to increase the visibility of the artifacts. The arrows and circles in the zoomed images indicate anatomies that could only be recovered by the iCTU-Net.

4.4 Discussion

We trained the iCTU-Net with metal-affected data, to investigate its ability to mitigate metal artifacts. The iCTU-Net outperformed the reference methods for reconstructions with severe metal artifacts. Similar results were found for the application of the iCTU-Net to sparse-angle CT reconstruction, where the iCTU-Net showed good performance for a small number of projections [77]. However, the iCTU-Net was not able to resolve small structures of only a few millimeters in size. The reconstructions were slightly blurred, which is probably the reason why the iCTU-Net could not match the quality of the FBP when no metal is present. In the ablation study, it was found that the loss function and the post-processing layers have a major impact on the quality of the reconstruction. We had attempted to sharpen the reconstructed image, by combining the SSIM loss with an additional gradient difference loss [84], but no substantial improvements were observed. In the future, we will investigate alternatives to the U-Net as post-processing layers to further optimize the network. The iCTU-Net was trained with a dataset of 3964 slices, of which only 310 contained no metal. Due to this small fraction of metal-free training data, the network might not be able to learn how to properly reconstruct metal-free sinograms. In the input study, we trained the reconstruction network exclusively with metal-free data to test this hypothesis. We found that the network trained with metal-free raw data did not perform better than the iCTU-Net for the no artifact category. Therefore, we can conclude that training the network with mainly metal-affected data does not degrade the quality of the reconstructions. Interestingly, the evaluation metrics for the moderate and severe artifact categories also did not differ substantially. Thus, the network trained with metal-affected input data is reconstructing images with metal-affected test data just as well as the network trained without metal is reconstructing images that do not include metal. This shows that the iCTU-Net is reliably reducing metal artifacts. This is confirmed by the fact that all networks in the input study performed very similarly for all severities of artifacts. The network seems to handle metal objects in the raw data very well.

The input study showed that the iCTU-Net performs similarly regardless of the sinogram input data used. Training the network with uncorrected metal sinograms revealed similar performance compared to the network trained with pre-corrected NMAR sinograms. This means that reconstruction without pre-correction is feasible, which reduces the complexity of the algorithm.

In the comparison study, a sinogram pre-processing and an image post-processing approach were investigated. We have found that the image-based post-processing deep learning approach provides better results than the sinogram pre-processing approach. This is consistent with the findings of Arabi *et al.* [85]. Since the reference methods are all FBP-based, they are superior to the iCTU-Net in the absence of artifacts due to the aforementioned blurring. However, the artifacts introduced by the FBP cannot be completely mitigated by the reference methods. The iCTU-Net is the only method that removes all metal artifacts and yields the best results of all methods for severe metal artifacts. Since the iCTU-Net is trained end-to-end, the network can fully utilize the raw data and learn to reconstruct an artifact-free image. The U-Net Sino learns to mitigate disturbances in the sinogram with the raw data as input. However, small errors in the sinogram can lead to significant deviations

in the reconstruction [77], which the U-Net Sino cannot correct. The U-Net Image only mitigates the artifacts in the image domain introduced by the FBP. In doing so, the network no longer has the original raw data to learn from.

The usage of digital XCAT phantom data for metal data simulation instead of real patient data has several advantages. First of all, with the organ masks provided by the XCAT, metal objects can automatically be inserted in specific body regions. In this work, we inserted iron into the blood vessels. For future studies it would be better to insert attenuation coefficients of materials that are commonly used for contrast agents and catheters. Moreover, for the simulation of polychromatic projections, it is not necessary to segment the images into soft tissue, bone, and metal to assign the corresponding attenuation coefficients, as is done in several other works [63, 70, 85]. Instead, the organ masks of the XCAT allow the insertion of energy-dependent attenuation coefficients for every organ. In the future, it is desirable to test the iCTU-Net on experimental raw data instead of simulated data. However, this requires the iCTU-Net to be adapted to work with the raw data of multirow detector CT scanners. The two-dimensional projection data might lead to restrictions due to GPU memory limitations. Since dual-energy CT has been shown to help reduce metal artifacts [54, 59], the iCTU-Net should benefit from the additional spectral information. Photon-counting CT is another spectral technology that can be used to reduce metal artifacts [86]. The energy of individual photons can be measured by energy-resolving detectors [87]. The iCTU-Net is readily applicable to energy resolved raw data by including the energy information in separate input channels. The additional spectral information in the raw data is expected to mitigate beam hardening artifacts.

We will also investigate the ability of the iCTU-Net to simultaneously mitigate different kinds of artifacts. This is achievable by using a training dataset that contains a combination of artifacts. Promising results for the isolated mitigation of artifacts with the iCTU-Net in low-dose CT and sparse-angle CT have already been shown [77].

4.5 Conclusions

The presented end-to-end deep learning CT reconstruction algorithm was trained with simulated interventional data to mitigate metal artifacts during reconstruction. We showed that the iCTU-Net reconstruction MAR approach is better suited to mitigate metal artifacts than commonly used sinogram pre-processing and image post-processing deep learning approaches. The iCTU-Net is the only studied method that can eliminate the metal streak artifacts. However, the end-to-end reconstruction approach performs worse than the other approaches when no artifacts are present. Reconstructions without any metal showed that the iCTU-Net is prone to blurring. Because the whole reconstruction is learned, the network is able to fully utilize the raw data, which benefits the removal of metal artifacts. In the future we will try to improve the network architecture by investigating alternative loss functions and post-processing layers to avoid blurring. We will also train networks with data including different kinds of artifacts to investigate simultaneous mitigation of several types of artifacts.

5. ”Development of an abdominal phantom for the validation of an oligometastatic disease diagnosis workflow”, *submitted to Medical Physics,* *19.05.2021*

5.1 Introduction

The liver is a common site for metastatic disease due to its rich blood supply via the hepatic artery and portal vein [88]. Liver metastases originate from a variety of cancers such as colon, rectum, pancreas, lung, and breast [89]. More than half of all patients with colorectal cancer develop liver metastases, which are the leading cause of death in colorectal cancer patients [90]. In general, metastatic liver disease is a challenging and life-threatening condition with a grim prognosis and outcome [5]. Resection of the liver metastases is the most promising treatment approach [91]. Unfortunately, systemic spread of the disease is already evident in most patients with liver metastases, which means that surgical resection is often no longer an option [88]. Radiofrequency ablation is becoming a popular treatment modality, but should only be used in combination with resection or when resection is not safely possible [92]. Oligometastatic disease (OMD) is defined by the localization of a small number of sites and lesions where local ablative treatment approaches promise improved disease control and thus an optimistic clinical outcome [7]. A gold standard workflow for the diagnosis of OMD is not available due to the high individuality of this disease. There are guidelines based on expert consensus that suggest an initial CT scan [7]. When in doubt, a stepwise imaging using MRI, US or positron emission tomography CT (PET-CT) is recommended. Contrast-enhanced US (CEUS) can be used to characterize focal liver lesions, but is often unable to correctly diagnose small and benign lesions [93]. Repeated imaging with dynamic CEUS for early treatment response assessment provides complementary information to CT and MRI [94].

The research campus “Mannheim Molecular Intervention Environment” (M²OLIE) proposes a standardized workflow towards a personalized medicine approach to OMD, which allows patient-specific treatment planning by combining morphological, molecular, and functional information about the lesions [29]. The envisioned workflow comprises the acquisition of a multimodal CT, MRI, and PET-CT dataset, novel image processing techniques, and robotically assisted CBCT-guided biopsy for subsequent molecular subtyping. The detection of metastases is not comprehensively covered by a single imaging technique; thus the multimodal imaging approach is necessary to identify all lesions [95] and to assist the interventionalist. Comprehensive detection of metastases and subsequent subtyping is critical to prolong survival via

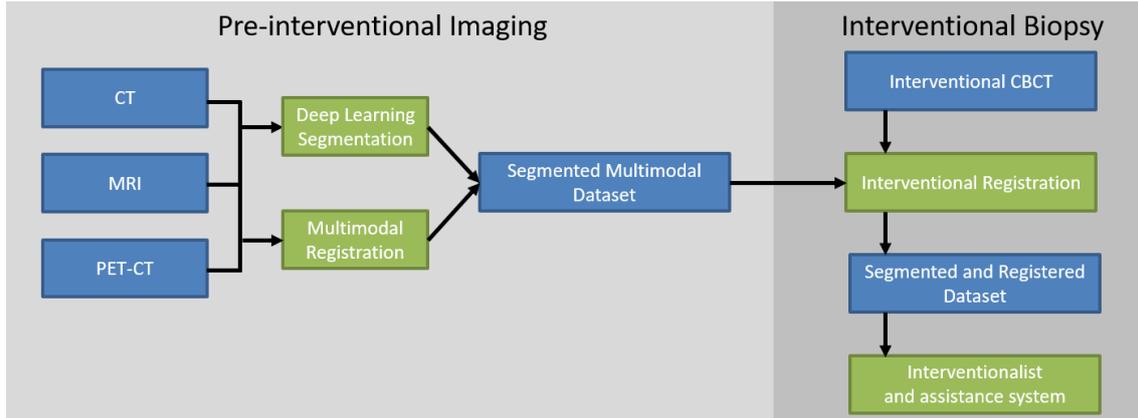


Figure 5.1: The envisioned OMD diagnosis workflow starts with CT, MRI, and PET-CT imaging, followed by multimodal image registration and a deep learning segmentation. The obtained segmented multimodal dataset provides information to assist the clinician during intervention.

local treatment [96–98]. Molecular differentiation of the lesions during the diagnosis stage provides crucial information to the deciding physicians.

Preliminary validation using imaging phantoms is a vital step for the introduction of the workflow into clinical routine. Numerous abdominal liver phantoms have been developed for different imaging modalities. However, most of them are monomodal and the design of multimodal phantoms is not trivial [99]. Many liver phantoms in literature are either not anthropomorphic [100], not embedded in a realistic abdomen [101] or do not include liver lesions [102, 103]. Using 3D printing with radiopaque ink on paper enables the production of CT biopsy phantoms with highly detailed arbitrary patient anatomies [104]. However, these phantoms are not designed for other modalities, such as MRI. The Triple Modality 3D Abdominal Phantom (Model 057A, CIRS Inc.) is a commercially available phantom suitable for CT, MRI, and US. It uses a simplified anthropomorphic geometry and among other organs includes a liver with six puncturable liver lesions, which are visible in all imaging modalities. Because the visibility of small or low-contrast lesions in real patients may be limited in CT scans [95], a phantom should contain lesions with different visibility depending on the imaging modality to reflect clinical reality.

In this work, we present an anthropomorphic multimodality phantom for percutaneous liver interventions. The phantom is puncturable and the liver contains a vascular tree and six lesions, with variable visibility for CT, MRI, and PET-CT. We used the phantom to evaluate the proposed standardized OMD diagnosis workflow, including multimodal imaging, image processing, and robotically assisted CBCT-guided biopsy. The biopsy was conducted for all six lesions, two of which are visible only in MRI and not in CBCT. Multimodal image registration was deployed to locate all lesions during intervention.

5.2 Materials and Methods

5.2.1 OMD diagnosis workflow

The envisioned diagnosis workflow includes multimodal imaging, state of the art image processing, and an interventional liver biopsy, as depicted in Figure 5.1. It starts with pre-interventional CT, T1w and T2w MRI, and PET-CT imaging to locate all liver lesions and to support the CBCT-guided biopsy. Each imaging modality has different sensitivities to certain anatomies; thus, the detection of lesions cannot be robustly performed by only utilizing one imaging technique [95]. A CT is characterized by its stable geometry, high-spatial resolution. The acquisition of a pre-interventional CT facilitates the registration of the pre-interventional data to the interventional CBCT data. MRI provides excellent soft tissue contrast and is able to reveal smaller low-contrast lesions which are invisible in CT [95, 105], without exposing patients to ionizing radiation. MRI is additionally able to provide functional parameters for diffusion or perfusion [106]. The use of PET-CT is recommended when the patient has potential extrahepatic lesions [7, 107] and provides an overview of metabolic activity that can be used to determine to what extent a biopsy should be performed.

To combine the advantages of the different imaging modalities, the acquired dataset is registered [29] and anatomical structures such as liver, liver lesions, and blood vessels are segmented using novel deep learning algorithms [108]. The segmentations provide location and size of biopsy targets and organs at risk, which must be considered during needle placement.

Tissue biopsies allow to sample parts of a tumor or lesion to provide crucial molecular information for the diagnosis and treatment of OMD [97, 109]. As tumors evolve in malignancy, early and comprehensive detection of metastases is critical to improve survival rates [96]. Furthermore, sampling of metastases is an important step towards a radiomics approach for the diagnosis and treatment of OMD [110, 111]. For a meaningful sampling of the lesion or organ, the biopsy needle placement needs to be accurate. By registering the pre-interventional patient data to the interventional image data, the interventionalist and the robotic system are provided with the necessary information for the needle path planning. Robotic assisted percutaneous interventions provide increased accuracy and less time spent on the procedure [112]. Compared to CT-guided biopsies, CBCT guidance offers a larger working space and 3D imaging during the procedure with similar diagnostic performance and effective dose [113].

5.2.2 Phantom design and construction

The proposed workflow was executed with an anthropomorphic phantom under clinically relevant measurement conditions. The phantom was designed to mimic the human abdomen incorporating a 3D-printed rib cage and spine, an aorta, lungs, a liver, six hepatic lesions, and a hepatic vessel tree. Three lesion types with different visibility in CT and MRI were created, one of which should be invisible in CT. Processing the proposed phantom via the above described imaging chain provides a proof-of-principle for the advantage of multimodal imaging, since it demonstrates the complementary information of each modality.

A CAD model of the phantom is displayed in Figure 5.2 (a) and the manufactured phantom is shown in Figure 5.2 (b). All organ contours were obtained from patient segmentations. Synthetic ballistic gelatin (10% ballistic gelatin, Clear Ballistics) was used as a tissue surrogate. Ballistic gelatin liquefies at 110°C and can therefore be poured into the desired shape, which it retains after cooling down. The remaining phantom components and the body mold must be heat-resistant, as they come into contact with the molten ballistic gelatin. In order to cast the ballistic gelatin into the shape of an abdomen, a body hull casting mold made of heat-resistant synthetic resin (High Temp Resin V2, Formlabs) was manufactured using SLA. The body hull casting mold consists of several individual parts shown in Figure 5.3 (a). This facilitated the removal of the body hull after molding without damaging it. The hull was firmly assembled with screws and additionally sealed with high-temperature silicone to prevent leakage, see Figure 5.3 (b). The bottom of the mold was closed with an extruded acrylic plate displayed in Figure 5.3 (c) and sealed with high-temperature silicone. The plate was laser cut to match the base of the phantom and was not removed after the completion of the phantom, because it serves as a convenient bearing surface for storage. The rib cage, aorta, and lungs were glued to the inside of the base plate using heat-resistant adhesive. The complete torso was not cast in one go, but in layers. This made it possible to embed the liver between two casting steps and to check each layer for bubbles and remove them.

The rib cage and spine displayed in Figure 5.4 (a) were produced by SLS using polyamide powder filled with glass particles (PA-GF, Materialise). The material yields high attenuation coefficients in CT, whilst not showing a signal in MRI. It exhibits a high thermal resistance and is not puncturable. The lungs shown in Figure 5.4 (b) were printed via FFF with blue high impact polystyrene (EasyFil HIPS, Formfutura). The lungs were thinly coated with Shore A 5 hardness silicone (Silicone Addition Colorless 5, Siliconesandmore) to prevent leakage of air bubbles and of the blue color into the hot ballistic gelatin. The aorta is a plastic tube with high heat resistance and an inner and outer diameter of 12 mm and 14 mm, respectively. It was placed in front of the spine and extends through the entire phantom, allowing flow measurements.

In Figure 5.4 (c) parts of the liver mold, the hepatic vessel tree, and the six hepatic lesions are displayed. The liver mold was printed via SLA with synthetic resin (Clear Resin V4, Formlabs). Like the body hull, the liver mold was printed in smaller parts that can be assembled with screws, allowing an easy release of the mold. The vessel tree was also printed via SLA with synthetic resin (White Resin V4, Formlabs). For easy positioning, the vessel tree was attached to the liver cast mold. Two liver lesions of each of the three lesion types were created and the lesions were colored with silicone dye (Type A: dark blue, Type B: light blue and Type C: green). The ovoid lesions have a length of 3 cm and a diameter of 2 cm. Type A and B were made of Shore A 5 silicone and type C of Shore A 19 silicone (TFC silicone rubber type 19 clear, Trollfactor). A harder silicone was chosen for type C to achieve a varying stiffness of the lesions. To ensure visibility in MRI, 0.4% MRI contrast agent (Dotarem, 0.5 mmol/ml gadoteric acid, Guerbet) was added to all lesion types. CT contrast was enhanced by adding 2.4% CT contrast agent (Imeron 400, Bracco Imaging) to type B and C. No Imeron was added to type A,

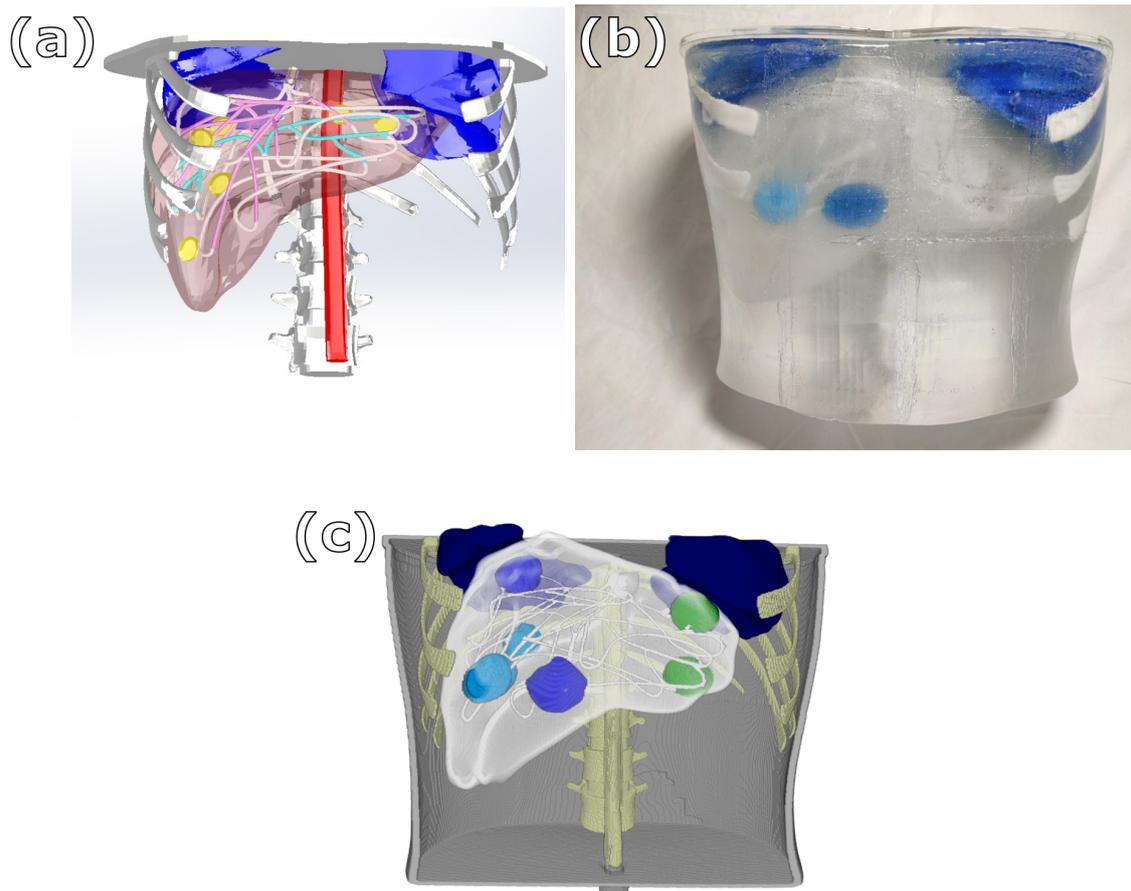


Figure 5.2: (a) CAD model of the phantom. The position of the lesions in the CAD model do not correspond to the final positions in the manufactured phantom. (b) Phantom front view. (c) Phantom organ segmentations.

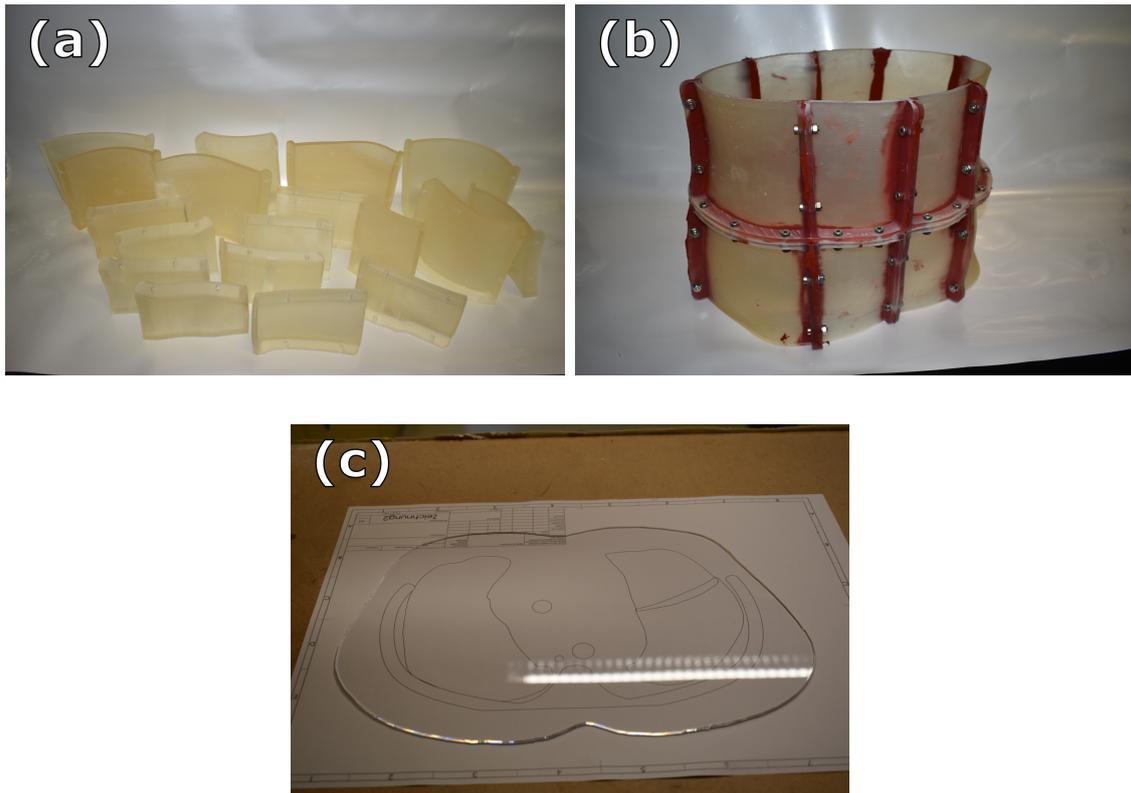


Figure 5.3: (a) Body hull pieces. (b) Assembled body hull. (c) Acrylic bottom plate.

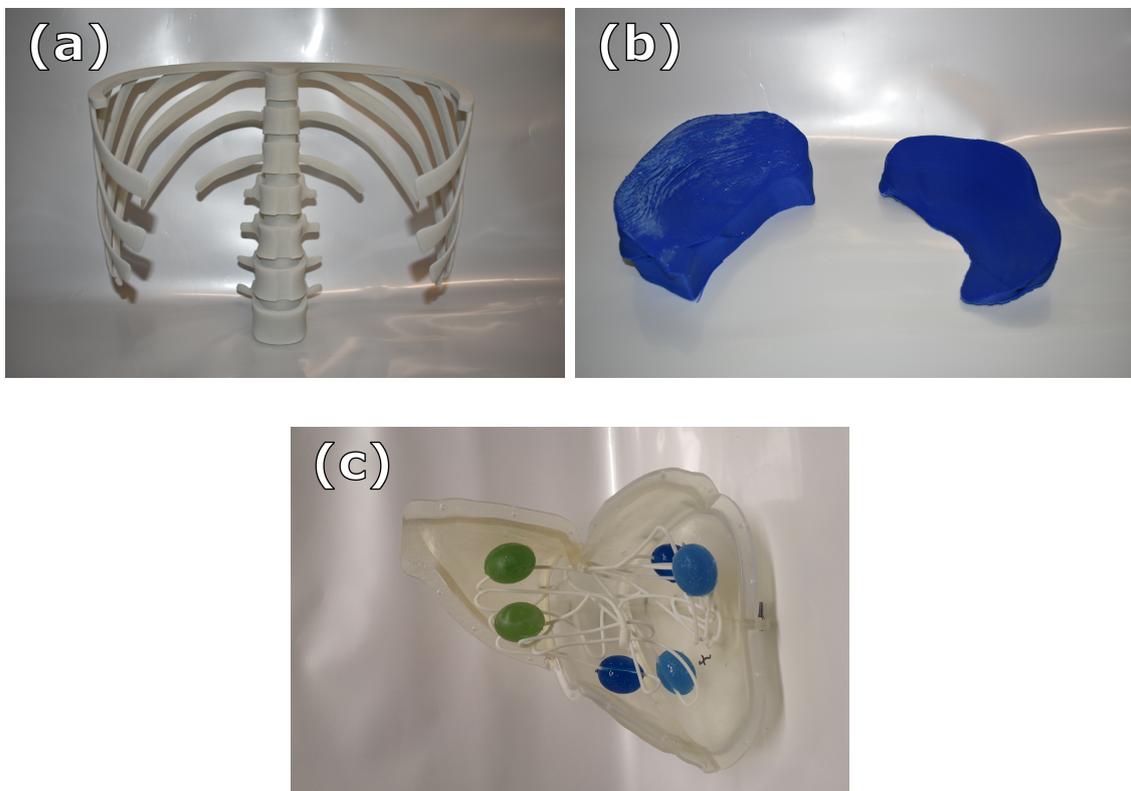


Figure 5.4: (a) Rib cage and spine. (b) Lungs. (c) Posterior view of the liver hull with liver lesions and hepatic vessel tree.

Table 5.1: MRI and CT contrast agent concentration and hardness of the hepatic lesions.

	Silicone	MRI	CT
	Shore A hardness	Contrast Agent	Contrast Agent
Type A (dark blue)	5	0.4% Dotarem	-
Type B (light blue)	5	0.4% Dotarem	2.4% Imeron
Type C (green)	19	0.4% Dotarem	2.4% Imeron

making them invisible in CT and CBCT. The contrast agent concentrations and hardness grades of the lesion are summarized in Table 5.1. The six liver lesions are cut open and wrapped around the vascular structures. This simulates the typical hypervascularity of hepatic lesions [114] while also facilitating their positioning. The liver parenchyma was molded with 2 kg of translucent Shore A 5 silicone.

5.2.3 Image acquisition

T1w and T2w MRI, CT, PET-CT, and CBCT scans of the phantom were acquired using clinically relevant parameter settings. The most important parameters are listed in Table 5.2. All scans were acquired on whole body clinical devices (Siemens Healthineers, Forchheim, Germany; MRI: MAGNETOM Aera; CT: SOMATOM Force; PET-CT: Biograph mCT; CBCT: Artis zee). Axial, coronal, and sagittal slices of the scans are shown in Figure 5.5. The slice positions were chosen such that all liver lesions are included and the lesion types are labeled in the T2w MRI scan. To produce a PET signal, a 2 ml vial was filled with a ^{68}Ga tracer. The vial was implanted into a type B lesion located a few cm below the phantom surface. After the PET-CT measurement, the vial was removed from the phantom, leaving behind no remainder of the tracer or its stable decay product ^{68}Zn .

Table 5.2: Summary of image acquisition parameters.

	Sequence	TE	TR	FA	Voltage	Spacing	Image size
		[ms]	[ms]	[°]	[kVp]	[mm ³]	[x/y/z]
T1w MRI (1.5 T)	Spin Echo	3.6	500	120	-	1 × 1 × 1	384 × 288 × 288
T2w MRI (1.5 T)	Spin Echo	81	6000	180	-	1 × 1 × 2	384 × 288 × 288
CT	-	-	-	-	90	0.97 × 0.97 × 0.5	512 × 512 × 605
PET-CT (CT)	-	-	-	-	120	0.73 × 0.73 × 1.5	512 × 512 × 264
PET-CT (PET)	-	-	-	-	-	4.07 × 4.07 × 3	200 × 200 × 115
CBCT	-	-	-	-	90	0.48 × 0.48 × 0.48	512 × 512 × 397

5.2.4 Image processing

The intended imaging workflow includes not only dedicated image acquisition, but also specified and extended image processing to extract and process all available information which equips the physician with complementary morphological and functional information provided by the various modalities. The pre-interventional CT, PET-CT, MRI, and interventional CBCT data are fused into a multiparametric data

set using image registration, and image information is extracted via image segmentation. For this phantom study, image registration and segmentation algorithms were used that are not applicable to patient scans.

5.2.4.1 Image segmentation

The extraction of information includes the application of automated deep learning image segmentation algorithms to identify different organs, such as the liver [108], and pathologies in the image data. This information is not only important for diagnosis, but also benefits the planning and performance of image-guided procedures by providing three-dimensional anatomical models of target and risk structures. Deep learning segmentation algorithms rely on a high similarity between training and test data. Since the similarity between patient data and the phantom was not high enough, traditional image processing methods were used instead in this study. Due to the sufficiently homogeneous signal intensities of the organs, it was possible to segment all organs in the phantom via region growing using the RegionGrowing-Macro module in MeVisLab 3.3. All segmentations shown in Figure 5.2 (c) were performed in CT, except for the two type A lesions and the hepatic vessel tree, which were segmented in T1w MRI.

5.2.4.2 Image registration

To continue the demonstration of the intended imaging workflow, the image data of the anthropomorphic phantom was registered. Due to the shape stability of the phantom, the application of a rigid registration method was sufficient. We implemented the registration in Python 3.5 using the SimpleITK 1.2.4 package [115]. We used the Euler3DTransform class, a rigid 3D transform that applies a 3D translation and a rotation described by three Euler angles. The registration scenario was divided into two steps. First, a registered pre-interventional CT, MRI, and PET-CT dataset was prepared. The T1w MRI was registered to CT using the Mattes Mutual Information (MMI) registration metric and the same transform was applied to T2w MRI. The PET-CT was registered to the CT using the normalized cross correlation (NCC) metric [84]. In a second step, the CT was registered to CBCT using the NCC metric. The obtained transformation was then applied to the MRI and PET-CT. A gradient descent optimizer was used for all registrations. This data set includes all available information about the patient’s morphology and pathology and builds the basis for path planning and image guidance during the intervention.

5.2.5 Lesion visibility

The three lesion types in the phantom were designed to be visible only in certain modalities. In order to show the usefulness of multimodal data in a CBCT-guided intervention, it is crucial that the two type A lesions are visible in MRI and invisible in CT and CBCT. To quantify the visibility of the three liver lesion types the contrast-to-noise ratio (CNR) in CT, CBCT, T1w and T2w MRI was calculated according to

$$CNR = \frac{|S_{lesion} - S_{liver}|}{\sigma_{noise}}, \quad (5.1)$$

where S_{lesion} and S_{liver} are signal intensities in the liver lesions and liver, respectively, and σ_{noise} is the standard deviation of the liver signal intensity. Additionally, the contrast of the liver to the surrounding tissue was determined by replacing S_{lesion} with S_{tissue} . The previously obtained organ segmentations were utilized to determine the signal intensities and noise.

5.2.6 Interventional biopsy experiment

A needle puncture experiment was performed in the interventional hybrid MRI suite of the research campus M²OLIE, consisting of Artis zee CBCT and MAGNETOM Aera MRI scanners. The CBCT-guided needle placement was performed with robotic assistance, where all six lesions were punctured with coaxial biopsy needles. Our biopsy setup is shown in Figure 5.6 (a). A vacuum mattress was used to stabilize the phantom. Before the needle placement, a CBCT was acquired and the pre-interventional dataset was registered to the CBCT scan to locate all lesions and plan the needle path. In this initial CBCT scan, both the robot and the phantom were present. This calibration scan was used to determine the position of the robot relative to the phantom. The needle trajectory was primarily planned on the modalities which have the best contrast for the targeted lesion. The type A MRI lesions were punctured with MRI-compatible titanium needles. First, five of the six lesions were punctured via an anterior percutaneous route. Lesion A2 is of particular interest because of its invisibility in CBCT and its close proximity to a lung. It was punctured with two needles that were placed parallel at a small distance. Lesion B2 was only accessible from a posterior percutaneous route, due to being located in the right posterior section of the liver. Thus, the phantom was positioned slightly tilted in prone position as shown in Figure 5.6 (c). Since the lesion is located under the rib cage, the needle must be placed between two ribs.

After the placement of the needles, a CBCT control scan was acquired for both biopsy settings. For the anterior biopsy, the phantom was additionally moved to an MRI machine to acquire a control scan for the type A MRI lesions, as shown in Figure 5.6 (b). The non-MRI-compatible needles were removed before the MRI scan. The MRI acquisition coil was placed under the phantom to not interfere with the needles.

5.3 Results

5.3.1 Multimodal imaging

Axial, coronal, and sagittal slices of the phantom for T1w and T2w MRI, CT, PET-CT, and CBCT scans are shown in Figure 5.5. The images were registered to the CBCT scan. We used a registration that incorporates resampling, thus the registered images have the same image resolution as the CBCT scan. The windowing was chosen to optimize visibility of the liver lesions. All six lesions are clearly visible in the T2w MRI scan. In the T1w MRI scan only the type C lesions are visible. The CT, PET-CT, and CBCT scans show the type B and C lesions with a high contrast. As intended, type A lesions are completely invisible.

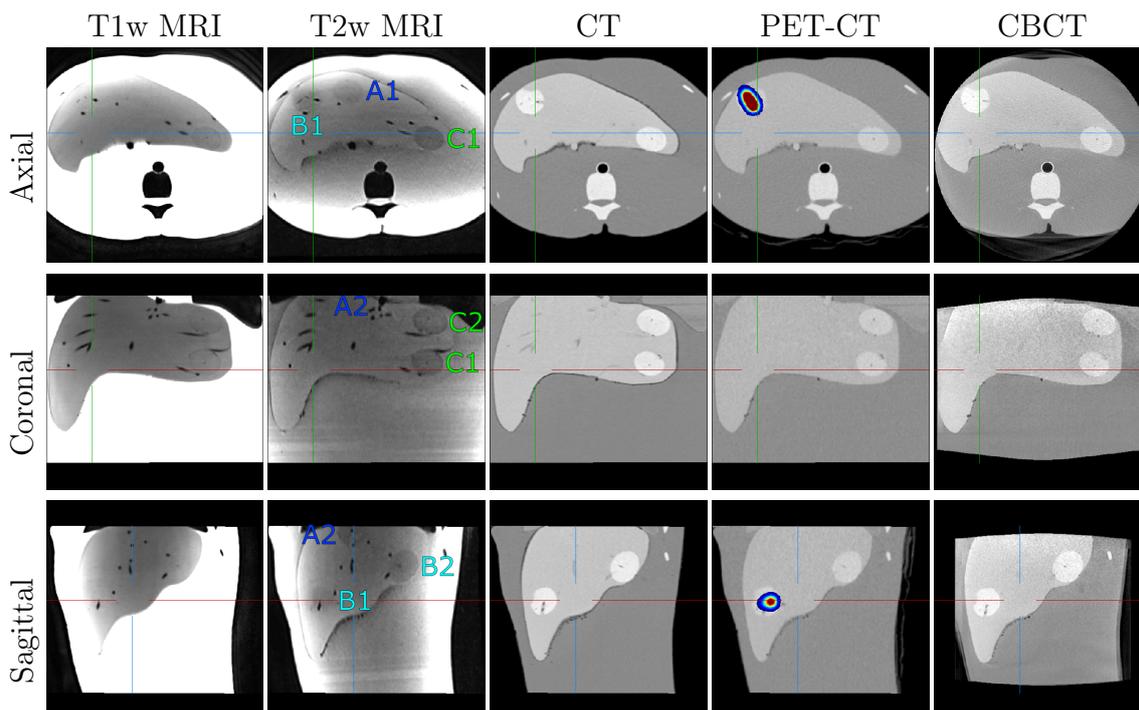


Figure 5.5: Clinical T1w MRI, T2w MRI, CT, PET-CT, and CBCT contrasts. Windowing was chosen to optimize visibility of liver lesions, resulting in oversaturation of MRI scans.

To quantify these observations, CNRs shown in Table 5.3 were calculated. CNRs between 1.0 and 1.4 confirm that every lesion is visible in T2w MRI. Objects with a CNR of 1.0 and the size of a few pixels would be difficult to detect, but the lesions occupy up to 1000 pixels per image slice in MRI, making them clearly visible. In T1w MRI, CNR is high for type C lesions and low for type A and B lesions. In CT and CBCT, a high CNR is found for type B and C lesions and a very low CNR for type A lesions. The signal difference between type A lesions and liver is only 2 HU for CT and CBCT. The CNR is generally lower in CBCT due to a high noise magnitude of 45 HU. Therefore, the CNR confirms the quantitative observations in Figure 5.5 and especially that the type A lesions are undetectable in CT and CBCT. Additionally, a high CNR for the liver means excellent discrimination of the liver from surrounding tissues for all modalities. The liver vessel tree is barely visible in CT, which is realistic for non-contrast enhanced imaging. The vessels yield no signal in MRI, which is not the case for real patients. Just like real bones, the PA-GF bones provide no signal in both MRI scans and a high signal of 400 HU is observed in the CT modalities. The realistic signal and high contrast of the bones in CT and MRI facilitates the multimodal registration, since the bones provide natural registration landmarks. The implanted PET tracer in the type B lesion shows a strong signal within the lesion. The PET-CT measurement of the phantom proves that the phantom is suitable for functional PET imaging and thus can be used to validate the entire diagnostic workflow.

Table 5.3: CNR of the liver and the liver lesion types.

	T1w MRI	T2w MRI	CT	CBCT
Type A (dark blue)	0.3	1.0	0.1	0.04
Type B (light blue)	0.2	1.4	12	4.6
Type C (green)	1.3	1.4	11	2.5
Liver	72	10	17	8

5.3.2 Interventional biopsy experiment

All six lesions of the phantom were punctured and control scans were used to assess the needle positions. CBCT and MRI control scans for the anterior biopsy are shown in Figure 5.6 (d) and (e), respectively. In the CBCT control scan, two needles targeting lesion A2 are shown. Additionally a needle tip inside lesion C2 can be seen, which was removed before acquiring the MRI scan. Because lesion A2 was not visible on the CBCT scan, the needle location was confirmed with the MRI scan. Instead of placing the needle directly in the center of the lesion, the target location was placed in front of the hepatic vessel that runs through the lesion. The MRI control scan shows that both needles, which appear thick due to metal extinction artifacts, hit their target. This demonstrates the reliability of the registered multimodal image information and the great accuracy of the needle placement, even though the path planning for both needles was limited by the proximity to the lung. The remaining needle positions are not displayed here, but the control scans showed that all needles hit the targeted lesions. A CBCT control scan in Figure 5.6 (f) proves that the lesion B2 was accurately hit through two ribs in the posterior experiment.

5.4 Discussion

By combining several modern additive manufacturing techniques and molding, the manufacturing process of the phantom is reproducible and allows to incorporate patient-specific anatomies. Planning and designing the phantom took about six months and was the most time-consuming part. The 3D printed parts and the liver were built within two weeks. Once all the components were ready, it only took two days to put them together and complete the phantom. Here, the trickiest part was assembling and sealing the body casting hull. The hull pieces were printed with high accuracy using the SLA printing technique. Nevertheless, the parts slightly deformed during the UV curing process, which complicated the assembly and required careful sealing. Printing the hull with material that does not deform as much would make assembly less prone to error. Casting the hot ballistic gelatin into the torso in layers increased the safety in case of leaks in the hull and it facilitated the positioning of organs. In an initial prototype, we positioned each lesion within the liver with a pair of thin metal rods, which were removed after the liver was completed. This approach was inconvenient and left visible tracks in the liver, so we switched to attaching the lesions directly to the vascular tree.

The ballistic gelatin - which was used as a tissue surrogate - mimics the ballistic properties of muscle tissue. As demonstrated by Pepley *et al.*, the ballistic gelatin provides a needle resistance similar to cadaveric tissue [116]. However, due to the

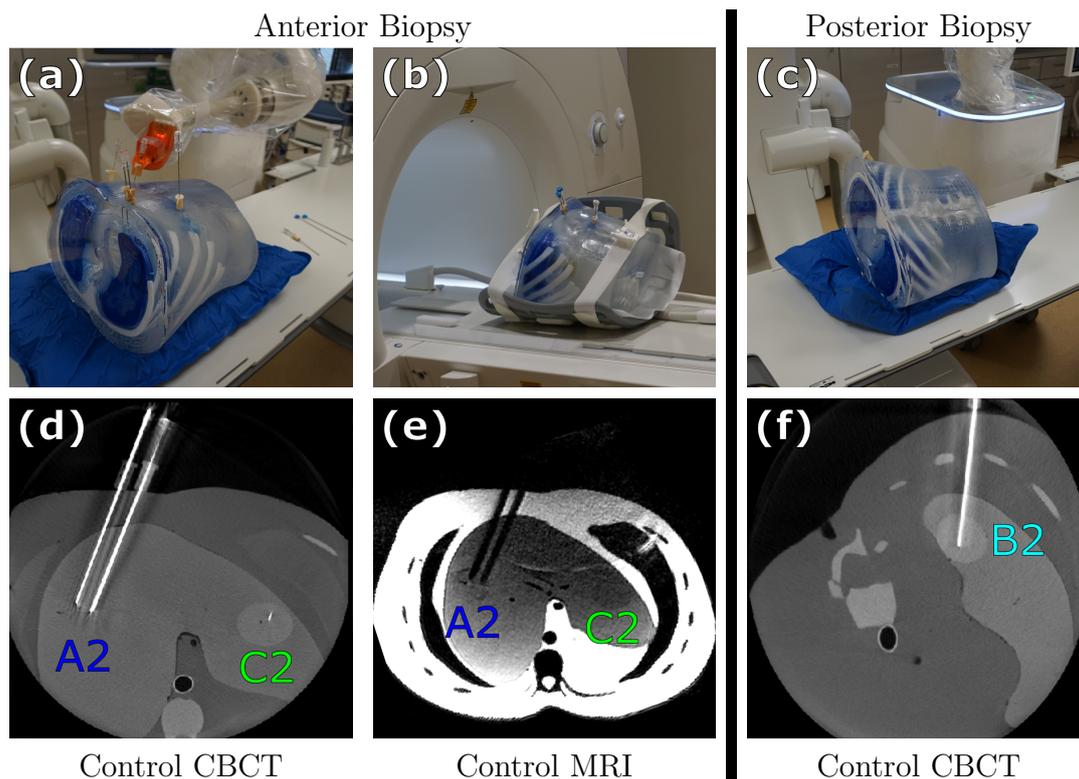


Figure 5.6: (a)-(c) Experimental biopsy setup. (d)-(f) CBCT and MRI control scans.

homogeneity of ballistic gelatin and the missing skin layer, the realism of the haptic needle feedback is limited. The ballistic gelatin also tears more easily than soft tissue and skin. Yoon *et al.* report a Shore OO hardness of 15.06 ± 2.64 for livers without hepatic fibrosis [117]. They also observed a correlation between hepatic hardness and degree of liver fibrosis. In our phantom, a silicone with Shore A 5 hardness was used to cast the liver. According to the ASTM D2240-97 standard, no simple relationship between the Shore A and Shore OO hardness scales is known to exist. However, an empirical conversion table for control purposes is provided, which states that the hardness of Shore A 5 is similar to Shore OO 45. This suggests that the silicone liver in our phantom is harder than a healthy human liver and more closely resembles liver fibrosis. During the design phase of the phantom, we investigated softer silicones for the liver. We cast a liver with Shore A 0 silicone, but the material tore after curing, resulting in large air pockets inside the liver. The type A and B liver lesions were cast with the same Shore A 5 silicone as the liver. This is not realistic, because tumors are stiffer than normal tissue [118]. Thus, a harder silicone with Shore A 19 was used for the type C liver lesions. The lungs, bones, and liver vessels serve as realistic inhibitions for needle-based intervention planning. Since they are solid, the operator immediately notices if an organ at risk was hit during needle placement. Due to the realistic size of the corpus and the organs, the phantom can be used to train clinicians in liver biopsies and other interventions. To make the phantom even more realistic, an integration of breathing movements is conceivable for the future. This has already been implemented in other phantoms [103], but simultaneous suitability for biopsies and inclusion of bones is difficult to realize.

The reusability of the phantom is limited by the extent of damage in the silicone liver and silicone lesions, because puncture channels in the silicone cannot be repaired. The extent of damage depends on the number of needle insertions, the needle gauge, and whether a biopsy sample was taken. We found that 16 gauge needles (G16) left visible tracks in the phantom. For voxel sizes below 1 mm, the tracks can be visualized in CT and MRI. Damages to the ballistic gelatin can be removed by remelting the material. This can be achieved via a hot metal rod, such as a soldering iron. Blemishes on the surface can be corrected with a hot air dryer. In case of more severe damage, the phantom can be remelted as a whole in an oven by reapplying the body hull mold. The separation of the body hull into several individual parts allows to reuse the mold shown in Figure 5.3 (b), because a non-destructive removal is possible after molding. If a hull part does get damaged, it can be replaced quickly and cost-effectively.

In our phantom experiment, we were able to hit all six liver lesions with robotic assistance. No corrections to the needle positions were necessary after initial needle placement. Because the phantom does not simulate realistic patient motion, these results cannot be directly applied to patient biopsies. In order to make statements about the reliability of biopsies in patients, a patient study must be carried out. Lungs, bones, and hepatic vessels were visible in the multimodal image data and could thus be bypassed. Using the phantom for the biopsy experiment allowed us to acquire CBCT and MRI control scans to assess the needle positions. The performance of the envisioned imaging workflow on the anthropomorphic phantom reflects the clinical reality where extensive multimodal imaging is beneficial prior to surgery to gather all necessary imaging information.

The liver lesions were embedded into the liver, which leads to a boundary that is visible under US. In reality, a large portion of liver tumors are not visible under US, which is the main reason for image fusion for intra-operative guidance [119]. To make the phantom suitable for use in US, it would be desirable to include lesions in the phantom that are invisible under US. It is a difficult task to remove the boundary between lesions and liver, but it might be achievable for the type A and B lesions, because they are made of the same Shore A 5 silicone as the liver. The boundary might become irrelevant, if the lesions would be embedded into the liver while the lesions are still in the curing process. This is difficult to achieve, because the time window during which the silicone already exhibits sufficient form stability during the curing process is short.

Moreover, the multimodal properties of the phantom enable the verification of MRI-based radiotherapy treatment planning methods. Currently, CT images are the basis of dose calculations in treatment planning. A crucial step towards MRI-only radiotherapy treatment planning is accurate MRI-to-CT synthesis [120], for which this phantom provides co-registered ground truth data.

5.5 Conclusions

Hitherto, there was a lack of appropriate multimodal phantoms with liver lesions of variable visibility. We designed and manufactured an abdominal phantom and used it to validate a standardized OMD diagnosis workflow. Utilizing the phantom,

we were able to show that a multimodal imaging pipeline is advantageous for a comprehensive detection of hepatic lesions. In a CBCT-guided needle placement experiment we punctured lesions that are invisible in CBCT using registered pre-interventional MRI scans for needle path planning. Accurate diagnosis and biopsy of hepatic lesions are important steps in diagnosing OMD and thus prolong the survival of patients. In the future, we plan to further modify the visibility of the liver lesions to make the phantom ultrasound compatible for percutaneous needle insertion.

6. ”An anthropomorphic pelvis phantom for MR-guided prostate interventions”, *Magn Reson Med, doi: 10.1002/mrm.29043*

6.1 Introduction

Prostate cancer is the second most common cancer in men and comprises highly aggressive and indolent varieties [121]. Multiparametric magnetic resonance imaging (mpMRI) is of high value for the detection and characterization of prostate cancer and reduces both overdiagnosis and overtreatment of indolent disease [122–124]. Moreover, mpMRI can be applied for prostate cancer management and therapy planning to localize index lesions [125, 126]. Accurate localization of the index lesion via mpMRI enables targeted MRGB and allows for optimized radiation dose planning for brachytherapy [127, 128].

Compared with traditional transrectal ultrasound (TRUS)-biopsy, which involves a random sampling of the prostate [129], mpMRI combined with targeted MRGB results in significantly fewer clinically insignificant prostate cancer (insignPCa) cases while maintaining an identical detection rate of clinically significant prostate cancer (csPCa) [130]. The improved disease localization of mpMRI additionally enables focal therapies for intermediate-risk prostate cancer, providing an alternative to radical treatment. This reduces overtreatment and associated permanent side-effects and is an important step toward personalized medicine [131]. Examples for minimal invasive focal therapies are cryoablation [132], focal laser ablation (FLA) [133], transurethral ultrasound ablation (TULSA) [134], high-intensity focused ultrasound (HIFU) [135], irreversible electroporation (IRE) [136] and brachytherapy [128].

These minimal invasive therapies often benefit from new technologies, such as robotic needle guidance, which supports in-bore MRI interventions. Robotic navigation systems improve the precision of needle placements and reduce intervention time [112, 137]. Another advantage of robots is that the needle probe can be remotely controlled from the MRI control room with real-time MRI guidance [138]. Preliminary validation using imaging phantoms is a vital step for the introduction of such technologies into the clinical routine. Numerous prostate phantoms have been developed for different imaging modalities. However, these are either not embedded in a realistic model of the pelvis [139–143] or are not puncturable [144, 145] and thus not suitable for needle interventions. Additionally, the design and development of multimodal phantoms incorporating lesions is not trivial [99] and none of the mentioned phantoms include prostate lesions. To the best of our knowledge, only the Tissue Equivalent Ultrasound Prostate Phantom (CIRS Model 053L, CIRS Inc., Norfolk, USA) - a commercially available prostate phantom - features puncturable prostate lesions and is suitable for CT, MR and US imaging. It also includes the urethra,

the rectal wall and seminal vesicles, but no bones. The phantom is disposable and must be stored in an air-tight container to minimize desiccation. The organs are embedded in a small and lightweight box with dimensions of $11.5 \times 7 \times 9.5 \text{ cm}^3$ and weight of approximately 900 g. However, realistic body shape and size are critical to the development of imaging protocols, and sufficient weight is required to simulate patient placement and prevent the phantom from moving during robotic needle placement [146]. Nevertheless, the CIRS phantom is an appropriate phantom when realistic patient size, shape and weight are not relevant to the measurement.

In this work, we present an anthropomorphic pelvis phantom for transperineal and transrectal prostate interventions. The phantom is puncturable and the prostate and the contained four lesions are accessible through a hollow rectum. The size and shape is similar to that of a male pelvis and realistic contrast is provided in MR, CT and US imaging. Thus, the phantom can be used for the validation of workflows in a variety of minimal invasive interventions, such as TRUS, IRE, brachytherapy seed placement or MRGB. We used the phantom to perform an in-bore MRGB with assistance of an MRI-compatible RCM where we targeted a prostate lesion. The phantom does not require specific storage and is reusable even after puncturing multiple times, as the needle insertion channels can be largely removed by heating the phantom material. Furthermore, the manufacturing process allows for the creation of patient-specific phantoms.

6.2 Materials and Methods

6.2.1 Phantom design

The phantom was designed to mimic the human pelvis incorporating bones, bladder, prostate with four lesions, urethra, arteries, veins and six lymph nodes embedded in ballistic gelatin [147]. A CAD model of the phantom is displayed in Figure 6.1 (a) and the manufactured phantom is shown in Figure 6.1 (c). Except for the bladder and prostate, all organ models were obtained from segmentation of the digital XCAT phantom [148], which provides realistic human anatomies. Because prostate cancer incidence is strongly correlated with age [149], a 62 year old male model was used. The prostate and bladder models were created by rotating 2D prostate and bladder silhouettes. The prostate has an enlarged volume of 95 ml, because older age is associated with prostate enlargement [150].

We used synthetic ballistic gelatin (10% ballistic gelatin, Clear Ballistics, Greenville, USA) as a tissue surrogate. Ballistic gelatin liquefies at 110°C and therefore can be poured into the desired shape, which it retains after cooling down. The other phantom components and the body mold must be heat-resistant, as they come into contact with the molten ballistic gelatin.

6.2.2 Phantom construction

In order to cast the ballistic gelatin into the shape of a pelvis, a body hull casting mold made of heat-resistant synthetic resin (High Temp Resin V2, Formlabs, Somerville, USA) was manufactured using the SLA 3D printing method. The body hull casting mold consists of 16 individual parts. This facilitated the removal of the

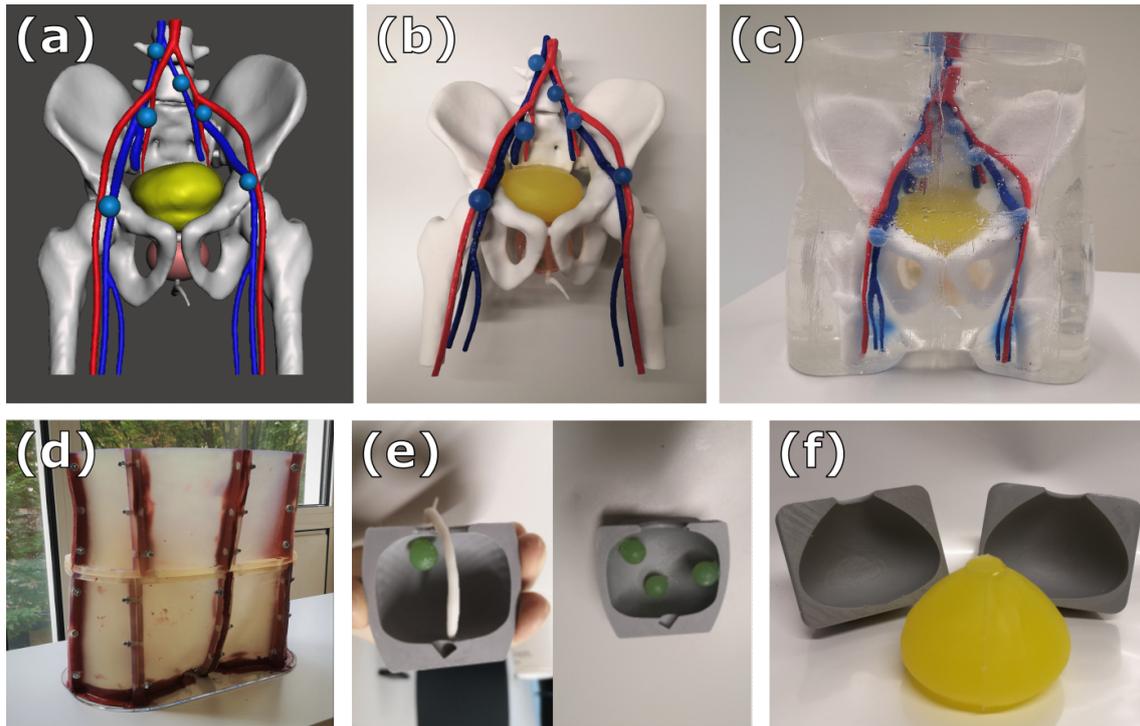


Figure 6.1: (a) 3D model of the organs. (b) Manufactured organs. (c) Final phantom: Organs embedded in ballistic gelatin. (d) Body casting mold including base plate. (e) Prostate casting mold including prostate lesions and urethra. The lesions and urethra were placed inside the mold before casting. (f) Bladder and bladder casting mold.

hull after casting without damaging it. The hull shown in Figure 6.1 (d) was firmly assembled with screws and sealed with high-temperature silicone. The bottom of the mold was closed with an extruded acrylic plate and sealed with high-temperature silicone. The base plate was removed after casting. Using heat-resistant adhesive, all organs were first glued together as shown in Figure 6.1 (b) and then glued on the inside of the base plate. A rectum was printed with the same material as the body hull and glued to the inside of the hull. The rectum was removed after casting, leaving behind a hollow transrectal access to the prostate. The complete pelvis was not cast in one step, but in several layers, which allows the removal of air bubbles from the intermediate layers.

The bones were 3D printed via SLS using polyamide powder filled with glass particles (PA-GF, Materialise, Leuven, Belgium). The material has excellent stiffness, high thermal resistance and tensile strength. The inside of the bones was printed less densely to simulate bone marrow. The artery and vein were printed with red and blue High Impact Polystyrene (EasyFil HIPS, Formfutura, Nijmegen, The Netherlands) via FFF. The urethra was printed with synthetic resin (White Resin V4, Formlabs, Somerville, USA) via SLA.

The prostate and bladder molds displayed in Figure 6.1 (e) and (f) were printed with polylactic acid (PLA) via FFF. The bladder and prostate were molded with silicones of Shore A hardness 0 and 5, respectively. The bladder mold was coated with a silicone of Shore A hardness 13 (SF13 - RTV2, Silikonfabrik, Ahrensburg,

Germany) to obtain a slightly harder bladder wall. The prostate lesions and lymph nodes were also molded with the Shore A 13 silicone. Silicone color was used to color the bladder, prostate, prostate lesions and lymph nodes yellow, red, green and blue, respectively. The diameter of the spherical prostate lesions is 1.2 cm and 1.5-2 cm for the lymph nodes.

6.2.3 Multimodal imaging

T1w, T2w and diffusion weighted (DWI) MRI at 3T and CT scans of the phantom were acquired and parametric T1-, T2- and ADC-maps were calculated. Furthermore, field inhomogeneities were characterized by calculating a B_0 -map from a double-echo gradient echo sequence [151]. The scanning parameters of the CT and MRI measurements are summarized in Table 6.1. The T1-, T2-times and the CT numbers for the prostate, prostate lesions and torso gel were determined by manual segmentation in the relaxometry maps and CT image, respectively. Additionally, a transcutaneous US scan of the phantom was obtained using a DC-N3 (Mindray, Shenzhen, China) US system.

Table 6.1: Summary of the CT and MRI scan parameters. Pre-interventional and interventional MRI measurements were acquired on the Skyra and Aera scanner, respectively. All scanners are manufactured by Siemens Healthineers, Germany. The used MRI sequences are GRE: Gradient Echo, SE: Spin Echo, VFA: Variable Flip Angle and EPI: Echo Planar Imaging. For the ADC measurement, $b = [50, 150, 300, 600, 800]$ was used [152].

	Reco. Kernel	Voltage [kVp]	Current [mA]	Pitch factor	Resolution [mm ³]/ Matrix size	Scanner Model
CT	Br36	90	105	0.35	$0.7 \times 0.7 \times 0.5$ / [512, 512, 620]	SOMATOM Force
	Sequence/ Plane	TE [ms]	TR [ms]	FA [°]	Resolution [mm ³]/ Matrix size	Scanner Model
T1w MRI	2D GRE/ axial	2.48	509	35	$1.3 \times 1.3 \times 3$ / [256, 192, 90]	MAGNETOM Skyra
T2w MRI	2D SE/ axial	187	10780	150	$1.3 \times 1.3 \times 3$ / [256, 192, 90]	MAGNETOM Skyra
T1-Map	2D VFA/ axial	2.48	509	35/70	$1.3 \times 1.3 \times 3$ / [256, 192, 60]	MAGNETOM Skyra
T2-Map	2D SE/ axial	33/110/187	10780	150	$1.3 \times 1.3 \times 3$ / [256, 192, 90]	MAGNETOM Skyra
B_0 -Map	2D GRE/ axial	4.92/7.38	793	60	$4.7 \times 4.7 \times 3.75$ / [64, 64, 75]	MAGNETOM Skyra
ADC-Map	2D EPI/ axial	61	3400	90	$1.8 \times 1.8 \times 3.5$ / [114, 102, 20]	MAGNETOM Skyra
T2w MRI (Interventional)	2D SE/ sagittal	112	7500	160	$0.6 \times 0.6 \times 3.5$ / [320, 320, 24]	MAGNETOM Aera

6.2.4 MRI-guided biopsy experiment

To demonstrate the feasibility of the phantom for minimal invasive MRI-guided interventions, a targeted in-bore MRGB was performed in a MAGNETOM Aera MRI scanner (Siemens Healthineers, Erlangen, Germany). The needle probe was rectally inserted and guided using an MRI-compatible remote controlled manipulator

(RCM, Soteria Medical, Arnhem, The Netherlands) [137]. The phantom was placed head first and prone in the scanner and the RCM was placed between the phantom's legs. The biopsy setup of the robot and the phantom is shown in Figure 6.2 (a).

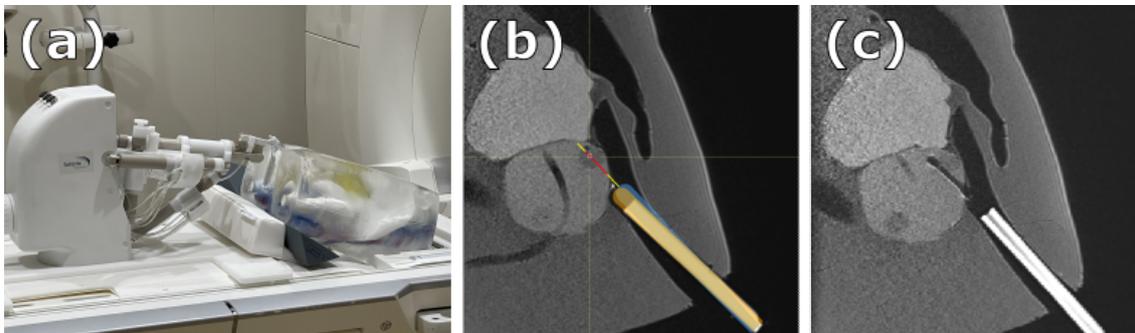


Figure 6.2: (a) In-bore MRGB setup. (b) Initial T2w MRI scan, where the current position of the needle probe is shown in yellow and the target position in blue. (c) Control T2w MRI scan during biopsy with the biopsy needle in place.

A T2w MRI image was acquired to localize the lesions and the initial position of the needle probe. The scan parameters of the interventional MRI are listed in Table 6.1. The T2w images were sent to a standalone computer located in the MRI control room, on which dedicated software automatically detected the needle probe. The biopsy target was determined in the software and the needle probe was remotely steered towards the target position. In our experiment we only punctured one lesion using an MRI-compatible titanium biopsy needle (Innovative Tomography Products, Bochum, Germany). After the needle application, a T2w MRI control scan was acquired to confirm the needle position.

6.3 Results

6.3.1 Multimodal imaging

In Figure 6.3 a CT scan, T1w and T2w MRI scans, T1-, T2-, B_0 - and ADC-maps of the phantom are shown. The positions of the prostate lesions are indicated by red arrows. The prostate lesions are visible in CT and T2w MRI and are not detectable in T1w MRI. Only small air bubbles on the surface of the lesions indicate the position of the lesions in the T1w image. The same air bubbles are visible to a lesser extent in the CT and T2w images. With a CT number of 500 HU, the bones have a high contrast in CT imaging and yield no signal in MRI.

The phantom's quantified T1-, T2-relaxation times and CT numbers are compared to literature values in Table 6.2. Since the lesions in the T1w MRI image and the T1-map showed no contrast to the prostate, the same T1-time was measured for the prostate and the prostatic lesions. The phantom's T1-times are 21% and 7% smaller than the lowest values and 44% and 28% smaller than the median values reported in human prostate and prostate lesions, respectively. T2 times are in very good agreement with literature values, and an increased T2-time of the prostate lesions results in their desired hypointense contrast between

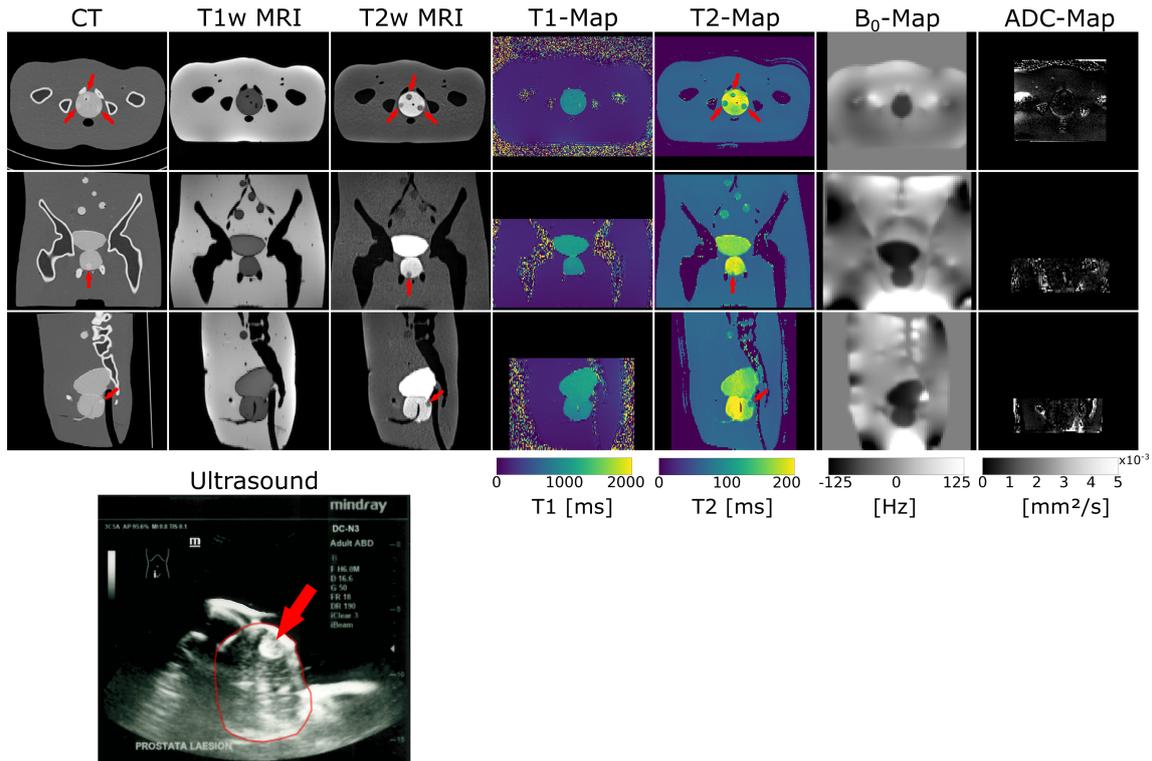


Figure 6.3: CT scan, T1w and T2w MRI scans, T1-, T2-, B₀- and ADC-maps of the pelvis phantom. Below a US scan of the prostate is shown in which the prostate is outlined in red. The prostate segmentation was performed manually. The unedited ultrasound scan is shown in Figure 10.4. In images where prostate lesions are visible, the position of the lesions is indicated by red arrows.

prostate and lesions is achieved in CT, however, the absolute CT numbers measured in the phantom are 140 HU higher than for real tissue. The B₀-map indicates high frequencies between the phantom's legs due to the air gap and low frequencies for the silicone organs. The low ADC values in the ADC-map show that there is very little diffusion in the phantom, which is underlined by a reference measurement of an agarose diffusion phantom in Figure 10.5.

Table 6.2: T1-, T2-relaxation times and CT numbers of prostate, prostate lesions and torso gelatin. The values and uncertainty are either given as mean \pm standard deviation or as median (min-max).

	Prostate		Prostate Lesions		Ballistic Gelatin
	Phantom	Literature	Phantom	Literature	Phantom
T1 [ms]	961 \pm 69	1666 (1222-2343) [153]	961 \pm 69	1328 (1037-1532) [153]	230 \pm 8
T2 [ms]	158 \pm 9	139 \pm 26 [154]	101 \pm 11	107 \pm 18 [154]	62 \pm 2
CT # [HU]	185 \pm 16	45 \pm 17 [155]	259 \pm 17	Hyperdense [156]	-180 \pm 16

A US scan of the phantom is shown in Figure 6.3, where the prostate is outlined in red. The good visibility in US makes the phantom viable for interventional US applications.

6.3.2 MRI-guided biopsy experiment

In a transrectal MRGB experiment we punctured a prostate lesion with a biopsy needle. The T2w MRI scan in Figure 6.2 (b) shows the initial position of the needle probe in yellow and the target position in blue. The whole cross section of the rectrum, needle probe, prostate and prostate lesion are visible in this sagittal scan. The position of the placed biopsy needle is displayed in the T2w MRI control scan in Figure 6.2 (c). The needle appears thick due to susceptibility artifacts. The scan shows that the lesion was hit, but the center was missed by 3.5 mm. We did not take a biopsy sample in order to spare the phantom from unnecessary damage.

6.4 Discussion

An anthropomorphic pelvis phantom incorporating a prostate with four lesions was developed. The realistic size of the pelvis allows evaluation of complete clinical workflows from patient positioning to imaging and intervention. The manufacturing process includes additive manufacturing techniques, enabling reproducibility and patient-specific anatomies.

The phantom is reusable, because most of the damages can be removed after usage by remelting the ballistic gelatin. Damages on the surface can be eradicated with a hot air dryer and puncture channels can be removed using a hot metal rod. The phantom can also be remelted as a whole in an oven after reapplying the body hull. However, damage to the silicone structures (prostate, lesions, lymph nodes and bladder) cannot be repaired.

Since ballistic gelatin mimics the properties of muscle tissue, the phantom simulates realistic feedback during needle interventions. Pepley *et al.* reported similar needle insertion forces for ballistic gelatin and cadaveric tissue [116]. Compared to the prostate, a harder silicone was used for the prostate lesions, because tumors are stiffer than normal prostate tissue [157]. The harder silicone has a higher density and therefore yields a higher signal in CT and enables lesion visibility in US and potentially in elastography. The bladder, urethra and bones serve as risk structures for needle interventions. Since bones and the urethra are solid, they realistically restrict possible needle paths. The lymph nodes act as additional percutaneous needle targets, with blood vessels as risk structures. With its realistic needle feedback, the phantom is useful for training surgical staff in prostate biopsies and other procedures. In the future, it is conceivable to integrate deformations to simulate organ motion [145, 158]. However, simultaneous suitability for needle interventions is difficult to achieve.

Our results demonstrate that the artificial pelvis shows human-like contrast in several MRI sequences and CT. Due to their high contrast, the bones provide high quality landmarks in MRI and CT, which facilitates multimodal image registration. The high visibility of the prostate lesions in T2w MRI scans in combination with the hollow rectum is central for targeted prostate interventions. The contrast of the lesions is consistent with reality, as prostatic lesions appear hypointense in T2w scans and are not detectable in T1w scans [124]. T1w imaging is mainly used to delineate bones and lymph nodes, for which the phantom provides an excellent contrast in the T1w scan. Quantification of relaxation times for prostate and prostate

showed excellent agreement of T2-times compared to literature values, whereas the T1-times of the phantom are too small. Nonetheless, the discrepancy of T1-times is still acceptable, since we achieved desired contrasts in the T1w and T2w MRI scans. Prostate lesions usually exhibit a decrease in ADC value [152]. Our phantom shows very little diffusion and the lesions are not visible in the ADC-map. Nevertheless, since the lesions are well detectable in T2w imaging, we do not need to rely on DWI to locate the lesions. Manipulating the relaxation times and diffusion properties of silicones is a challenging task, because additives like contrast agent often do not mix well and settle during the curing process. Agarose gels are a great alternative to silicones, because it is easy to manipulate T1-, T2-times and ADC value by varying the concentrations of agarose, gadolinium trichloride and sucrose [159, 160]. However, agarose gels were not suitable for our biopsy phantom, because as shown in Figure 10.3, they suffer from dehydration and mold-buildup, which substantially reduces their life span. The synthetic ballistic gelatin and silicones used in our phantom do not require specific storage. Our oldest phantom made from these materials is now three years old and we have yet to see any noticeable aging.

In the transrectal MRGB experiment we hit the targeted prostate lesion, but missed the center by 3.5 mm. The reason for this inaccuracy is that the RCM stops moving when the needle probe experiences resistance, and therefore cannot reach its target position. This is an intended feature of the robot to increase patient safety [137]. The limited size of the rectum therefore represents a realistic impairment for the needle placement. In future experiments we will lubricate the needle probe to reduce the resistance. Additionally, the center of rotation of the needle probe can be adjusted manually. By iteratively moving the probe and acquiring T2w MRI scans, the needle placement accuracy can be increased. In this regard, the limited size of the rectum is valuable for the training of interventionalists, as they can gain experience with such a challenging situation without causing discomfort or harm to patients. Another reason for the inaccurate needle placement could be needle bending, but we believe its contribution to be minor.

The good visibility of prostate and prostate lesions in US makes it possible to use the phantom for the validation of US imaging workflows and US-guided interventions. These include MRI-US fusion and TRUS-biopsy, for which the transrectal access is particularly convenient. Moreover, the multimodal properties of the phantom enable the verification of MRI-to-CT synthesis, which is essential for MRI-only radiotherapy treatment planning [120].

6.5 Conclusions

We have designed and manufactured a reusable pelvis phantom that is suitable for MRI, CT and US imaging and exhibits realistic imaging properties for prostate and prostate lesions. The phantom has been used to perform an in-bore MRGB with assistance of an MRI-compatible RCM. We have been able to hit a prostate lesion after locating it with MR imaging. Difficulties targeting the lesion have been identified and can be mitigated in the future. The phantom can be used for a variety of minimal invasive interventions and in the future will be used for the validation of an MRI-guided brachytherapy seed placement workflow using the RCM.

7. Summary

In the first part of this thesis, two frameworks for the generation of synthetic data were presented. The digital phantoms are useful to overcome the data sparsity in medical image processing. In [Chapter 3](#), a CycleGAN network was used to synthesize abdominal MRI, CT, and CBCT data. Since the images are inherently registered and organ masks are available, they can be used as a ground truth dataset for image registration and image segmentation. The dataset was eventually used to optimize a multimodal non-rigid image registration by evaluating different parameter settings. In [Chapter 4](#), a pipeline for the simulation of metal artifacts was developed. Metal objects were added into specified positions in the XCAT phantom and noise and beam hardening were simulated. The data was used to train an end-to-end deep learning CT reconstruction for the correction of metal artifacts. Since both frameworks are based on the XCAT phantom, arbitrary patient anatomies can be generated. This allows to create large datasets with highly individualized patient geometries, which can be used for the development of a wide range of medical imaging algorithms.

The second part of this thesis focuses on the manufacturing of physical phantoms for the validation of interventional workflows. An abdomen and a pelvis phantom were presented in [Chapter 5](#) and [Chapter 6](#), respectively. Both phantoms were made with similar materials and a similar manufacturing process. Lesions are embedded inside the liver of the abdomen phantom and inside the prostate of the pelvis phantom. Image-guided biopsies targeting the lesions were performed on both phantoms. The suitability of the phantoms for MRI and CT imaging enabled the acquisition of a multimodal pre-interventional dataset, which was fused to the interventional image modality via image registration. Thus, the phantoms allow to perform the whole imaging pipeline, including patient positioning, image acquisition, and image processing. The use of imaging phantoms allowed to safely validate robotic needle guidance systems. Medical personnel can be trained for various procedures using the phantoms, without risking harm to patients. Potential dangers can be discovered at this stage and avoided in the future.

A detailed summary of the four scientific studies presented in [Chapter 3](#) to [Chapter 6](#) is provided below.

Generation of annotated multimodal ground truth datasets for abdominal medical image registration,

Int J Comput Assist Radiol Surg, doi: 10.1007/s11548-021-02372-7

In [Chapter 3](#), a deep learning framework was developed to generate multimodal synthetic data. A CycleGAN network was used to transfer the style of real patient images onto the geometry provided by the XCAT phantom. The use of custom intensity and gradient loss functions improved the preservation of the XCAT geometry. A dataset consisting of abdominal T1-weighted MRI, CT, and CBCT images was created. Realistic noise texture and magnitude was achieved for all modalities.

The multimodal images are inherently registered and a simulation of breathing motion and heartbeat is provided. This makes them ideal to use as a ground truth

dataset for image registration. In addition, organ masks are provided that can be used to evaluate registration accuracy for individual organs or to train deep learning segmentation networks. To demonstrate the usefulness of the dataset for registration algorithm optimization, we performed a multimodal non-rigid registration for different parameter settings. The MRI, CT, and CBCT images in the inhaled state were registered to the CT image in the exhaled state. We assessed the quality of the registrations by calculating the overlap of the post-registration and ground truth liver masks. An optimal parameter setting was found for the registration of each modality. For CT and CBCT the MMI registration metric with small B-spline grid point spacing provided the best results, while for MRI the NC metric with large grid point spacing performed best.

The presented framework can not only be used to create completely new datasets, but also to extend existing datasets. An extension of small datasets using synthetic images can be used in combination with classical data augmentation to increase the training dataset of deep learning networks.

End-to-End Deep Learning CT Image Reconstruction for Metal Artifact Reduction,

Appl Sci, doi: [10.3390/app12010404](https://doi.org/10.3390/app12010404)

In [Chapter 4](#), a deep learning CT reconstruction technique called iCTU-Net was trained for the correction of metal artifacts. Inspired by the classical FBP, the network includes filtering and back projection layers. The whole reconstruction was trained end-to-end with a learnable back projection operation. The inputs of the iCTU-Net are sinograms and the outputs are reconstructed images. This allows the network to freely adapt the reconstruction to imperfections of the sinogram data. The network was trained in a supervised manner with metal-affected sinogram input data and artifact-free reconstructions as labels. A CT data generation framework was established to simulate metal artifacts in XCAT phantoms. First metal objects were inserted into blood vessels by utilizing the organ masks of the XCAT. Then a polychromatic forward projection was performed to obtain sinograms containing beam hardening. Finally, Poisson and Gaussian noise was added to the sinograms.

In a configuration study, different post-processing layers and loss functions for the iCTU-Net were investigated. Using a U-Net for post-processing after the back projection and using the SSIM loss provided the best results. In an input study, it was found that the network performs similarly for metal-free, metal-affected and NMAR pre-corrected sinogram input data. This showed that the NMAR pre-correction is not necessary. In a comparison study, we implemented NMAR and two state-of-the-art deep learning MAR methods. The iCTU-Net was the only MAR method that could provide reconstructions without any metal artifacts. For severe artifacts $SSIM = 0.970 \pm 0.009$ and $PSNR = 40.7 \pm 1.6$ was achieved, whereas the second best method only achieved $SSIM = 0.944 \pm 0.024$ and $PSNR = 39.8 \pm 1.9$. However, the reconstructions of the iCTU-Net were slightly blurry.

Development of an abdominal phantom for the validation of an oligometastatic disease diagnosis workflow,
submitted Med Phys, 19.05.2021

An abdominal phantom including a liver with six lesions and a hepatic vessel tree, a portal vein, bones, and lungs was presented in [Chapter 5](#). The organs were produced via additive manufacturing with organ contours obtained from patient segmentations. This makes the manufacturing process reproducible and allows to create patient-specific anatomies. The phantom was used to validate an OMD diagnosis workflow that includes multimodal imaging, image registration, and interventional liver biopsy for subsequent molecular subtyping. Multimodal imaging is required because lesions cannot be comprehensively detected with only one imaging modality. Three different types of lesions with different visibility for MRI and CT were developed for the phantom.

Multimodal imaging of the phantom showed that all six lesions are visible in T2w MRI, whereas only lesion type C is visible in T1w MRI. Lesion types B and C are visible in CT and CBCT, but lesion type A is invisible. This is confirmed by a CNR of 0.1 and 0.04 for lesion type A in CT and CBCT, respectively. The signal difference between type A lesions and the liver is only 2 HU for CT and CBCT.

In a CBCT-guided biopsy, we targeted all six liver lesions with the assistance of a robot. To plan the needle path for the type A lesions, pre-interventional MRI data needed to be registered to the interventional CBCT. CBCT control scans showed that all type B and C lesions were hit. The needle position for lesion type A was verified in a MRI control scan. Lesion B2 was accessible only through a posterior percutaneous route, in which the needle was placed between two ribs. Bones and lungs served as risk structures and could successfully be avoided. The biopsy experiment highlighted the benefit of multimodal image data for image-guided interventions. The reliability of multimodal image registration and good accuracy of needle placement were demonstrated.

An Anthropomorphic Pelvis Phantom for MR-guided Prostate Interventions,
Magn Reson Med, doi: 10.1002/mrm.29043

[Chapter 6](#) introduced an anthropomorphic pelvis phantom for transperineal and transrectal prostate interventions. The phantom incorporates bones, a bladder, a urethra, arteries, veins, six lymph nodes, and a prostate with four lesions. Similar to the abdomen phantom in [Chapter 5](#), several 3D printing and casting techniques were used to manufacture the phantom. The 3D models for all organs, except the bladder and prostate, were obtained from segmentations of a 62-year-old male model of the XCAT phantom. The organs were embedded in ballistic gelatin in the shape of a human-sized pelvis. A hollow rectum enables transrectal access to the prostate and prostate lesions. The phantom is suitable for needle interventions, because the silicone prostate and prostate lesions and the ballistic gelatin are puncturable. Damages to the rectum and the torso material can be repaired by remelting the ballistic gelatin.

Multimodal imaging showed realistic contrasts in MRI, CT, and US imaging. Bones yield no signal in MRI and have a CT number of 500 HU. Lymph nodes are visible

in T1w and T2w MRI and CT. The prostate lesions are conspicuous in CT and T2w MRI, and not detectable in T1w MRI. This is consistent with scans of patients, where T2w MRI is used for detection of lesions and T1w MRI is used to delineate bones and lymph nodes. Quantitative evaluation of relaxation times showed fantastic agreement of T2-times for prostate and prostate lesions with literature values, whereas T1-times are too small. While the CT numbers for prostate and prostate lesions are slightly elevated, a realistic hyperdense signal of the lesions is achieved.

The phantom was used to perform an in-bore MRGB. A target inside a prostate lesion was defined in a T2w MRI scan and the needle probe was automatically steered by an RCM. A T2w MRI scan after needle placement confirmed that the targeted lesion was hit. The solid bones and the urethra served as risk structures and had to be avoided by the needle path planning. This was easily realizable, because they were visible in MRI. It was found that the limited size of the rectum triggered a safety mechanism of the RCM that caused the needle sample to stop abruptly before it could reach its target position. This is a scenario that can also occur in real patients. The phantom allows medical personnel to be prepared for such situations and thus the procedure can be performed more safely and quickly.

8. Outlook

In [Chapter 3](#) a framework for the synthesis of multimodal image data using a CycleGAN network was presented to provide a solution to the data sparsity in medical imaging. In recent years a plethora of GANs for image-to-image translation have been introduced [161]. To optimize the performance of the presented image generation framework, it is reasonable to explore different GAN models in the future. A promising candidate is the StyleGAN2, which is able to generate realistic high-resolution images of human faces [162] and was already used for MRI-to-CT and CT-to-MRI translation [163]. The introduction of custom identity and gradient losses helped to preserve the geometry provided by the XCAT phantom. The preservation of the XCAT geometry is crucial in order to apply the XCAT organ masks to the synthetic images. Additional loss functions are summarized by Li *et al.* in a review paper about GANs in medical imaging [161]. The impact of these loss functions on the presented image generation framework will be investigated in the future. While abdominal data were generated in the present work, synthesis of data from other body regions such as the thorax or pelvis is also possible. In addition, the framework can be expanded to include other modalities such as T2w-MRI, PET-CT, and US. This allows to create datasets for a variety of interventional imaging applications. The presented use case was the validation of multimodal image registration. The organ masks obtained within the framework also allow to use the data for the training of deep learning segmentation networks. This was demonstrated for blood vessel segmentation in CT images, where extending a patient dataset with synthetic data improved the segmentation performance [38]. Such segmentation networks can be trained for any organ and modality without the need for labor-intensive manual labeling of data.

The deep learning reconstruction in [Chapter 4](#) was trained to mitigate metal artifacts in the abdomen. With the artifact simulation framework used to generate the training data, it is possible to readily simulate data of the whole body. Thus, it is easy to simulate data for different scenarios, for example dental CT with high-density dental prostheses. It is also possible to simulate training data with different kinds of artifacts. For example, by decreasing the incident photon flux in [Equation 4.1](#) and [Equation 4.2](#), low-dose CT images with increased noise can be simulated. The iCTU-Net has already been successfully trained with low-dose CT data and was able to substantially reduce noise [77, 164]. The ability of the iCTU-Net to simultaneously mitigate different kinds of artifacts will be investigated in the future. This can be achieved by simulating a training dataset that contains a combination of artifacts, such as low-dose and metal artifacts. To improve the iCTU-Net in general, alternative loss functions and post-processing layers will be investigated. It has been found that these have a great influence on the quality of the reconstruction. With the U-Net, a rather deep network architecture was used for post-processing, which might have contributed to the blurring of the image. As an alternative, three simple convolutional layers were investigated, which were not able to provide artifact-free images. Post-processing layers more complex than the three convolutional layers but shallower than the U-Net may be able to effectively remove artifacts without blurring the image. Thus, post-processing layers with varying complexity will be

thoroughly investigated in future studies. Adding custom loss functions that focus on edge information to the used SSIM loss could also help to reduce blurring. So far, the iCTU-Net was trained and tested only on simulated data, but it is desirable to test the network on real scanner data. However, in its current form, the back projection layers in the iCTU-Net are not applicable to the two-dimensional projection data of multirow detector CT scanners, as they require a three-dimensional reconstruction. In theory, the back projection layers can easily be adapted to reconstruct three-dimensional images. However, the GPU's memory requirements are immense and cannot be handled with the currently available hardware. In order to use data from real scanners, the sinogram data must be downsampled in the first layers of the network or in a pre-processing step before training. Further modifications to the back projection layers might still be necessary to meet the memory constraints. Using separate input channels, the iCTU-Net is readily applicable to energy resolved raw data of dual energy and photon-counting CT. The additional energy information is expected to benefit the mitigation of beam hardening and thus will be investigated in the future.

By using modern additive manufacturing techniques and molding, it was possible to manufacture highly-individual physical phantoms in [Chapter 5](#) and [Chapter 6](#). While an abdomen and a pelvis phantom were presented in this work, it is feasible to manufacture phantoms of any other body region in the future. To make the phantoms more realistic, it would be desirable to implement respiratory movements. This is especially of interest for the abdomen phantom, because patient breathing is an important factor to consider for interventions of the liver. For the pelvic phantom, respiration is of little importance because the prostatic movements caused by respiration are negligible [165]. Organ motion, however, is difficult to realize if the phantom is supposed to be suitable for needle interventions. The ballistic gelatin and the silicones in the phantoms tear easily and the rigid bones are not flexible. The variable visibility of the liver lesions in the abdomen phantom help to show the advantages of a multimodal imaging pipeline for a comprehensive detection of lesions. However, because there is a boundary between the lesions and the liver, all lesions are visible under US, which does not reflect clinical reality. To make the lesions invisible under US, this boundary needs to be removed. In the future, this could be achieved for lesion type A and B, because they are made of the same Shore A 5 silicone as the liver. If the lesions would be embedded while they are still curing, the boundary might cure with the liver material, making the boundary negligible. However, this is difficult to achieve because the period during which the lesions are still curing but are already form stable is very short. Both phantoms presented in this work were used to validate a biopsy workflow using robotic assistance. The prostate biopsy was an initial experiment to set up the robotic system. Ultimately, the phantom will be used for the validation of MRI-guided prostate brachytherapy seed placements using the RCM. Of special interest will be the application of markers that are visible in MRI. Exploiting the multimodal properties of the pelvis phantom, it will be possible to compare the seed position in CT to the marker position in MRI.

9. Bibliography

- [1] T. R. Geiger and D. S. Peeper, “Metastasis mechanisms,” *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, vol. 1796, no. 2, pp. 293–308, 2009. (cited on Page 1)
- [2] C. L. Chaffer and R. A. Weinberg, “A perspective on cancer cell metastasis,” *science*, vol. 331, no. 6024, pp. 1559–1564, 2011. (cited on Page 1)
- [3] A. M. Clark, B. Ma, D. L. Taylor, L. Griffith, and A. Wells, “Liver metastases: Microenvironments and ex-vivo models,” *Experimental Biology and Medicine*, vol. 241, no. 15, pp. 1639–1652, 2016. (cited on Page 1)
- [4] S. R. Horn, K. C. Stoltzfus, E. J. Lehrer, L. A. Dawson, L. Tchelebi, N. J. Gusani, N. K. Sharma, H. Chen, D. M. Trifiletti, and N. G. Zaorsky, “Epidemiology of liver metastases,” *Cancer epidemiology*, vol. 67, p. 101760, 2020. (cited on Page 1)
- [5] M. Koshariya, R. B. Jagad, J. Kawamoto, P. Papastratis, H. Kefalourous, T. Porfiris, C. Tzouma, and N. Lygidakis, “An update and our experience with metastatic liver disease.,” *Hepato-gastroenterology*, vol. 54, no. 80, pp. 2232–2239, 2007. (cited on Page 1 and 47)
- [6] R. Konopke, J. Roth, A. Volk, S. Pistorius, G. Folprecht, K. Zoephel, C. Schuetze, M. Laniado, H.-D. Saeger, and S. Kersting, “Colorectal liver metastases: an update on palliative treatment options.,” *Journal of Gastrointestinal & Liver Diseases*, vol. 21, no. 1, 2012. (cited on Page 1)
- [7] E. Van Cutsem, A. Cervantes, R. Adam, A. Sobrero, J. Van Krieken, D. Aderka, E. A. Aguilar, A. Bardelli, A. Benson, G. Bodoky, *et al.*, “Esmo consensus guidelines for the management of patients with metastatic colorectal cancer,” *Annals of Oncology*, vol. 27, no. 8, pp. 1386–1422, 2016. (cited on Page 1, 47, and 49)
- [8] D. Li, J. Kang, B. J. Golas, V. W. Yeung, and D. C. Madoff, “Minimally invasive local therapies for liver cancer,” *Cancer biology & medicine*, vol. 11, no. 4, p. 217, 2014. (cited on Page 1)
- [9] D. Shen, G. Wu, and H.-I. Suk, “Deep learning in medical image analysis,” *Annual review of biomedical engineering*, vol. 19, pp. 221–248, 2017. (cited on Page 1)
- [10] J. Ker, L. Wang, J. Rao, and T. Lim, “Deep learning applications in medical image analysis,” *Ieee Access*, vol. 6, pp. 9375–9389, 2017. (cited on Page 1)
- [11] F. G. Zöllner, A. Šerifović-Trbalić, G. Kabelitz, M. Kociński, A. Materka, and P. Rogelj, “Image registration in dynamic renal mri—current status and prospects,” *Magn. Reson. Mater. Phy.*, vol. 33, pp. 33–48, 2020. (cited on Page 1 and 17)

- [12] T. D. Do, J. Heim, S. Skornitzke, C. Melzig, D. F. Vollherbst, M. Faerber, P. L. Pereira, H.-U. Kauczor, and C. M. Sommer, "Single-energy versus dual-energy imaging during ct-guided biopsy using dedicated metal artifact reduction algorithm in an in vivo pig model," *Plos one*, vol. 16, no. 4, p. e0249921, 2021. (cited on Page 1 and 31)
- [13] Z. Yaniv, P. Cheng, E. Wilson, T. Popa, D. Lindisch, E. Campos-Nanez, H. Abeledo, V. Watson, K. Cleary, and F. Banovac, "Needle-based interventions with the image-guided surgery toolkit (igstk): From phantoms to clinical trials," *IEEE transactions on biomedical engineering*, vol. 57, no. 4, pp. 922–933, 2009. (cited on Page 2)
- [14] V. P. Grover, J. M. Tognarelli, M. M. Crossey, I. J. Cox, S. D. Taylor-Robinson, and M. J. McPhail, "Magnetic resonance imaging: principles and techniques: lessons for clinicians," *Journal of clinical and experimental hepatology*, vol. 5, no. 3, pp. 246–255, 2015. (cited on Page 5)
- [15] R. Schofield, L. King, U. Tayal, I. Castellano, J. Stirrup, F. Pontana, J. Earls, and E. Nicol, "Image reconstruction: Part 1—understanding filtered back projection, noise and image acquisition," *Journal of cardiovascular computed tomography*, vol. 14, no. 3, pp. 219–225, 2020. (cited on Page 6)
- [16] F. E. Boas, D. Fleischmann, *et al.*, "Ct artifacts: causes and reduction techniques," *Imaging Med*, vol. 4, no. 2, pp. 229–240, 2012. (cited on Page 9 and 31)
- [17] T. Behr, T. P. Pusch, M. Siegfarth, D. Hüsener, T. Mörschel, and L. Karstensen, "Deep reinforcement learning for the navigation of neurovascular catheters," *Current Directions in Biomedical Engineering*, vol. 5, no. 1, pp. 5–8, 2019. (cited on Page 10)
- [18] T. P. Lillicrap, A. Santoro, L. Marris, C. J. Akerman, and G. Hinton, "Back-propagation and the brain," *Nature Reviews Neuroscience*, vol. 21, no. 6, pp. 335–346, 2020. (cited on Page 10)
- [19] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014. (cited on Page 13)
- [20] I. Gibson, D. W. Rosen, and B. Stucker, *Additive Manufacturing Technologies: Rapid Prototyping to Direct Digital Manufacturing*. Springer US, 2010. (cited on Page 14)
- [21] H. K. Lam, L. Ding, T. Cheng, and H. Zhou, "The impact of 3d printing implementation on stock returns: A contingent dynamic capabilities perspective," *International Journal of Operations & Production Management*, 2019. (cited on Page 15)
- [22] C.-Y. Liaw and M. Guvendiren, "Current and emerging applications of 3d printing in medicine," *Biofabrication*, vol. 9, no. 2, p. 024102, 2017. (cited on Page 15)

-
- [23] E. D. Ehler, "Utilization of 3d printing in clinical medical physics," in *AIP Conference Proceedings*, vol. 2348, p. 020004, AIP Publishing LLC, 2021. (cited on Page 15)
- [24] S. Singh, G. Singh, C. Prakash, and S. Ramakrishna, "Current status and future directions of fused filament fabrication," *Journal of Manufacturing Processes*, vol. 55, pp. 288–306, 2020. (cited on Page 16)
- [25] H. Kodama, "Automatic method for fabricating a three-dimensional plastic model with photo-hardening polymer," *Review of scientific instruments*, vol. 52, no. 11, pp. 1770–1773, 1981. (cited on Page 16)
- [26] J.-P. Kruth, X. Wang, T. Laoui, and L. Froyen, "Lasers and materials in selective laser sintering," *Assembly Automation*, 2003. (cited on Page 16)
- [27] C. M. Tempany, J. Jayender, T. Kapur, R. Bueno, A. Golby, N. Agar, and F. A. Jolesz, "Multimodal imaging for improved diagnosis and treatment of cancers," *Cancer*, vol. 121, no. 6, pp. 817–827, 2015. (cited on Page 17)
- [28] K. Memon, R. J. Lewandowski, L. Kulik, A. Riaz, M. F. Mulcahy, and R. Salem, "Radioembolization for primary and metastatic liver cancer," *Semin. Radiat. Oncol.*, vol. 21, no. 4, pp. 294–302, 2011. (cited on Page 17)
- [29] B. Waldkirch, S. Engelhardt, F. G. Zöllner, L. R. Schad, and I. Wolf, "Multimodal image registration of pre-and intra-interventional data for surgical planning of transarterial chemoembolisation," in *Proc. SPIE Med. Imag.*, vol. 10951, p. 109512U, SPIE, 2019. (cited on Page 17, 18, 47, and 49)
- [30] N. Spahr, S. Thoduka, N. Abolmaali, R. Kikinis, and A. Schenk, "Multi-modal image registration for liver radioembolization planning and patient assessment," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 2, pp. 215–225, 2019. (cited on Page 17)
- [31] J. P. Pluim, S. E. Muenzing, K. A. Eppenhof, and K. Murphy, "The truth is hard to make: Validation of medical image registration," in *2016 23rd International Conference on Pattern Recognition (ICPR)*, pp. 2294–2300, IEEE, 2016. (cited on Page 17)
- [32] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, G. B. Pike, and A. C. Evans, "Brainweb: Online interface to a 3d mri simulated brain database," *NeuroImage*, vol. 5, p. 425, 1997. (cited on Page 18)
- [33] S. Roy, A. Carass, A. Jog, J. L. Prince, and J. Lee, "Mr to ct registration of brains using image synthesis," in *Proc. SPIE Med. Imag.*, vol. 9034, p. 903419, SPIE, 2014. (cited on Page 18)
- [34] X. Cao, J. Yang, Y. Gao, Y. Guo, G. Wu, and D. Shen, "Dual-core steered non-rigid registration for multi-modal images via bi-directional image synthesis," *Med. Image Anal.*, vol. 41, pp. 18–31, 2017. (cited on Page 18)

- [35] J. M. Wolterink, A. M. Dinkla, M. H. F. Savenije, P. R. Seevinck, C. A. T. van den Berg, and I. Išgum, “Deep mr to ct synthesis using unpaired data,” in *Proc. Int. Workshop Simul. Synth. Med. Imag.*, pp. 14–23, 2017. (cited on Page 18)
- [36] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui, “4D XCAT phantom for multimodality imaging research,” *Med. Phys.*, vol. 37, pp. 4902–4915, Aug 2010. (cited on Page 18 and 35)
- [37] O. Tmenova, R. Martin, and L. Duong, “Cyclegan for style transfer in x-ray angiography,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 10, pp. 1785–1794, 2019. (cited on Page 18 and 23)
- [38] T. Russ, S. Goerttler, A.-K. Schnurr, D. F. Bauer, S. Hatamikia, L. R. Schad, F. G. Zöllner, and K. Chung, “Synthesis of ct images from digital body phantoms using cyclegan,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 10, pp. 1741–1750, 2019. (cited on Page 18, 24, and 73)
- [39] C. Paganelli, P. Summers, C. Gianoli, M. Bellomi, G. Baroni, and M. Riboldi, “A tool for validating mri-guided strategies: a digital breathing ct/mri phantom of the abdominal site,” *MBEC*, vol. 55, no. 11, pp. 2001–2014, 2017. (cited on Page 18 and 19)
- [40] L. Wissmann, C. Santelli, W. P. Segars, and S. Kozerke, “Mrxcat: Realistic numerical phantoms for cardiovascular magnetic resonance,” *J. Cardiovasc. Magn. Reson.*, vol. 16, no. 1, p. 63, 2014. (cited on Page 18 and 23)
- [41] S. Abbasi-Sureshjani, S. Amirrajab, C. Lorenz, J. Weese, J. Pluim, and M. Breeuwer, “4d semantic cardiac magnetic resonance image synthesis on xcat anatomical model,” *arXiv preprint arXiv:2002.07089*, 2020. (cited on Page 18)
- [42] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 2223–2232, 2017. (cited on Page 19)
- [43] D. Nie, R. Trullo, J. Lian, L. Wang, C. Petitjean, S. Ruan, Q. Wang, and D. Shen, “Medical image synthesis with deep convolutional adversarial networks,” *IEEE Trans. Biomed. Eng.*, vol. 65, no. 12, pp. 2720–2730, 2018. (cited on Page 20)
- [44] D. F. Bauer, A.-K. Schnurr, T. Russ, S. Goerttler, L. R. Schad, F. G. Zoellner, and K. Chung, “Synthesis of ct images using cyclegans: Enhancement of anatomical accuracy,” in *Proc. Int. Conf. Med. Imag. Deep Learning*, 2019. (cited on Page 21)
- [45] N. Jacobsen, A. Deistung, D. Timmann, S. L. Goericke, J. R. Reichenbach, and D. Güllmar, “Analysis of intensity normalization for optimal segmentation performance of a fully convolutional neural network,” *Z. Med. Phys.*, vol. 29, no. 2, pp. 128–138, 2019. (cited on Page 21)

-
- [46] G. J. Stanisiz, E. E. Odrobina, J. Pun, M. Escaravage, S. J. Graham, M. J. Bronskill, and R. M. Henkelman, “T1, t2 relaxation and magnetization transfer in tissue at 3t,” *Magn. Reson. Med.*, vol. 54, no. 3, pp. 507–512, 2005. (cited on Page 23)
- [47] C. M. De Bazelaire, G. D. Duhamel, N. M. Rofsky, and D. C. Alsop, “Mr imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 t: preliminary results,” *Radiology*, vol. 230, no. 3, pp. 652–659, 2004. (cited on Page 23)
- [48] Z. Wang, A. C. Bovik, and H. R. Sheikh, “Image quality assessment: From error measurement to structural similarity,” *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600 – 612, 2004. (cited on Page 24)
- [49] Lin Zhang, Lei Zhang, Xuanqin Mou, and D. Zhang, “FSIM: A Feature Similarity Index for Image Quality Assessment,” *IEEE Trans. Image Processing*, vol. 20, pp. 2378–2386, Aug 2011. (cited on Page 24)
- [50] L. Chen, F. Jiang, H. Zhang, S. Wu, S. Yu, and Y. Xie, “Edge preservation ratio for image sharpness assessment,” in *2016 12th World Congress on Intelligent Control and Automation (WCICA)*, pp. 1377–1381, IEEE, 2016. (cited on Page 24)
- [51] R. C. Geary, “The ratio of the mean deviation to the standard deviation as a test of normality,” *Biometrika*, vol. 27, no. 3/4, pp. 310–332, 1935. (cited on Page 28)
- [52] S. R. McWilliams, K. P. Murphy, S. Golestaneh, K. N. O’Regan, R. S. Arellano, M. M. Maher, and O. J. O’Connor, “Reduction of guide needle streak artifact in ct-guided biopsy,” *Journal of Vascular and Interventional Radiology*, vol. 25, no. 12, pp. 1929–1935, 2014. (cited on Page 31)
- [53] J. Stattaus, H. Kuehl, S. Ladd, T. Schroeder, G. Antoch, H. A. Baba, J. Barkhausen, and M. Forsting, “Ct-guided biopsy of small liver lesions: visibility, artifacts, and corresponding diagnostic accuracy,” *Cardiovascular and interventional radiology*, vol. 30, no. 5, pp. 928–935, 2007. (cited on Page 31)
- [54] K. R. Laukamp, D. Zopfs, A. Wagner, S. Lennartz, L. Pennig, J. Borggreffe, N. Ramaiya, and N. G. Hokamp, “Ct artifacts from port systems: Virtual monoenergetic reconstructions from spectral-detector ct reduce artifacts and improve depiction of surrounding tissue,” *European journal of radiology*, vol. 121, p. 108733, 2019. (cited on Page 31 and 46)
- [55] C. Kim, D. Kim, K. Y. Lee, H. Kim, J. Cha, J. Y. Choo, and P.-K. Cho, “The optimal energy level of virtual monochromatic images from spectral ct for reducing beam-hardening artifacts due to contrast media in the thorax,” *American Journal of Roentgenology*, vol. 211, no. 3, pp. 557–563, 2018. (cited on Page 31)

- [56] M. Lin, R. Loffroy, N. Noordhoek, K. Taguchi, A. Radaelli, J. Blijd, A. Balgund, and J.-F. Geschwind, “Evaluating tumors in transcatheter arterial chemoembolization (tace) using dual-phase cone-beam ct,” *Minimally Invasive Therapy & Allied Technologies*, vol. 20, no. 5, pp. 276–281, 2011. (cited on Page 31)
- [57] S. Jeong, S. H. Kim, E. J. Hwang, C.-i. Shin, J. K. Han, and B. I. Choi, “Usefulness of a metal artifact reduction algorithm for orthopedic implants in abdominal ct: phantom and clinical study results,” *American Journal of Roentgenology*, vol. 204, no. 2, pp. 307–317, 2015. (cited on Page 31)
- [58] R. N. Bismark, R. Frysch, S. Abdurahman, O. Beuing, M. Blessing, and G. Rose, “Reduction of beam hardening artifacts on real c-arm ct data using polychromatic statistical image reconstruction,” *Zeitschrift für Medizinische Physik*, vol. 30, no. 1, pp. 40–50, 2020. (cited on Page 31)
- [59] F. Bamberg, A. Dierks, K. Nikolaou, M. F. Reiser, C. R. Becker, and T. R. Johnson, “Metal artifact reduction by dual energy computed tomography using monoenergetic extrapolation,” *European radiology*, vol. 21, no. 7, pp. 1424–1429, 2011. (cited on Page 31 and 46)
- [60] W. A. Kalender, R. Hebel, and J. Ebersberger, “Reduction of ct artifacts caused by metallic implants.,” *Radiology*, vol. 164, no. 2, pp. 576–577, 1987. (cited on Page 31 and 36)
- [61] E. Meyer, R. Raupach, M. Lell, B. Schmidt, and M. Kachelrieß, “Normalized metal artifact reduction (nmar) in computed tomography,” *Medical physics*, vol. 37, no. 10, pp. 5482–5493, 2010. (cited on Page 31 and 36)
- [62] A. S. Lundervold and A. Lundervold, “An overview of deep learning in medical imaging focusing on mri,” *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 102–127, 2019. (cited on Page 31)
- [63] Y. Zhang and H. Yu, “Convolutional neural network based metal artifact reduction in x-ray computed tomography,” *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1370–1381, 2018. (cited on Page 31, 32, and 46)
- [64] M. U. Ghani and W. C. Karl, “Deep learning based sinogram correction for metal artifact reduction,” *Electronic Imaging*, vol. 2018, no. 15, pp. 472–1, 2018. (cited on Page 32)
- [65] M. U. Ghani and W. C. Karl, “Fast enhanced ct metal artifact reduction using data domain deep learning,” *IEEE Transactions on Computational Imaging*, vol. 6, pp. 181–193, 2019. (cited on Page)
- [66] L. Gjestebj, Q. Yang, Y. Xi, Y. Zhou, J. Zhang, and G. Wang, “Deep learning methods to guide ct image reconstruction and reduce metal artifacts,” in *Medical Imaging 2017: Physics of Medical Imaging*, vol. 10132, p. 101322W, International Society for Optics and Photonics, 2017. (cited on Page 32 and 33)

-
- [67] T. Lossau, H. Nickisch, T. Wissel, M. Morlock, and M. Grass, “Learning metal artifact reduction in cardiac ct images with moving pacemakers,” *Medical image analysis*, vol. 61, p. 101655, 2020. (cited on Page 32)
- [68] L. Gjestebj, Q. Yang, Y. Xi, B. Claus, Y. Jin, B. De Man, and G. Wang, “Reducing metal streak artifacts in ct images via deep learning: Pilot results,” in *The 14th international meeting on fully three-dimensional image reconstruction in radiology and nuclear medicine*, vol. 14, pp. 611–614, 2017. (cited on Page 32 and 33)
- [69] J. Wang, Y. Zhao, J. H. Noble, and B. M. Dawant, “Conditional generative adversarial networks for metal artifact reduction in ct images of the ear,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 3–11, Springer, 2018. (cited on Page 32)
- [70] X. Huang, J. Wang, F. Tang, T. Zhong, and Y. Zhang, “Metal artifact reduction on cervical ct images by deep residual learning,” *Biomedical engineering online*, vol. 17, no. 1, pp. 1–15, 2018. (cited on Page 32 and 46)
- [71] H. Liao, W.-A. Lin, S. K. Zhou, and J. Luo, “Adn: artifact disentanglement network for unsupervised metal artifact reduction,” *IEEE transactions on medical imaging*, vol. 39, no. 3, pp. 634–643, 2019. (cited on Page 32)
- [72] M. Nakao, K. Imanishi, N. Ueda, Y. Imai, T. Kirita, and T. Matsuda, “Regularized three-dimensional generative adversarial nets for unsupervised metal artifact reduction in head and neck ct images,” *IEEE Access*, vol. 8, pp. 109453–109465, 2020. (cited on Page)
- [73] J. Lee, J. Gu, and J. C. Ye, “Unsupervised ct metal artifact learning using attention-guided β -cycleGAN,” *IEEE Transactions on Medical Imaging*, 2021. (cited on Page 32)
- [74] W.-A. Lin, H. Liao, C. Peng, X. Sun, J. Zhang, J. Luo, R. Chellappa, and S. K. Zhou, “Dudonet: Dual domain network for ct metal artifact reduction,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10512–10521, 2019. (cited on Page 32)
- [75] Y. Li, K. Li, C. Zhang, J. Montoya, and G.-H. Chen, “Learning to reconstruct computed tomography images directly from sinogram data under a variety of data acquisition conditions,” *IEEE transactions on medical imaging*, vol. 38, no. 10, pp. 2469–2481, 2019. (cited on Page 32 and 33)
- [76] A. L. Maas, A. Y. Hannun, A. Y. Ng, *et al.*, “Rectifier nonlinearities improve neural network acoustic models,” in *Proc. icml*, vol. 30, p. 3, Citeseer, 2013. (cited on Page 33)
- [77] J. Leuschner, M. Schmidt, P. S. Ganguly, V. Andriashen, S. B. Coban, A. Denker, D. Bauer, A. Hadjifaradji, K. J. Batenburg, P. Maass, *et al.*, “Quantitative comparison of deep learning-based image reconstruction methods for low-dose and sparse-angle ct applications,” *Journal of Imaging*, vol. 7, no. 3, p. 44, 2021. (cited on Page 33, 38, 45, 46, and 73)

- [78] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015. (cited on Page 33)
- [79] J. F. Williamson, B. R. Whiting, J. Benac, R. J. Murphy, G. J. Blaine, J. A. O’Sullivan, D. G. Politte, and D. L. Snyder, “Prospects for quantitative computed tomography imaging in the presence of foreign metal bodies using statistical image reconstruction,” *Medical physics*, vol. 29, no. 10, pp. 2404–2418, 2002. (cited on Page 35)
- [80] G. Poludniowski, G. Landry, F. Deblois, P. Evans, and F. Verhaegen, “Spekcalc: a program to calculate photon spectra from tungsten anode x-ray tubes,” *Physics in Medicine & Biology*, vol. 54, no. 19, p. N433, 2009. (cited on Page 35)
- [81] J. Leuschner, M. Schmidt, D. O. Baguer, and P. Maaß, “The lodopab-ct dataset: A benchmark dataset for low-dose ct reconstruction methods,” *arXiv preprint arXiv:1910.01113*, 2019. (cited on Page 36)
- [82] P. J. La Rivière, J. Bian, and P. A. Vargas, “Penalized-likelihood sinogram restoration for computed tomography,” *IEEE transactions on medical imaging*, vol. 25, no. 8, pp. 1022–1036, 2006. (cited on Page 36)
- [83] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014. (cited on Page 36)
- [84] D. F. Bauer, T. Russ, B. I. Waldkirch, C. Tönnies, W. P. Segars, L. R. Schad, F. G. Zöllner, and A.-K. Golla, “Generation of annotated multimodal ground truth datasets for abdominal medical image registration,” *International journal of computer assisted radiology and surgery*, pp. 1–9, 2021. (cited on Page 45 and 54)
- [85] H. Arabi and H. Zaidi, “Deep learning-based metal artefact reduction in pet/ct imaging,” *European radiology*, pp. 1–13, 2021. (cited on Page 45 and 46)
- [86] R. A. Nasirudin, K. Mei, P. Panchev, A. Fehringer, F. Pfeiffer, E. J. Rummeny, M. Fiebich, and P. B. Noël, “Reduction of metal artifact in single photon-counting computed tomography by spectral-driven iterative reconstruction technique,” *PLoS One*, vol. 10, no. 5, p. e0124831, 2015. (cited on Page 46)
- [87] M. J. Willeminck, M. Persson, A. Pourmorteza, N. J. Pelc, and D. Fleischmann, “Photon-counting ct: technical principles and clinical prospects,” *Radiology*, vol. 289, no. 2, pp. 293–312, 2018. (cited on Page 46)
- [88] M. D. McCarter and Y. Fong, “Metastatic liver tumors,” in *Seminars in surgical oncology*, vol. 19, pp. 177–188, Wiley Online Library, 2000. (cited on Page 47)
- [89] H. Qiu, A. W. Katz, and M. T. Milano, “Oligometastases to the liver: predicting outcomes based upon radiation sensitivity,” *Journal of thoracic disease*, vol. 8, no. 10, p. E1384, 2016. (cited on Page 47)

-
- [90] J. Xu, J. Fan, X. Qin, J. Cai, J. Gu, S. Wang, X. Wang, S. Zhang, and Z. Zhang, “Chinese guidelines for the diagnosis and comprehensive treatment of colorectal liver metastases (version 2018),” *Journal of cancer research and clinical oncology*, vol. 145, no. 3, pp. 725–736, 2019. (cited on Page 47)
- [91] T. Ruers, F. Van Coevorden, C. J. Punt, J.-P. E. Pierie, I. Borel-Rinkes, J. A. Ledermann, G. Poston, W. Bechstein, M.-A. Lentz, M. Mauer, *et al.*, “Local treatment of unresectable colorectal liver metastases: results of a randomized phase ii trial,” *JNCI: Journal of the National Cancer Institute*, vol. 109, no. 9, p. djx015, 2017. (cited on Page 47)
- [92] P. Kron, M. Linecker, R. P. Jones, G. J. Toogood, P.-A. Clavien, and J. Lodge, “Ablation or resection for colorectal liver metastases? a systematic review of the literature,” *Frontiers in oncology*, vol. 9, p. 1052, 2019. (cited on Page 47)
- [93] T. Bartolotta, A. Taibbi, M. Midiri, L. La Grutta, M. De Maria, and R. La-galla, “Characterisation of focal liver lesions undetermined at grey-scale us: contrast-enhanced us versus 64-row mdct and mri with liver-specific contrast agent,” *La radiologia medica*, vol. 115, no. 5, pp. 714–731, 2010. (cited on Page 47)
- [94] A. El Kaffas, R. M. S. Sigrist, G. Fisher, S. Bachawal, J. Liau, H. Wang, A. Karanany, I. Durot, J. Rosenberg, D. Hristov, *et al.*, “Quantitative three-dimensional dynamic contrast-enhanced ultrasound imaging: first-in-human pilot study in patients with liver metastases,” *Theranostics*, vol. 7, no. 15, p. 3745, 2017. (cited on Page 47)
- [95] M. Cathomas, N. Mertineit, C. Kim-Fuchs, A. Lachenmayer, and M. Maurer, “Value of mri/ct image fusion for targeting “invisible” lesions in stereotactic microwave ablation (smwa) of malignant liver lesions: A retrospective analysis,” *Cardiovascular and interventional radiology*, vol. 43, no. 10, pp. 1505–1514, 2020. (cited on Page 47, 48, and 49)
- [96] R. R. Weichselbaum and S. Hellman, “Oligometastases revisited,” *Nature reviews Clinical oncology*, vol. 8, no. 6, p. 378, 2011. (cited on Page 48 and 49)
- [97] S. P. Pitroda, N. N. Khodarev, L. Huang, A. Uppal, S. C. Wightman, S. Ganai, N. Joseph, J. Pitt, M. Brown, M. Forde, *et al.*, “Integrated molecular subtyping defines a curable oligometastatic state in colorectal liver metastasis,” *Nature communications*, vol. 9, no. 1, pp. 1–8, 2018. (cited on Page 49)
- [98] D. R. Gomez, G. R. Blumenschein Jr, J. J. Lee, M. Hernandez, R. Ye, D. R. Camidge, R. C. Doebele, F. Skoulidis, L. E. Gaspar, D. L. Gibbons, *et al.*, “Local consolidative therapy versus maintenance therapy or observation for patients with oligometastatic non-small-cell lung cancer without progression after first-line systemic therapy: a multicentre, randomised, controlled, phase 2 study,” *The lancet oncology*, vol. 17, no. 12, pp. 1672–1682, 2016. (cited on Page 48)

- [99] A. Valladares, T. Beyer, and I. Rausch, “Physical imaging phantoms for simulation of tumor heterogeneity in pet, ct, and mri: An overview of existing designs,” *Medical physics*, vol. 47, no. 4, p. 2023, 2020. (cited on Page 48 and 61)
- [100] M. K. Chmarra, R. Hansen, R. Mårvik, and T. Langø, “Multimodal phantom of liver tissue,” *PloS one*, vol. 8, no. 5, p. e64180, 2013. (cited on Page 48)
- [101] M. S. Ahmad, N. Suardi, A. Shukri, N. N. A. N. Ab Razak, A. A. Oglat, O. Makhmrah, and H. Mohammad, “Dynamic hepatocellular carcinoma model within a liver phantom for multimodality imaging,” *European journal of radiology open*, vol. 7, p. 100257, 2020. (cited on Page 48)
- [102] Y. Liao, L. Wang, X. Xu, H. Chen, J. Chen, G. Zhang, H. Lei, R. Wang, S. Zhang, X. Gu, X. Zhen, and L. Zhou, “An anthropomorphic abdominal phantom for deformable image registration accuracy validation in adaptive radiation therapy,” *Medical Physics*, vol. 44, no. 6, pp. 2369–2378, 2017. (cited on Page 48)
- [103] S. Ehrbar, A. Jöhl, M. Kühni, M. Meboldt, E. Ozkan Elsen, C. Tanner, O. Goksel, S. Klöck, J. Unkelbach, M. Guckenberger, *et al.*, “Elpha: Dynamically deformable liver phantom for real-time motion-adaptive radiotherapy treatments,” *Medical physics*, vol. 46, no. 2, pp. 839–850, 2019. (cited on Page 48 and 58)
- [104] P. Jahnke, F. B. Schwarz, M. Ziegert, T. Almasi, O. Abdelhadi, M. Nunninger, B. Hamm, and M. Scheel, “A radiopaque 3d printed, anthropomorphic phantom for simulation of ct-guided procedures,” *European radiology*, vol. 28, no. 11, pp. 4818–4823, 2018. (cited on Page 48)
- [105] J.-E. Kim, S. H. Kim, S. J. Lee, and H. Rhim, “Hypervascular hepatocellular carcinoma 1 cm or smaller in patients with chronic liver disease: characterization with gadoteric acid-enhanced mri that includes diffusion-weighted imaging,” *American Journal of Roentgenology*, vol. 196, no. 6, pp. W758–W765, 2011. (cited on Page 49)
- [106] H. Chandarana and B. Taouli, “Diffusion and perfusion imaging of the liver,” *European journal of radiology*, vol. 76, no. 3, pp. 348–358, 2010. (cited on Page 49)
- [107] M. R. Oliva and S. Saini, “Liver cancer imaging: role of ct, mri, us and pet,” *Cancer imaging*, vol. 4, no. Spec No A, p. S42, 2004. (cited on Page 49)
- [108] A.-K. Schnurr, L. R. Schad, and F. G. Zöllner, “Sparsely connected convolutional layers in cnns for liver segmentation in ct,” in *Bildverarbeitung für die Medizin 2019*, pp. 80–85, Springer, 2019. (cited on Page 49 and 54)
- [109] B. Vogelstein, N. Papadopoulos, V. E. Velculescu, S. Zhou, L. A. Diaz, and K. W. Kinzler, “Cancer genome landscapes,” *science*, vol. 339, no. 6127, pp. 1546–1558, 2013. (cited on Page 49)

-
- [110] J. M. Franklin, R. A. Sharma, A. L. Harris, and F. V. Gleeson, “Imaging oligometastatic cancer before local treatment,” *The Lancet Oncology*, vol. 17, no. 9, pp. e406–e414, 2016. (cited on Page 49)
- [111] N. M. deSouza, Y. Liu, A. Chiti, D. Oprea-Lager, G. Gebhart, B. E. Van Beers, K. Herrmann, and F. E. Lecouvet, “Strategies and technical challenges for imaging oligometastatic disease: Recommendations from the european organisation for research and treatment of cancer imaging group,” *European journal of cancer*, vol. 91, pp. 153–163, 2018. (cited on Page 49)
- [112] A. Smakic, N. Rathmann, M. Kostrzewa, S. O. Schönberg, C. Weiß, and S. J. Diehl, “Performance of a robotic assistance device in computed tomography-guided percutaneous diagnostic and therapeutic procedures,” *Cardiovascular and interventional radiology*, vol. 41, no. 4, pp. 639–644, 2018. (cited on Page 49 and 61)
- [113] N. Rotolo, C. Floridi, A. Imperatori, F. Fontana, A. M. Ierardi, M. Mangini, V. Arlant, G. De Marchi, R. Novario, L. Dominioni, *et al.*, “Comparison of cone-beam ct-guided and ct fluoroscopy-guided transthoracic needle biopsy of lung nodules,” *European radiology*, vol. 26, no. 2, pp. 381–389, 2016. (cited on Page 49)
- [114] A. Kamaya, K. E. Maturen, G. A. Tye, Y. I. Liu, N. N. Parti, and T. S. Desser, “Hypervascular liver lesions,” in *Seminars in Ultrasound, CT and MRI*, vol. 30, pp. 387–407, Elsevier, 2009. (cited on Page 53)
- [115] R. Beare, B. Lowekamp, and Z. Yaniv, “Image segmentation, registration and characterization in r with simpleitk,” *Journal of statistical software*, vol. 86, 2018. (cited on Page 54)
- [116] D. F. Pepley, C. C. Sonntag, R. S. Prabhu, M. A. Yovanoff, D. C. Han, S. R. Miller, and J. Z. Moore, “Building ultrasound phantoms with modified polyvinyl chloride: a comparison of needle insertion forces and sonographic appearance with commercial and traditional simulation materials,” *Simulation in healthcare: journal of the Society for Simulation in Healthcare*, vol. 13, no. 3, p. 149, 2018. (cited on Page 57 and 67)
- [117] Y. C. Yoon, J. S. Lee, S. U. Park, J. H. Kwon, T. H. Hong, and D. G. Kim, “Quantitative assessment of liver fibrosis using shore durometer,” *Annals of surgical treatment and research*, vol. 93, no. 6, pp. 300–304, 2017. (cited on Page 58)
- [118] M. J. Paszek, N. Zahir, K. R. Johnson, J. N. Lakins, G. I. Rozenberg, A. Gefen, C. A. Reinhart-King, S. S. Margulies, M. Dembo, D. Boettiger, *et al.*, “Tensional homeostasis and the malignant phenotype,” *Cancer cell*, vol. 8, no. 3, pp. 241–254, 2005. (cited on Page 58)
- [119] A. Hakime, S. Yevich, L. Tselikas, F. Deschamps, D. Petrover, and T. De Baere, “Percutaneous thermal ablation with ultrasound guidance. fusion imaging guidance to improve conspicuity of liver metastasis,” *Cardiovas-*

- cular and interventional radiology*, vol. 40, no. 5, pp. 721–727, 2017. (cited on Page 59)
- [120] L. Zimmermann, M. Buschmann, H. Herrmann, G. Heilemann, P. Kuess, G. Goldner, T. Nyholm, D. Georg, and N. Nesvacil, “An mr-only acquisition and artificial intelligence based image-processing protocol for photon and proton therapy using a low field mr,” *Zeitschrift für Medizinische Physik*, vol. 31, no. 1, pp. 78–88, 2021. (cited on Page 59 and 68)
- [121] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, “Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” *CA: a cancer journal for clinicians*, vol. 71, no. 3, pp. 209–249, 2021. (cited on Page 61)
- [122] L. Dickinson, H. U. Ahmed, C. Allen, J. O. Barentsz, B. Carey, J. J. Futterer, S. W. Heijmink, P. J. Hoskin, A. Kirkham, A. R. Padhani, *et al.*, “Magnetic resonance imaging for the detection, localisation, and characterisation of prostate cancer: recommendations from a european consensus meeting,” *European urology*, vol. 59, no. 4, pp. 477–494, 2011. (cited on Page 61)
- [123] A. B. Rosenkrantz and S. S. Taneja, “Prostate mri can reduce overdiagnosis and overtreatment of prostate cancer,” *Academic radiology*, vol. 22, no. 8, pp. 1000–1006, 2015. (cited on Page)
- [124] A. Stabile, F. Giganti, A. B. Rosenkrantz, S. S. Taneja, G. Villeirs, I. S. Gill, C. Allen, M. Emberton, C. M. Moore, and V. Kasivisvanathan, “Multiparametric mri for prostate cancer diagnosis: current status and future directions,” *Nature reviews urology*, vol. 17, no. 1, pp. 41–61, 2020. (cited on Page 61 and 67)
- [125] J. D. Le, N. Tan, E. Shkolyar, D. Y. Lu, L. Kwan, L. S. Marks, J. Huang, D. J. Margolis, S. S. Raman, and R. E. Reiter, “Multifocality and prostate cancer detection by multiparametric magnetic resonance imaging: correlation with whole-mount histopathology,” *European urology*, vol. 67, no. 3, pp. 569–576, 2015. (cited on Page 61)
- [126] J. P. Radtke, C. Schwab, M. B. Wolf, M. T. Freitag, C. D. Alt, C. Kesch, I. V. Popeneciu, C. Huettenbrink, C. Gasch, T. Klein, *et al.*, “Multiparametric magnetic resonance imaging (mri) and mri–transrectal ultrasound fusion biopsy for index tumor detection: correlation with radical prostatectomy specimen,” *European urology*, vol. 70, no. 5, pp. 846–853, 2016. (cited on Page 61)
- [127] M. Noureldin, D. Eldred-Evans, C. Khoo, M. Winkler, H. Sokhi, H. Tam, and H. Ahmed, “Mri-targeted biopsies for prostate cancer diagnosis and management,” *World journal of urology*, pp. 1–7, 2020. (cited on Page 61)
- [128] J. Wang, K. Tanderup, A. Cunha, A. L. Damato, N. C. Gil’ad, R. J. Kudchadker, and F. Mourtada, “Magnetic resonance imaging basics for the prostate brachytherapist,” *Brachytherapy*, vol. 16, no. 4, pp. 715–727, 2017. (cited on Page 61)

-
- [129] B. R. Matlaga, L. A. Eskew, and D. L. McCullough, “Prostate biopsy: indications and technique,” *The Journal of urology*, vol. 169, no. 1, pp. 12–19, 2003. (cited on Page 61)
- [130] M. van der Leest, E. Cornel, B. Israël, R. Hendriks, A. R. Padhani, M. Hoogenboom, P. Zamecnik, D. Bakker, A. Y. Setiasti, J. Veltman, *et al.*, “Head-to-head comparison of transrectal ultrasound-guided prostate biopsy versus multiparametric prostate resonance imaging with subsequent magnetic resonance-guided biopsy in biopsy-naïve men with elevated prostate-specific antigen: a large prospective multicenter clinical study,” *European urology*, vol. 75, no. 4, pp. 570–578, 2019. (cited on Page 61)
- [131] M. Connor, M. Gorin, H. Ahmed, and R. Nigam, “Focal therapy for localized prostate cancer in the era of routine multi-parametric mri,” *Prostate cancer and prostatic diseases*, vol. 23, no. 2, pp. 232–243, 2020. (cited on Page 61)
- [132] D. A. Woodrum, A. Kawashima, R. J. Karnes, B. J. Davis, I. Frank, D. E. Engen, K. R. Gorny, J. P. Felmlee, M. R. Callstrom, and L. A. Mynderse, “Magnetic resonance imaging-guided cryoablation of recurrent prostate cancer after radical prostatectomy: initial single institution experience,” *Urology*, vol. 82, no. 4, pp. 870–875, 2013. (cited on Page 61)
- [133] S. E. Eggener, A. Yousuf, S. Watson, S. Wang, and A. Oto, “Phase ii evaluation of magnetic resonance imaging guided focal laser ablation of prostate cancer,” *The Journal of urology*, vol. 196, no. 6, pp. 1670–1675, 2016. (cited on Page 61)
- [134] J. L. Chin, M. Billia, J. Relle, M. C. Roethke, I. V. Popeneciu, T. H. Kuru, G. Hatiboglu, M. B. Mueller-Wolf, J. Motsch, C. Romagnoli, *et al.*, “Magnetic resonance imaging-guided transurethral ultrasound ablation of prostate tissue in patients with localized prostate cancer: a prospective phase 1 clinical trial,” *European urology*, vol. 70, no. 3, pp. 447–455, 2016. (cited on Page 61)
- [135] C. G. Chaussy and S. Thüroff, “High-intensity focused ultrasound for the treatment of prostate cancer: a review,” *Journal of endourology*, vol. 31, no. S1, pp. S–30, 2017. (cited on Page 61)
- [136] W. van den Bos, M. J. Scheltema, A. R. Siriwardana, A. M. Kalsbeek, J. E. Thompson, F. Ting, M. Böhm, A.-M. Haynes, R. Shnier, W. Delprado, *et al.*, “Focal irreversible electroporation as primary treatment for localized prostate cancer,” *BJU international*, vol. 121, no. 5, pp. 716–724, 2018. (cited on Page 61)
- [137] J. Bomers, D. Bosboom, G. Tigelaar, J. Sabisch, J. Fütterer, and D. Yakar, “Feasibility of a 2nd generation mr-compatible manipulator for transrectal prostate biopsy guidance,” *European radiology*, vol. 27, no. 4, pp. 1776–1782, 2017. (cited on Page 61, 65, and 68)
- [138] M. G. Schouten, J. Ansems, W. K. J. Renema, D. Bosboom, T. W. Scheenen, and J. J. Fütterer, “The accuracy and safety aspects of a novel robotic needle

- guide manipulator to perform transrectal prostate biopsies,” *Medical physics*, vol. 37, no. 9, pp. 4744–4750, 2010. (cited on Page 61)
- [139] T. Chiu, Z. Xiong, D. Parsons, M. R. Folkert, P. M. Medin, and B. Hrycushko, “Low-cost 3d print–based phantom fabrication to facilitate interstitial prostate brachytherapy training program,” *Brachytherapy*, vol. 19, no. 6, pp. 800–811, 2020. (cited on Page 61)
- [140] N. Hungr, J.-A. Long, V. Beix, and J. Troccaz, “A realistic deformable prostate phantom for multimodal imaging and needle-insertion procedures,” *Medical physics*, vol. 39, no. 4, pp. 2031–2041, 2012. (cited on Page)
- [141] U. Lindner, N. Lawrentschuk, R. A. Weersink, O. Raz, E. Hlasny, M. S. Sussman, S. R. Davidson, M. R. Gertner, and J. Trachtenberg, “Construction and evaluation of an anatomically correct multi-image modality compatible phantom for prostate cancer focal ablation,” *The Journal of urology*, vol. 184, no. 1, pp. 352–357, 2010. (cited on Page)
- [142] J.-A. Long, V. Daanen, A. Moreau-Gaudry, J. Troccaz, J.-J. Rambeaud, and J.-L. Descotes, “Prostate biopsies guided by three-dimensional real-time (4-d) transrectal ultrasonography on a phantom: comparative study versus two-dimensional transrectal ultrasound-guided biopsies,” *European urology*, vol. 52, no. 4, pp. 1097–1105, 2007. (cited on Page)
- [143] W. Neumann, A. Bichert, J. Fleischhauer, A. Stern, R. Figuli, M. Wilhelm, L. R. Schad, and F. G. Zöllner, “A novel 3d printed mechanical actuator using centrifugal force for magnetic resonance elastography: Initial results in an anthropomorphic prostate phantom,” *PloS one*, vol. 13, no. 10, p. e0205442, 2018. (cited on Page 61)
- [144] J. M. Cunningham, E. A. Barberi, J. Miller, J. P. Kim, and C. K. Glide-Hurst, “Development and evaluation of a novel mr-compatible pelvic end-to-end phantom,” *Journal of applied clinical medical physics*, vol. 20, no. 1, pp. 265–275, 2019. (cited on Page 61)
- [145] N. Niebuhr, W. Johnen, G. Echner, A. Runz, M. Bach, M. Stoll, K. Giske, S. Greilich, and A. Pfaffenberger, “The adam-pelvis phantom—an anthropomorphic, deformable and multimodal phantom for mrgt,” *Physics in Medicine & Biology*, vol. 64, no. 4, p. 04NT05, 2019. (cited on Page 61 and 67)
- [146] A. J. Doyle, F. Sullivan, J. Walsh, D. M. King, D. Cody, and J. E. Browne, “Development and preliminary evaluation of an anthropomorphic trans-rectal ultrasound prostate brachytherapy training phantom,” *Ultrasound in Medicine & Biology*, vol. 47, no. 3, pp. 833–846, 2021. (cited on Page 62)
- [147] D. F. Bauer, E. Oelschlegel, A.-K. Golla, A. Adlung, T. Russ, I. Hermann, I. Brumer, J. Rosenkranz, F. Tollens, S. Clausen, P. Aumüller, L. R. Schad, D. Nörenberg, and F. G. Zöllner, “An anthropomorphic pelvis phantom for prostate brachytherapy and biopsy,” *Proc. Intl. Soc. Mag. Reson. Med.* 29, 2021. (cited on Page 62)

-
- [148] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. Tsui, “4d xcat phantom for multimodality imaging research,” *Medical physics*, vol. 37, no. 9, pp. 4902–4915, 2010. (cited on Page 62)
- [149] M. F. Leitzmann and S. Rohrmann, “Risk factors for the onset of prostatic cancer: age, location, and behavioral correlates,” *Clinical epidemiology*, vol. 4, p. 1, 2012. (cited on Page 62)
- [150] S.-J. Zhang, H.-N. Qian, Y. Zhao, K. Sun, H.-Q. Wang, G.-Q. Liang, F.-H. Li, and Z. Li, “Relationship between age and prostate size,” *Asian journal of andrology*, vol. 15, no. 1, p. 116, 2013. (cited on Page 62)
- [151] P. Jezzard and R. S. Balaban, “Correction for geometric distortion in echo planar images from b0 field variations,” *Magnetic resonance in medicine*, vol. 34, no. 1, pp. 65–73, 1995. (cited on Page 64)
- [152] A. M. Weidner, H. J. Michaely, A. Lemke, L. Breiting, F. Wenz, A. Marx, S. O. Schoenberg, and D. J. Dinter, “Value of multiparametric prostate mri of the peripheral zone,” *Zeitschrift für Medizinische Physik*, vol. 21, no. 3, pp. 198–205, 2011. (cited on Page 64 and 68)
- [153] A. D. Baur, C. M. Hansen, J. Rogasch, H. Posch, S. Elezkurta, A. Maxeiner, K. Erb-Eigner, and M. R. Makowski, “Evaluation of t1 relaxation time in prostate cancer and benign prostate tissue using a modified look-locker inversion recovery sequence,” *Scientific reports*, vol. 10, no. 1, pp. 1–8, 2020. (cited on Page 66)
- [154] A. H. Dinh, R. Souchon, C. Melodelima, F. Bratan, F. Mège-Lechevallier, M. Colombel, and O. Rouvière, “Characterization of prostate cancer using t2 mapping at 3 t: a multi-scanner study,” *Diagnostic and interventional imaging*, vol. 96, no. 4, pp. 365–372, 2015. (cited on Page 66)
- [155] W. D. D’Souza, E. L. Madsen, O. Unal, K. K. Vigen, G. R. Frank, and B. R. Thomadsen, “Tissue mimicking materials for a multi-imaging modality prostate phantom,” *Medical physics*, vol. 28, no. 4, pp. 688–700, 2001. (cited on Page 66)
- [156] A. Batur, U. Kerimoglu, and H. Ataseven, “Hounsfield unit density in the characterisation of bile duct lesions,” *Polish journal of radiology*, vol. 84, p. e397, 2019. (cited on Page 66)
- [157] M. Zhang, P. Nigwekar, B. Castaneda, K. Hoyt, J. V. Joseph, A. di Sant’Agnese, E. M. Messing, J. G. Strang, D. J. Rubens, and K. J. Parker, “Quantitative characterization of viscoelastic properties of human prostate correlated with histology,” *Ultrasound in medicine & biology*, vol. 34, no. 7, pp. 1033–1042, 2008. (cited on Page 67)
- [158] W. Neumann, F. Lietzmann, L. R. Schad, and F. G. Zöllner, “Design of a multimodal (1h/23na mr/ct) anthropomorphic thorax phantom,” *Zeitschrift für Medizinische Physik*, vol. 27, no. 2, pp. 124–131, 2017. (cited on Page 67)

- [159] K. Hattori, Y. Ikemoto, W. Takao, S. Ohno, T. Harimoto, S. Kanazawa, M. Oita, K. Shibuya, M. Kuroda, and H. Kato, “Development of mri phantom equivalent to human tissues for 3.0-t mri,” *Medical physics*, vol. 40, no. 3, p. 032303, 2013. (cited on Page 68)
- [160] I. Lavdas, K. C. Behan, A. Papadaki, D. W. McRobbie, and E. O. Aboagye, “A phantom for diffusion-weighted mri (dw-mri),” *Journal of Magnetic Resonance Imaging*, vol. 38, no. 1, pp. 173–179, 2013. (cited on Page 68)
- [161] X. Yi, E. Walia, and P. Babyn, “Generative adversarial network in medical imaging: A review,” *Medical image analysis*, vol. 58, p. 101552, 2019. (cited on Page 73)
- [162] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, “Analyzing and improving the image quality of stylegan,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8110–8119, 2020. (cited on Page 73)
- [163] L. Fetty, M. Bylund, P. Kuess, G. Heilemann, T. Nyholm, D. Georg, and T. Löfstedt, “Latent space manipulation for high-resolution medical image synthesis via the stylegan,” *Zeitschrift für Medizinische Physik*, vol. 30, no. 4, pp. 305–314, 2020. (cited on Page 73)
- [164] D. F. Bauer, C. Ulrich, A.-K. Golla, T. Russ, C. Tönnies, J. Leuschner, M. Schmidt, L. R. Schad, and F. G. Zöllner, “Image reconstruction using end-to-end deep learning for low-dose ct,” in *CARS 2021: Computer Assisted Radiology and Surgery Proceedings of the 35th International Congress and Exhibition*, 2021. (cited on Page 73)
- [165] J. Dinkel, C. Thieke, C. Plathow, P. Zamecnik, H. Prüm, P. E. Huber, H.-U. Kauczor, H.-P. Schlemmer, and C. M. Zechmann, “Respiratory-induced prostate motion,” *Strahlentherapie und Onkologie*, vol. 187, no. 7, pp. 426–432, 2011. (cited on Page 74)

10. Appendix

10.1 Supplementary Material Chapter 3

10.1.1 ASSD and DSC Registration Evaluation Metrics

To evaluate the proof of principle registration, we calculated the Dice similarity coefficient (DSC) in the manuscript:

$$DSC(X, Y) = \frac{2 |X \cap Y|}{|X| + |Y|}, \quad (10.1)$$

where X are the registered liver masks and Y are the ground truth CT liver masks. In this document we additionally provide the average symmetric surface distance (ASSD):

$$ASSD(X, Y) = \frac{\sum_{x \in X} \min d(x, Y) + \sum_{y \in Y} \min d(y, X)}{|X| + |Y|}, \quad (10.2)$$

In the manuscript we presented boxplots of the DSC shown in [Figure 3.4](#). Here, we additionally show boxplots for the ASSD ([Figure 10.1](#)) and mean values for the DSC ([Table 10.1](#)) and ASSD ([Table 10.2](#)). The distributions in both boxplots are very similar and the outliers (failed registrations) stem from the same cases. The mean values shown in the tables show the same tendencies as the boxplots for ASSD and DSC. We therefore argue, that both metrics lead to the same conclusions regarding the quality of the registrations. Thus, we decided that only presenting the boxplot of the DSC in the manuscript is sufficient and avoids redundancy.

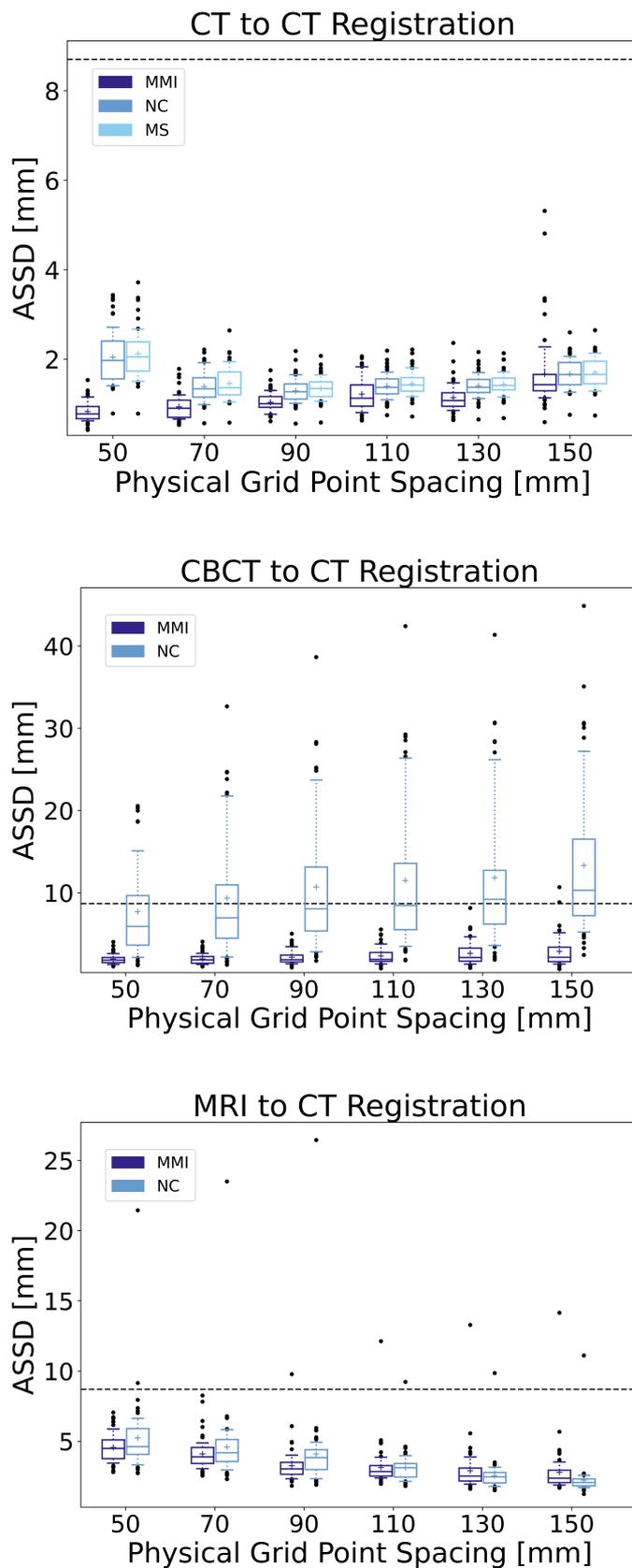


Figure 10.1: ASSD for the proof of principle registrations with 56 data points each. The mean is marked as a "+" and the whiskers indicate the 10th and 90th percentile. All outliers are depicted as black dots. The dashed horizontal line shows the mean pre-registration DSC.

Table 10.1: Registration evaluation DSC metric. The best result for each modality is highlighted. Mean pre-registration DSC: 0.667 ± 0.035 .

DSC (CT to CT)						
Grid Point Spacing [mm]	50	70	90	110	130	150
MMI	0.960 ± 0.010	0.956 ± 0.013	0.952 ± 0.010	0.945 ± 0.015	0.948 ± 0.013	0.926 ± 0.031
NC	0.913 ± 0.020	0.938 ± 0.013	0.941 ± 0.010	0.937 ± 0.011	0.936 ± 0.012	0.925 ± 0.015
MS	0.910 ± 0.019	0.935 ± 0.013	0.939 ± 0.010	0.935 ± 0.011	0.935 ± 0.012	0.924 ± 0.015

DSC (CBCT to CT)						
Grid Point Spacing [mm]	50	70	90	110	130	150
MMI	0.915 ± 0.019	0.915 ± 0.020	0.907 ± 0.028	0.903 ± 0.034	0.893 ± 0.045	0.886 ± 0.058
NC	0.784 ± 0.116	0.752 ± 0.134	0.721 ± 0.143	0.700 ± 0.148	0.697 ± 0.149	0.662 ± 0.139

DSC (MRI to CT)						
Grid Point Spacing [mm]	50	70	90	110	130	150
MMI	0.811 ± 0.045	0.832 ± 0.048	0.862 ± 0.054	0.867 ± 0.061	0.876 ± 0.068	0.879 ± 0.073
NC	0.791 ± 0.069	0.816 ± 0.070	0.836 ± 0.077	0.868 ± 0.043	0.891 ± 0.042	0.903 ± 0.044

Table 10.2: Registration evaluation ASSD metric. The best result for each modality is highlighted. Mean pre-registration ASSD: 8.7 mm.

		ASSD (CT to CT)					
Grid Point Spacing [mm]		50	70	90	110	130	150
MMI		0.8 ± 0.2	0.9 ± 0.3	1.0 ± 0.2	1.2 ± 0.4	1.1 ± 0.3	1.7 ± 0.8
NC		2.0 ± 0.6	1.4 ± 0.3	1.3 ± 0.3	1.4 ± 0.3	1.4 ± 0.3	1.7 ± 0.3
MS		2.1 ± 0.5	1.5 ± 0.4	1.3 ± 0.3	1.5 ± 0.3	1.4 ± 0.3	1.7 ± 0.3

		ASSD (CBCT to CT)					
Grid Point Spacing [mm]		50	70	90	110	130	150
MMI		1.9 ± 0.6	2.0 ± 0.6	2.2 ± 0.8	2.3 ± 1.1	2.6 ± 1.5	2.9 ± 1.9
NC		7.7 ± 5.5	9.4 ± 7.2	10.7 ± 8.1	11.5 ± 8.7	11.8 ± 8.7	13.3 ± 9.0

		ASSD (MRI to CT)					
Grid Point Spacing [mm]		50	70	90	110	130	150
MMI		4.6 ± 1.0	4.1 ± 1.1	3.3 ± 1.2	3.1 ± 1.4	2.9 ± 1.6	2.8 ± 1.7
NC		5.2 ± 2.6	4.6 ± 2.8	4.1 ± 3.2	3.1 ± 1.1	2.5 ± 1.1	2.2 ± 1.2

10.1.2 Res-Net network architecture

For the generators in the CycleGAN network, we used a Res-Net architecture with an encoding stage, 9 residual blocks and a decoding stage. A detailed schematic of the Res-Net is shown in Figure 10.2. In the encoding stage, the downsampling is performed via strided convolutions. The upsampling in the decoding stage is performed via a bilinear interpolation instead of a deconvolution, in order to avoid checkerboard artifacts. All convolutional layers use the ReLU activation function, except for the final convolution, which employs a hyperbolic tangent (tanh).

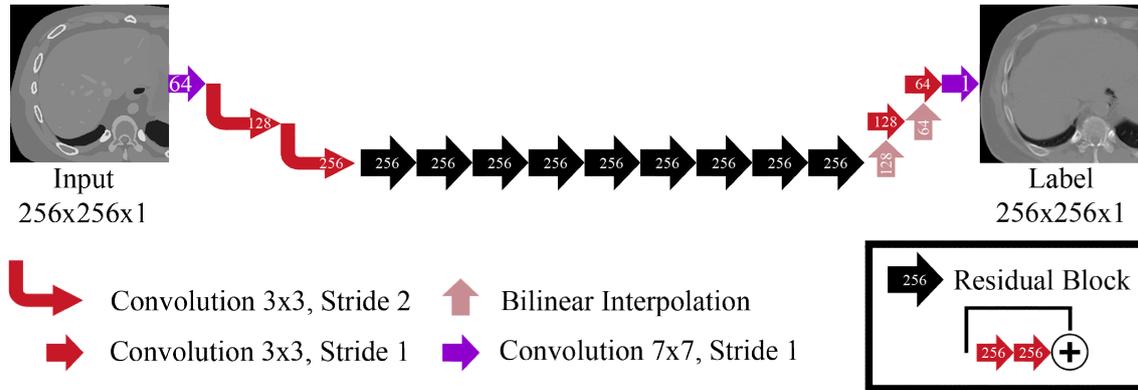


Figure 10.2: Res-Net architecture used for the CycleGAN generators. The numbers inside the arrows indicate the number of output channels of an operation.

10.2 Supplementary Material Chapter 6

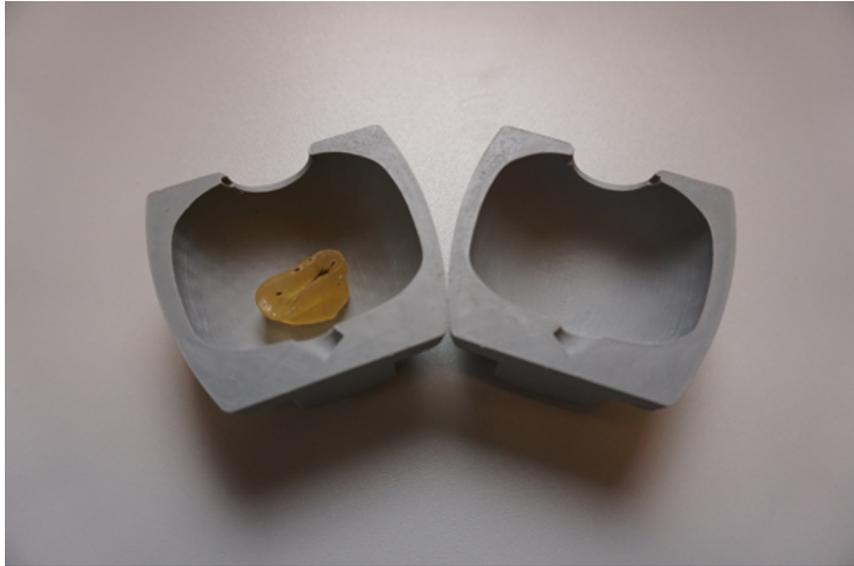


Figure 10.3: Dehydrated and moldy agarose prostate inside the prostate mold.



Figure 10.4: Unedited ultrasound image of the prostate without annotations.

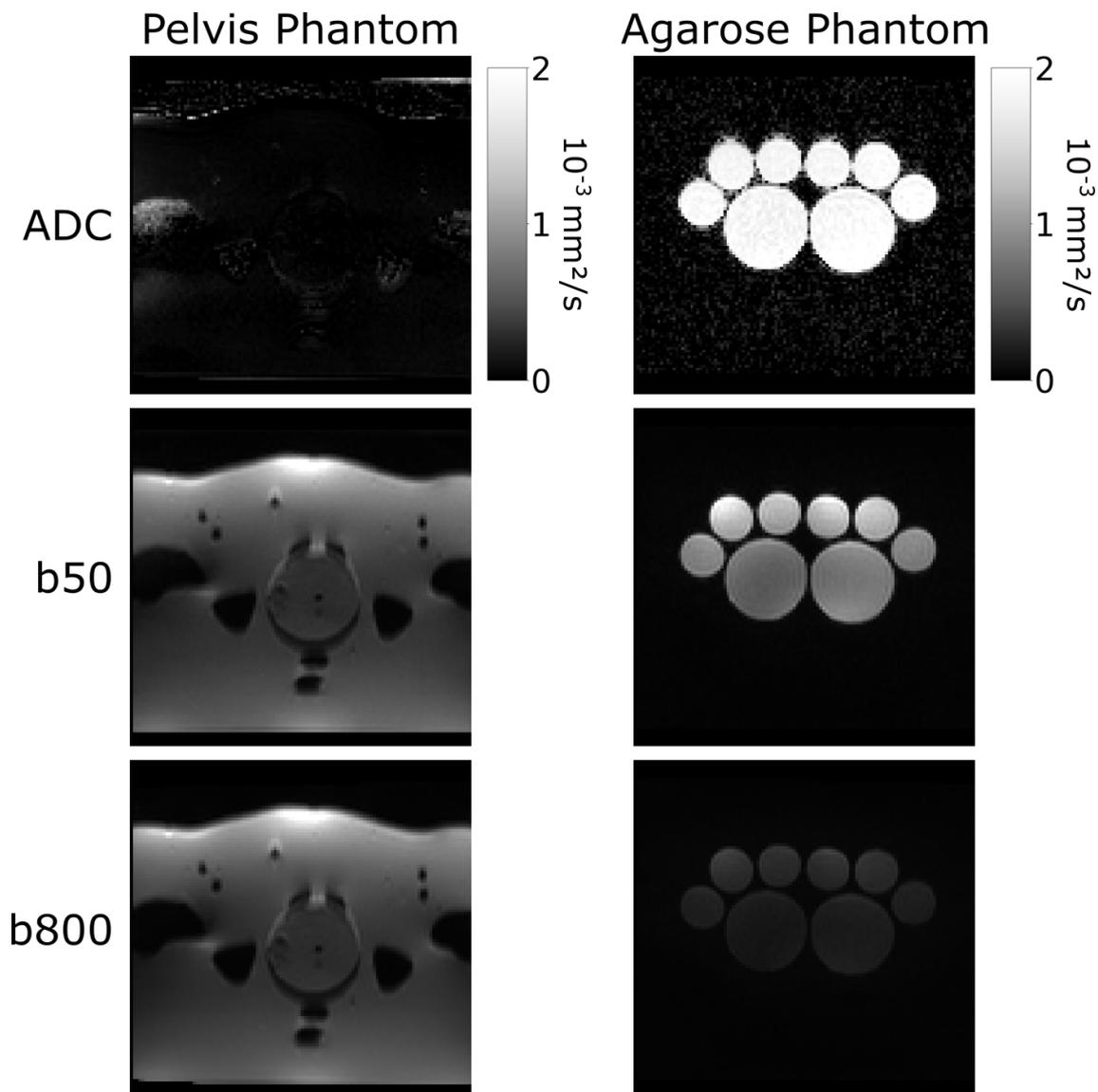


Figure 10.5: Diffusion-weighted imaging (DWI) of the pelvis phantom (left) and reference agarose phantom (right). The ADC-maps in the top row show that there is very little diffusion in our pelvis phantom ($ADC \approx 0.05 \times 10^{-3} \text{ mm}^2/\text{s}$) compared to the agarose phantom ($ADC \approx 2 \times 10^{-3} \text{ mm}^2/\text{s}$). Typical ADC values for the prostate range between $1 \times 10^{-3} \text{ mm}^2/\text{s}$ and $2 \times 10^{-3} \text{ mm}^2/\text{s}$. The b50 and b800 scans of the pelvis phantom confirm this observation, as they hardly differ from each other.

11. Publications

Journal Papers

- **D. F. Bauer**, J. Rosenkranz, A.-K. Golla, C. Tönnies, I. Hermann, T. Russ, G. Kabelitz, A. J. Rothfuss, L. R. Schad, J. L. Stallkamp and F. G. Zöllner. Development of an abdominal phantom for the validation of an oligometastatic disease diagnosis workflow. Submitted to Medical Physics, 19.05.2021.
- **D. F. Bauer**, C. Ulrich, T. Russ, A.-K. Golla, L. R. Schad and F. G. Zöllner. End-to-End Deep Learning CT Image Reconstruction for Metal Artifact Reduction. Applied Sciences, 12(1), 404, doi.org/10.3390/app12010404, 2022.
- **D. F. Bauer**, A. Adlung, I. Brumer, A.-K. Golla, T. Russ, E. Oelschlegel, F. Tollens, S. Clausen, P. Aumüller, L. R. Schad, D. Nörenberg and F. G. Zöllner. An anthropomorphic pelvis phantom for MR-guided prostate interventions. Magnetic Resonance in Medicine, 87(3), pp.1605-1612, doi.org/10.1002/mrm.29043, 2021.
- A.-K. Golla, C. Tönnies, T. Russ, **D. F. Bauer**, M. F. Frölich, S. J. Diehl, S. O. Schönberg, M. Keese, L. R. Schad, F. G. Zöllner and J. S. Rink. Automated Screening for Abdominal Aortic Aneurysm in CT Scans under Clinical Conditions Using Deep Learning. Diagnostics, 11(11), 2131, doi.org/10.3390/diagnostics11112131, 2021
- F. Tollens, N. Westhoff, J. von Hardenberg, S. Clausen, M. Ehmman, F. G. Zöllner, A. Adlung, **D. F. Bauer**, S. O. Schönberg and D. Nörenberg. MRT-gestützte minimal-invasive Therapie des Prostatakarzinoms. Der Radiologe, 61, pp.829–838, doi.org/10.1007/s00117-021-00883-7, 2021.
- **D. F. Bauer**, T. Russ, B. I. Waldkirch, C. Tönnies, W. P. Segars, L. R. Schad, F. G. Zöllner and A.-K. Golla. Generation of annotated multimodal ground truth datasets for abdominal medical image registration. International Journal of Computer Assisted Radiology and Surgery, 16, pp.1277–1285, doi: 10.1007/s11548-021-02372-7, 2021.
- A. Adlung, N. K. Paschke, A.-K. Golla, **D. F. Bauer**, S. A. Mohamed, M. Samartzi, M. Fatar, E. Neumaier-Probst, F. G. Zöllner and L. R. Schad. ^{23}Na MRI in ischemic stroke: Acquisition time reduction using postprocessing with convolutional neural networks. NMR in Biomedicine, 34(4), e4474, doi: 10.1002/nbm.4474, 2021.
- A.-K. Golla, **D. F. Bauer**, R. Schmidt, T. Russ, D. Nörenberg, K. Chung, C. Tönnies, L. R. Schad, and F. G. Zöllner. Convolutional Neural Network Ensemble Segmentation with Ratio-based Sampling for the Arteries and Veins in Abdominal CT Scans. IEEE Transactions on Biomedical Engineering, 68(5), pp.1518-1526, doi: 10.1109/TBME.2020.3042640, 2021.
- J. Leuschner, M. Schmidt, P. S. Ganguly, V. Andriiashen, S. B. Coban, A. Denker, **D. F. Bauer**, A. Hadjifaradji, K. J. Batenburg, P. Maass, M. van Eijnatten. Quantitative Comparison of Deep Learning-Based Image Reconstruction Methods for Low-Dose and Sparse-Angle CT Applications. Journal of Imaging, 7(3), 44, doi: 10.3390/jimaging7030044, 2021.
- T. Russ, S. Goerttler, A.-K. Schnurr, **D. F. Bauer**, S. Hatamikia, L. R. Schad, F. G. Zöllner and K. Chung. Synthesis of CT images from digital

body phantoms using CycleGAN. *International Journal of Computer Assisted Radiology and Surgery*, 14 (10), pp.1741-1750, doi: 10.1007/s11548-019-02042-9, 2019.

Conference Abstracts

- T. Russ, W. Wang, A.-K. Golla, **D. F. Bauer**, M. Tivnan, C. Tönnnes, Y. W. Ma, T. Reynolds, S. Hatamikia, L. R. Schad, F. G. Zöllner, G. J. Gang, and J. W. Stayman, "Fast CBCT reconstruction using convolutional neural networks for arbitrary robotic C-arm orbits, SPIE Medical Imaging, San Diego, USA, 2022.
- B. Grüne, R. Burger, **D. F. Bauer**, A. Schäfer, A. Rothfuss, J. Stallkamp, M. C. Kriegmair, M.-C. Rassweiler-Seyfried. Robotic-assisted versus manual Uro-Dyna CT-guided puncture in an ex-vivo kidney model. 37th Annual European Association of Urology Congress, Amsterdam, The Netherlands, 2022.
- I. Brumer, **D. F. Bauer**, L. R. Schad and F. G. Zöllner. Synthetic Kidney ASL Data for Evaluation & Comparison of Processing Pipelines. ISMRM Workshop on Perfusion MRI, Los Angeles, United States, 2022.
- I. Brumer, **D. F. Bauer**, L. R. Schad and F. G. Zöllner. Synthetic Arterial Spin Labelling MR Images. ISMRM Workshop on Kidney MRI Biomarkers: The Route to Clinical Adoption, Lisbon, Portugal, 2021.
- **D. F. Bauer**, I. Brumer, A. Adlung, A.-K. Golla, T. Russ, E. Oelschlegel, F. Tollens, S. Clausen, P. Aumüller, L. R. Schad, D. Nörenberg and F. G. Zöllner. An Anthropomorphic Pelvis Phantom for MR-guided Prostate Biopsy. Proc. 5th Conference on Image-Guided Interventions, Magdeburg, Germany, 2021.
- **D. F. Bauer**, T. Russ, C. Ulrich, A.-K. Golla, L. R. Schad and F. G. Zöllner. Metal Artifact Correction in End-to-end Deep Learning CT Image Reconstruction. Proc. 5th Conference on Image-Guided Interventions, Magdeburg, Germany, 2021.
- **D. F. Bauer**, C. Ulrich, A.-K. Golla, T. Russ, C. Tönnnes, J. Leuschner, M. Schmidt, L. R. Schad and F. G. Zöllner. Image reconstruction using end-to-end deep learning for low-dose CT. Proc. Computer Assisted Radiology and Surgery, Munich, Germany, 2021.
- T. Russ, W. Wang, A.-K. Golla, **D. F. Bauer**, M. Tivnan, C. Tönnnes, Y. W. Ma, T. Reynolds, S. Hatamikia, L. R. Schad, F. G. Zöllner, G. J. Gang, J. W. Stayman. Fast Reconstruction of non-circular CBCT orbits using CNNs. Proc. Fully3D Congress, Online, 2021.
- **D. F. Bauer**, E. Oelschlegel, A.-K. Golla, A. Adlung, T. Russ, I. Hermann, I. Brumer, J. Rosenkranz, F. Tollens, S. Clausen, P. Aumüller, L. R. Schad, D. Nörenberg and F. G. Zöllner. An Anthropomorphic Pelvis Phantom for Prostate Brachytherapy and Biopsy. Proc. ISMRM Congress, Online, 2021.
- A. Adlung, N. Paschke, A.-K. Golla, **D. F. Bauer**, S. Mohamed, M. Samartzi, M. Fatar, E. Neumaier Probst, F. G. Zöllner, L. R. Schad. Accuracy of quantified ^{23}Na MRI in ischemic stroke with varying undersampling factors and CNN postprocessing. Proc. ISMRM Annual Meeting, Online, 2021.
- **D. F. Bauer**, T. Russ, W. P. Segars, A.-K. Schnurr, L. R. Schad and F. G. Zöllner. Generation of Fully Annotated Abdominal T1-weighted MR Ground

-
- Truth Data for Image Segmentation and Registration using CycleGANs and the XCAT Phantom. Proc. ESMRMB Congress, Online, 37, pp.26-27, 2020.
- T. Russ, A. M. Abdelrehim, **D. F. Bauer**, S. Hatamikia, L. R. Schad, F. G. Zöllner and K. Chung. CBCT image quality and dose simulations for arbitrary source-detector trajectories with GATE. EuroSafe Imaging, doi: 10.26044/esi2020/ESI-05719, 2020.
 - **D. F. Bauer**, A.-K. Schnurr, T. Russ, B. Waldkirch, L. R. Schad, F. G. Zöllner and K. Chung. Synthesis of CBCT Images from Digital Phantoms Using CycleGANs. Proc. 4th Conference on Image-Guided Interventions, Mannheim, Germany, 2019.
 - G. Kabelitz, **D. F. Bauer**, A.-K. Schnurr, T. Russ, I. Hermann, F. G. Zöllner, L. R. Schad and K. Chung. Evaluation phantom for multimodal imaging and image-guided needle interventions. Proc. 4th Conference on Image-Guided Interventions, Mannheim, Germany, 2019.
 - T. Russ, A. M. Abdelrehim, **D. F. Bauer**, L. R. Schad, F. G. Zöllner and K. Chung. Framework for image quality and physical dose predictions of arbitrary CBCT source-detector trajectories with GATE. Proc. 4th Conference on Image-Guided Interventions, Mannheim, Germany, 2019.
 - B. Waldkirch, **D. F. Bauer**, A.-K. Schnurr, S. Engelhardt, F. G. Zöllner, L. R. Schad and I. Wolf. Point-based Evaluation of Multimodal Non-Rigid Image Registration of Synthetic Abdominal Data Generated with a Digital Phantom and a Cycle-Consistent Network. Proc. 4th Conference on Image-Guided Interventions, Mannheim, Germany, 2019.
 - **D. F. Bauer**, A.-K. Schnurr, T. Russ, S. Goerttler, L. R. Schad, F. G. Zöllner and K. Chung. Synthesis of CT Images Using CycleGANs: Enhancement of Anatomical Accuracy. Proc. International Conference on Medical Imaging with Deep Learning, London, United Kingdom, 2019

Supervised Theses

- **Master's Thesis** "CT Image Reconstruction using End-to-End Deep Learning for Low Dose and Sparse View Sinograms", Constantin Ulrich, 2020.

12. Curriculum Vitae

Personal Data

Name **Dominik Fabian Bauer.**

Date of Birth **18. June 1993.**

Place of Birth **Schwäbisch Hall.**

Nationality **German.**

Education

2018 – today **Doctoral Candidate: Dr. sc. hum.,**
Ruprecht Karl University of Heidelberg,
Topic: Digital and Physical Phantoms for
Image-guided Interventions,
Supervisor: Prof. Dr. rer. nat. Lothar R. Schad.

2016 – 2018 **Master of Science: Physics,**
Ruprecht Karl University of Heidelberg, Grade: 1.2
Master's thesis: Study of the Octavius ionization chamber
array as a film replacement for clinical ion beam quality
assurance, performed at Deutsche Krebsforschungszentrum
(DKFZ).

2012 – 2016 **Bachelor of Science: Physics,**
Ruprecht Karl University of Heidelberg, Grade: 1.9
Bachelor's thesis: Gefrieren einzelner Wassertropfen auf
monokristallinem Saphir in einer Coldstage, performed at
Karlsruhe Institute of Technology (KIT).

2009–2012 **Abitur,**
Wirtschaftsgymnasium Schwäbisch Hall, Grade: 1.6.

Experience

2022 – today **Lecturer,**
Baden-Wuerttemberg Cooperative State University (DHBW)
Stuttgart,
Subject: Physics.

2018 – today **Research Assistant,**
Computer Assisted Clinical Medicine, Medical Faculty
Mannheim, Ruprecht Karl University of Heidelberg,
since 06.2021: Project Leader „Image Guided Interventions“,
since 10.2019: Lecturer ”Physics of Imaging Systems”.

- 2018 – 2022 **Lecturer**,
MTA School at Ludwigshafen Municipal Hospital,
Subject: Physics Practical Course and Radiophysics.
- 2017 – 2018 **Student Employee**,
Physikalisch-Technische Bundesanstalt (PTB).
- 2016 – 2018 **Student Employee**,
Heidelberg Ion-Beam Therapy Center (HIT).
- 2016 – 2018 **Student Assistant**,
Ruprecht Karl University of Heidelberg.

Scholarships & Grants

- 2021 **ISMRM Interventional MR Study Group Award**,
Best Interventional MR Abstract (Hardware Development and Integration),
An Anthropomorphic Pelvis Phantom for Prostate Brachytherapy and Biopsy.
- 2021 **ISMRM Magna Cum Laude Merit Award**,
An Anthropomorphic Pelvis Phantom for Prostate Brachytherapy and Biopsy.
- 2021 **ISMRM Educational Stipend.**
- 2019 **Erasmus+ Mobility Program.**
- 2019 **NVIDIA GPU Grant**,
Titan V graphics card.

13. Acknowledgements

First and foremost, I want to thank Prof. Dr. Lothar R. Schad for the opportunity to start my PhD thesis in his research group and for the supervision of this thesis.

I take this opportunity to express my gratitude to Prof. Dr. Frank Zöllner for his scientific guidance and his mentorship.

Many thanks to all current and former members of the CKM. The coffee breaks in front of house 8 or in the seminar room were always a welcome distraction and were filled with exciting conversations.

Special thanks go to Tom, for always having my back. Your constant pursuit of self-improvement is both inspiring and motivating.

Alena, thank you so much for your tireless support at all stages of my PhD thesis. Without your groundwork, many deep learning projects in this thesis and in our group would not have been possible.

Thank you, Christian, for always taking the time to help me with all kinds of problems.

Last but not least, I want to thank my family for all the support and encouragement. Above all, I am grateful for the tireless encouragement of my wife Aylin.