DISSERTATION

submitted to the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of Heidelberg University, Germany

for the degree of

Doctor of Natural Sciences

Put forward by

Tom Russ

Born in Heidelberg, Germany

Oral Examination: February 15$^{\text{th}}$ 2022

Interventional Cone-Beam CT
with Arbitrary Acquisition Orbits:
Simulation and Reconstruction

Interventionelle Kegelstrahl-CT mit beliebigen Aufnahme-Orbits: Simulation und Rekonstruktion

Die Cone-Beam-CT (CBCT) wird aufgrund ihrer Geschwindigkeit, Mobilität und räumlichen Auflösung häufig in der bildgestützten Chirurgie eingesetzt. Das Aufkommen robotischer interventioneller Röntgensysteme hat die Flexibilität beim Design von CBCT-Aufnahmeorbits drastisch erhöht. Alternative Umlaufbahnen ermöglichen es, das CBCT-System um störende chirurgische Geräte herum zu steuern und strahlenundurchlässige Strukturen in der Nähe von Bildgebungsinteressen zu vermeiden, um die Bildqualität zu verbessern. Aufgrund der zeitlichen Beschränkungen im Operationssaal werden Echtzeit-Algorithmen zur Bestimmung der idealen Aufnahmegeometrie und zur Rekonstruktion der entsprechenden Projektionsdaten benötigt, um beliebige Trajektorien in der klinischen Praxis umzusetzen. In dieser Arbeit wird die beste Alternative zu einer kreisförmigen Standardgeometrie durch die Simulation von CBCT-Bildern mit Hilfe eines Projektionsoperator-Ansatzes ermittelt, der mit robusten Monte-Carlo-Simulationen (MC) validiert wird. Die Vorhersagen der beiden Simulationsmethoden stimmen überein, wobei der Projektionsoperator-Ansatz nur sieben Minuten benötigt. Zweitens wird eine Methode zur schnellen Rekonstruktion beliebiger Orbits auf der Grundlage von Convolutional Neural Networks (CNNs) vorgestellt. Das vorgeschlagene Rekonstruktionsverfahren liefert eine ähnliche Bildqualität (nRMSE=0.060) im Vergleich zu state-of-the-art Rekonstruktionen (nRMSE=0.045) und reduziert gleichzeitig die Rechenzeit drastisch um 90%. In der vorliegenden Arbeit werden schnelle Methoden entwickelt, um geeignete Alternativen zur Standardkreisbahn zu bestimmen und adäquate Rekonstruktionen zu erstellen. Diese Methoden fördern die Implementierung beliebiger CBCT-Orbits, um die Flexibilität und Präzision während bildgesteuerter Operationen zu erhöhen.

Interventional Cone-Beam CT with Arbitrary Acquisition Orbits: Simulation and Reconstruction

Cone-beam CT (CBCT) is commonly used in image-guided surgery due to its speed, mobility and spatial resolution. The advent of robotic interventional X-ray systems has opened the door to dramatically increased flexibility in the design of CBCT acquisition orbits. Alternative orbits allow to steer the CBCT system around interfering surgical equipment and to avoid radiopaque structures near imaging tasks to improve image quality. Due to time constraints in the operating room, real-time algorithms to determine the ideal acquisition geometry and reconstruct the corresponding projection data are needed to implement arbitrary trajectories in clinical practice. In this thesis, the best alternative to a standard circular orbit is determined by simulating CBCT images using a projection operator approach, which is validated with robust Monte Carlo (MC) simulations. The predictions of both simulation methods coincide, with the projection operator approach requiring only seven minutes. Second, a method for fast reconstruction of arbitrary orbits based on convolutional neural networks (CNNs) is presented. The proposed reconstruction technique provides similar image quality (nRMSE=0.060) when compared to state-of-the-art reconstructions (nRMSE=0.045), while drastically reducing the computation time by 90%. The presented work develops fast methods to determine suitable alternatives to the standard circular orbit and to provide adequate reconstructions. These methods facilitate the implementation of arbitrary CBCT orbits to increase flexibility and precision during image-guided procedures.

Acknowledgements

Contents

1. Introduction

Cone-beam computed tomography (CBCT) is commonly used in image-guided surgery [1],[2],[3], image-guided radiation therapy [4],[5] as well as for imaging of musculoskeletal extremity [6], breast [7], head and dental imaging [8],[9] due to its speed, mobility, spatial resolution and contrast. For image-guided procedures, CBCT systems installed on mobile C-arms are highly flexible imaging solutions and support minimally invasive surgeries, which are increasingly performed with robotic assistance in the modern, hybrid intervention room; such surgeries are designed to minimize tissue damage for improved patient recovery times and comfort. Common image-guided procedures include pedicle screw insertion, percutaneous needle placement and endovascular aneurysm repair, to name a few. In particular, robot-assisted surgeries have become increasingly popular in recent years - in Germany alone, the number of robot-assisted surgeries has been rising from 8600 in 2013 to more than 37000 in 2020 [10],[11] - due to their potential to improve precision, accuracy and safety during interventions, as well as overall procedure time [12],[13]. Previously, robot-assisted image-guided interventions have been shown to reduce imaging dose and procedure duration while providing comparable accuracy and complication rates [14],[15],[16].

The abundance of surgical equipment limits the available space in the intervention room. In most scenarios, the standard circular trajectory used in CBCT imaging is not feasible without rearranging equipment *and/or* personnel. Not only interfering equipment, but also patient size can impose such challenges [17]. Alternative data acquisition orbits can help by steering the CBCT system around any interfering structure. Even in cases with sufficient range of motion, suboptimal positioning of the imaging task near metal implants, surgical tools, or other radiopaque structures such as bone can result in inadequate image quality for CBCT images acquired with the standard circular trajectory [18]. The deterioration of image quality originating from metal artifacts and radiopague structures results from a bias *and/or* discrepancy between the actual physical processes of image formation (*i.e.* the forward model) and the assumed model used for reconstruction of the projections (*i.e.* the inverse model) [19]. In some cases, non-circular orbits can be used to substantially improve image quality [20].

The advent of robotic interventional X-ray systems has opened the door to dramatically increased flexibility in the design of CBCT acquisition orbits. As a result, a variety of alternative orbits - all of which can theoretically be implemented on a robotic C-arm CBCT - have recently been investigated to address various issues: Sinusoidal orbits of different frequencies and combinations thereof to improve image quality [21],[22],[20], a combination of multiple arcs to avoid otherwise interfering structures [23],[17], circular tomosynthesis to reduce imaging dose [24], and line-ellipse-line or reverse helical orbits to increase the field-of-view (FOV) for long objects [25],[26],[27]. It has also been generally stated that the adjustment of the orbit to the specific patient anatomy, surgical procedure, and available actuation space should be based on pre-interventional imaging [21],[23]. Consequently, the

ideal X-ray source trajectory can be determined either between pre-interventional imaging and surgery (*i.e.* offline) or intra-interventional (*i.e.* online).

Adaptive orbit optimization algorithms attempt to find the best possible source trajectory, typically by exhaustively testing possible candidates from a general class of geometries. In recent years, there has been continuous work on the formulation of sophisticated objective functions for orbit optimization. Cascaded systems analysis as well as local noise and spatial resolution approximations have been used to calculate the detectability index of tomosynthesis [28],[29] and task-based orbits [30],[21] to optimize data collection. The detectability index incorporates knowledge of the imaging task, as well as the spatial-frequency dependent transfer of the spatial resolution and noise by the imaging system; the use of these quantities is generally accepted as the ideal definition of imaging performance [31]. In general, such computations are very time consuming.

However, there is an urgent need for an optimization algorithm that accommodates the severe time constraints of intra-interventional implementation; due to unanticipated changes in surgical equipment placement *and/or* patient size, orientation, and fixation [32], online trajectory optimization is a particularly valuable tool. Recently, efforts have been made to perform orbit optimization on-the-fly to reduce metal artifacts by leveraging the speed of convolutional neural networks (CNN) to predict the next best view angle [33] and to match circular CBCT image quality while acquiring undersampled projection data by combining multiple arc trajectories [17]. To the author's knowledge, an online algorithm for predicting the general image quality for arbitrary source orbits under actuation space constraints has not been developed yet.

To validate the prediction of the imaging outcome, most studies use data acquired with clinical CBCT devices [29],[34],[22]. Evaluating the quality of clinical images is primarily done by comparing spatial resolution metrics such as the modulation transfer function (MTF) and noise metrics such as the noise power spectrum (NPS). However, the experimental acquisition of such data presents some challenges, as manufacturers currently do not fully support the implementation of arbitrary source orbits on their devices. Consequently, the implementation of arbitrary trajectories on clinical devices is an ongoing effort [20],[35],[27] and faces major challenges such as robust calibration due to geometrical instabilities of CBCT systems [36],[24]. As a result, CBCT simulation can be a valuable tool in the study of arbitrary source orbits. For this reason, a fast but robust simulation strategy is needed to optimize and validate data acquisition strategies with non-circular CBCT orbits. The first goal of this thesis is to establish an online optimization scheme for arbitrary CBCT orbits by performing computed tomography (CT) simulation and subsequent image quality evaluation studies.

Modeling of fundamental physical processes such as photon scattering is essential for CT image simulation. The current gold standard for robust photon interaction simulations is the integration of morphological photon attenuation data into a Monte Carlo (MC) framework [37]. The general idea of MC analysis is to model a specific

CT imaging system, *i.e.* X-ray spectrum, pre-filtering, collimation, detector pixels, etc., and to propagate each X-ray photon through the system by sampling the probability density functions of the physical interactions involved (e.g. photoelectric effect, etc.). This process is computationally expensive as it is repeated for the huge number of photons normally involved in CT imaging (around 20 billion per CBCT projection) to form a projection image [38].

A common research effort has been the development of open-source MC techniques, which resulted in the availability of multiple alternatives such as MCNP [39], EGS4 [40] and GEANT4 [41]; however, most implementations require intricate knowledge of FORTRAN or C++ languages to run. The GEANT4 application for tomographic tmission (GATE) on the other hand, provides a framework for MC algorithms based on GEANT4 [42]. GATE was developed to accommodate the unique simulation needs of medical imaging and is executed using an easy-to-learn script language that provides the ability to design a wide range of experimental configurations. GATE was originally developed for emission tomography, but was later revised for CT simulation [43]. Since these simulations are very time consuming (around 7000 photons per second per central processing unit (CPU) core/thread), there are ongoing efforts to increase the computation speed of GATE using graphical processing units (GPU) [44] and the European Grid Infrastructure [45],[46].

Recent progress in the field of machine learning demonstrated results comparable to MC simulations, for instance when estimating the extent of photon scattering artifacts based on digital phantom data [47]. The use of generative adversarial networks (GAN) for simulating photo-realistic retinal images from morphological data also showed promising results [48]. In general, simulating medical images, for instance via the synthesis of magnetic resonance imaging (MRI) with the GAN architecture, achieves image quality close to real-world data. However, a major issue of this technique is that the generated images often fail to preserve anatomical features [49]; this can be partly addressed by modifying the objective function during network training, as Olut *et al.* have shown for the synthesis of MRI angiography data based on T1- and T2-weighted images [50]. Especially for medical imaging applications, where matched training data is scarce, the GAN architecture is not ideal due to the need for paired training samples. An extension of the GAN in the form of the Cycle-consistent generative-adversarial networks (CycleGAN) architecture can overcome this limitation by learning a reciprocal mapping between the two domains using unpaired samples [51]. Using this architecture already showed feasible results for the translation from the MRI to the CT domain [52].

An often used alternative to MC simulations and CNNs are projection operators, especially for CT reconstruction. The fastest way to compute a forward projection is a voxel-ray intersection approach based on the Siddon ray-tracing algorithm [53],[54]; it computes the distance of two intersection points of orthogonal, uniformly spaced and parallel planes (as an alternative representation of pixels) of an infinitesimally narrow ray within a given voxel and multiplies it by the voxel intensity. However, due to the finite voxel size, Siddon ray-tracing is known to cause discretization artifacts,

especially for larger voxels. Therefore, alternative algorithms have been proposed based on trilinear interpolation methods [55] or on approximating the voxel footprint by two separable functions [56]. In this way, projection operators remain a fast way to compute CBCT projections while maintaining geometric accuracy. However, additional processing is required to account for physical effects such as beam hardening or noise.

Apart from the prospective optimization of orbit geometry before or during surgery, the subsequent reconstruction of the projections acquired with non-circular trajectories remains a challenge in the operating room. In general, reconstruction algorithms for non-circular data rely on both analytical and model-based methods. Theories for exact solutions exist for certain classes of non-circular orbits; some are a type of filtered backprojection [57],[58] or differentiated backprojection - with a subsequent inverse Hilbert transform in the image domain [59]. However, for an accurate reconstruction of a region, there is a general requirement that the region is covered in so-called R-lines; therefore, analysis of R-line coverage is essential when investigating new source trajectories [60],[61].

Alternatively, model-based iterative reconstruction (MBIR) techniques can be applied to arbitrary orbit data without the need for adjustments; MBIR provides a general best-estimate based on the available data, as it can incorporate knowledge of the stochastic process of image formation and thus improved noise suppression [62],[21]. However, due to their iterative nature and the associated repeated forward and backprojections, these algorithms are computationally expensive, which is a major limitation, especially for interventional applications. The recent proliferation of data-driven and machine learning-based reconstruction methods offers opportunities for increased reconstruction speed *and* image quality comparable to MBIR.

Data-driven methods for 3D CT reconstruction have attracted much interest after the recent success of deep learning techniques, mainly because of their speed and accuracy. There are several types of proposed network architectures, for instance the direct reconstruction of sinogram data using CNNs [63]. Another approach is to use CNNs to unroll iterative algorithms, which can improve convergence paths for MBIR, among others [64],[65]. Since these algorithms are still iterative in nature and often combine CNN processing with domain transfers in the form of forward- and backprojection operators, they are still computationally intensive, especially for the huge image matrices associated with 3D image reconstruction. Consequently, the second goal of this thesis is to develop a data-driven approach for fast reconstruction of projection data acquired with arbitrary source orbits.

This dissertation's main goal is to facilitate the clinical implementation of arbitrary CBCT orbits for interventional imaging. To this end, this thesis addresses two important scientific questions:

1. Can CT simulations using either CNN deployment or projection operators predict the imaging performance for specific imaging tasks and acquisition orbits online, and can these predictions be verified by robust MC simulations?
2. Can the computational efficiency of CNNs be used to accelerate the reconstruction of arbitrary CBCT projections and provide image quality comparable to MBIR for a variety of different orbits without requiring adjustments?

To explore different CT simulation options for image quality predictions, three different methods are evaluated. First, a CNN-based approach using the CycleGAN architecture to generate CT images from the digital extended cardiac-torso phantom (XCAT) [66] is investigated. Specifically, various CNN configurations and architecture specifications are tested and optimized, including novel adaptions of the loss function to produce more accurate, task-specific CT images. A novel evaluation framework is proposed, consisting of already established image quality and similarity metrics, as well as newly introduced means to evaluate anatomical accuracy and noise realism in the simulated CT images. Second, an approach utilizing the MC algorithm to simulate CBCT projections is investigated. The potential of the European Grid Infrastructure is exploited with GateLab to accelerate GATE simulations. Third, CBCT projections are simulated using projection operators with the Astra toolbox [67],[68] in combination with a realistic noise model. The simulation methods are then applied to predict the imaging performance of a standard circular orbit and two alternative orbits for a specific imaging scenario in which the available actuation space is insufficient to realize a circular orbit. A digital phantom in the form of a slanted tungsten wire surrounded by a water cylinder was used to evaluate imaging performance using a supersampling scheme of line-spread function (LSF) curves to determine the MTF and thereby image quality.

To address the second goal of this thesis, a reconstruction procedure that uses CNNs is proposed for the fast reconstruction of arbitrary CBCT projections. In particular, a processing chain is developed in which the backprojection of the projection data is followed by a shift-invariant deconvolution step to both capture the dependence on the system geometry and to remove the blur associated with backprojection, followed by CNN processing. The deconvolution is based on the orbital trajectory and the intrinsic system response, but is only approximate. The CNN step is trained to mitigate the deficiencies in this approximate deconvolution. Each of these steps is computationally efficient and non-iterative, resulting in a fast processing chain. First, this technique is applied to individual (or pairs of) predefined orbits. Then, this technique is improved by training the networks on a variety of random orbits to facilitate the processing of new random orbits that were not used in CNN training. The proposed reconstruction scheme is applied to procedurally generated tetrahedral phantoms as well as to anthropomorphic data. The simulated data were further refined by adding low frequency textures (i.e., Gaussian peaks) to the tetrahedral phantoms and using a realistic noise model for all simulated projections. The proposed pipeline is compared to iterative reconstruction techniques in terms of quantitative performance measures relative to ground truth (GT).

2. Theoretical Background

This chapter first outlines some mathematical basics in Section 2.1, gives a brief overview of the fundamental CT concepts in Section 2.2, describes the methods for simulation and reconstruction used in this thesis in Section 2.3 and introduces basic principles when applying CNNs for image processing in Section 2.4.

2.1   Mathematical Basics

2.1.1   Fourier Transform

The Fourier transform $\mathcal{F}\{\cdot\}$ of a real or complex function $f(\mathbf{x})$ of an n-dimensional vector $\mathbf{x} = (x_1, ..., x_n)$ is defined as

$$\mathcal{F}\{f(\mathbf{x})\} = \left(\frac{\alpha}{2\pi}\right)^{n/2} \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} f(\mathbf{x}) e^{-\alpha i \mathbf{ux}} \, \mathrm{d}x_1 ... \, \mathrm{d}x_n = F(\mathbf{u}). \tag{2.1}$$

If the constant α is set to 2π, the Fourier transform simplifies to

$$F(\mathbf{u}) = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} f(\mathbf{x}) e^{-2\pi i \mathbf{ux}} \, \mathrm{d}x_1 ... \, \mathrm{d}x_n. \tag{2.2}$$

In the sense of a change of basis, the Fourier transform of a spatial signal can be understood as a decomposition into basis frequencies $e^{-2\pi i \xi}$. The inversion of $F(\mathbf{u})$ is equally defined as a decomposition

$$\mathcal{F}^{-1}\{F(\mathbf{u})\} = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} F(\mathbf{u}) e^{2\pi i \mathbf{ux}} \, \mathrm{d}u_1 ... \, \mathrm{d}u_n = f(\mathbf{x}). \tag{2.3}$$

In image processing, Fourier transforms are commonly used to extract and analyze the frequency content, *i.e.* periodic basis functions, of images. This is done by discrete Fourier transformation (DFT) due to the discrete nature of the voxel-based image representations in medical imaging. In the case of a one-dimensional discretely sampled function $f_a(n)$ Equation 2.2 can be rewritten as:

$$F_a(u) = \sum_{n=-\infty}^{+\infty} f_a(n\Delta\xi) e^{-2\pi i u n \Delta\xi}, \tag{2.4}$$

where the discretely sampled function is expressed via dirac-delta sifting:

$$f_a(n\Delta\xi) = f(\xi) \sum_{n=-\infty}^{+\infty} \delta(\xi - n\Delta\xi) \tag{2.5}$$

2.1.2 Convolution Theorem

The convolution of two functions $f(x)$ and $g(x)$ is a particular kind of integral transform and defined as

$$(f * g)(x) = \int_{-\infty}^{+\infty} f(\xi)g(x - \xi) \, d\xi = h(x). \tag{2.6}$$

The convolution theorem states that

$$\begin{aligned} \mathcal{F}\{(f * g)(x)\} &= F(u) \cdot G(u), \\ \mathcal{F}\{(f \cdot g)(x)\} &= F(u) * G(u), \end{aligned} \tag{2.7}$$

holds and its inverse is given by

$$\begin{aligned} \mathcal{F}^{-1}\{F(u) \cdot G(u)\} &= (f * g)(x), \\ \mathcal{F}^{-1}\{F(u) * G(u)\} &= (f \cdot g)(x). \end{aligned} \tag{2.8}$$

This process can be inversed to yield a so-called deconvolution operation:

$$\begin{aligned} (f * g)(x) &= h(x), \\ \mathcal{F}^{-1}\{F(u) \cdot G(u)\} &= \mathcal{F}^{-1}\{H(u)\}, \\ \mathcal{F}^{-1}\{F(u)\} &= \mathcal{F}^{-1}\left\{\frac{H(u)}{G(u)}\right\}, \\ f(x) &= \mathcal{F}^{-1}\left\{\frac{H(u)}{G(u)}\right\} \end{aligned} \tag{2.9}$$

2.2 Computed Tomography (CT)

Wilhelm Conrad Röntgen discovered a new form of radiation he called X-ray in 1895, and was promptly awarded the first Nobel Prize in Physics for his discovery shortly after in 1901. His observations: X-rays are invisible to the human eye, they can penetrate matter and they are able to create biological changes in living tissue [69]. It did not take long until the first medical X-ray images were taken. Although the idea for cross-sectional imaging was already existing and a mathematical basis to calculate cross-sections from transmission measurements was already established by Johann Radon in 1917 [70], the first CT scanner was developed much later in 1972 by Godfrey N. Hounsfield. Together with Allan M. Cormack, Hounsfield received the Nobel Prize in Physiology or Medicine for the invention of the CT scanner in 1979. In medical CT imaging, the patient is irradiated with X-rays from multiple directions and the number of X-rays transmitted through the patient are recorded. The X-ray measurements are then used to calculate cross-sectional distributions of the attenuation coefficient, allowing for a non-invasive study of the patient's anatomy.

The following subsections briefly introduce the physical concepts of CT imaging, *i.e.* relevant electron and photon interactions with matter, which are based on the books by Buzug, Hsieh and Maier *et al.* [71],[72],[73]. In addition, some concepts of interventional cone-beam CT, CT image noise and common image artifacts are summarized.
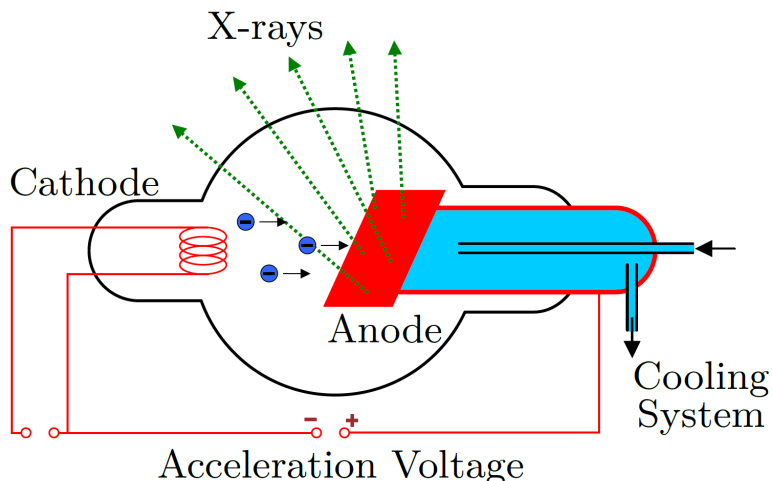
Figure 2.1: Production of X-rays in the X-ray tube. Free electrons are created at the anode by heating the filament. Electrons are accelerated towards the anode, where there kinetic energy is partly converted to X-ray radiation. As most of the energy is converted to heat, an anode material with high melting point is chosen and cooling systems are implemented. Figure reprinted from [73] under Creative Commons Attribution 4.0 International License: https://creativecommons.org/licences/by/4.0/

2.2.1 Production of X-rays

The X-rays used for CT imaging are commonly produced in an X-ray tube by ir-radiating a metal anode with an accelerated beam of electrons. The electrons are generated in a previous step by heating a filament at the cathode. In the heating process, the thermal energy of the electrons in the filament exceeds the binding energy (or work function) of the material and free electrons are ejected from the filament. Subsequently, electrons are accelerated in the X-ray tube via an external electric field of voltage U. After acceleration, the electrons possess a kinetic energy $E_{e,kin}$ calculated via:

$$E_{e,kin} = \frac{1}{2}m_e c^2 = e \cdot U, \tag{2.10}$$

where m_e denotes the electron mass and e the elementary charge of $1.602 \cdot 10^{19}$ Coulomb. The X-ray tube and its components are depicted in Figure 2.1. The ki-netic energy of the electrons is converted into different forms of energies via several interaction mechanisms. X-rays are generated mainly by Bremsstrahlung (braking radiation) and ionisation of atomic (mainly inner-shell due to higher binding energy) electrons. Most of the primary electron's energy, however, is lost to the anode mate-rial via inelastic scattering with atomic electrons of the outer shells. Consequently, the secondary electrons are ejected and create so-called delta rays, cascades of low energy electrons. Eventually, these electrons are converted to vibrations of anode atoms, *i.e.* heat. Bremsstrahlung is the phenomenon of X-ray emission by decelera-tion of the primary electrons in the electric field of anode atoms and electrons. This deceleration of a charged particle is known to create an electrical dipole which in turn emits electromagnetic radiation. In case of a complete stopping of the primary
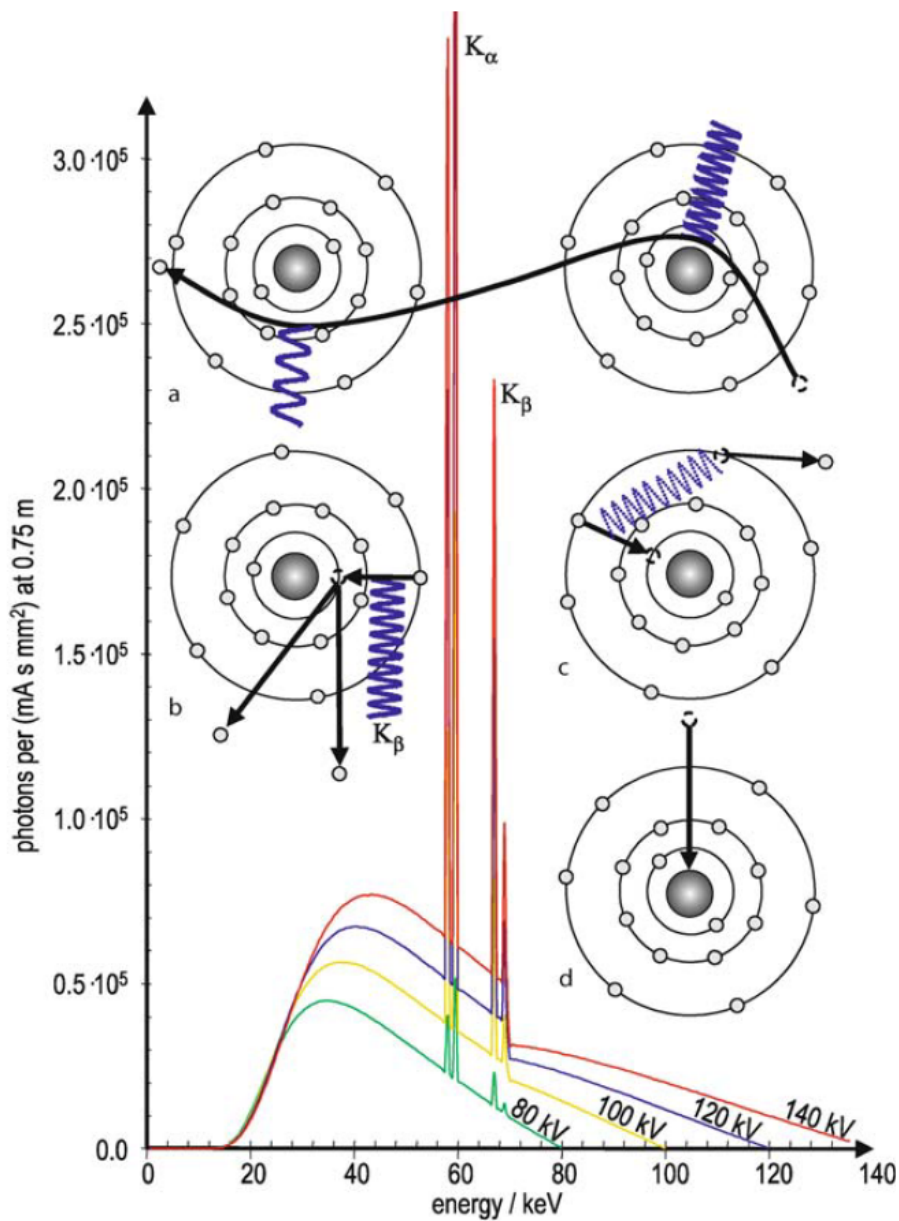
Figure 2.2: X-ray spectra of a tungsten anode at various tube voltages/electron energies. Note that the spectra are filtered by 2 mm of aluminum to remove low energy X-rays. The characteristic K_α and K_β lines are shown in the form of photon intensity peaks. The physical processes responsible for the production of the X-ray spectra are depicted: (a) Deflection/Deceleration in the electric field of the atom/electron producing Bremsstrahlung; Ionisation of inner-shell electron followed by filling of vacant position by outer-shell electron and subsequent X-ray emission (b) or emission of an Auger electron (c); (d) Direct collision of electron with atomic nucleus, thereby transferring maximum energy in the form of Bremsstrahlung. Reprinted with permission from [71]. Copyright ©2008, Springer-Verlag Berlin Heidelberg.

electron, the maximum energy of the electron is converted to X-ray radiation. The maximum energy of the emitted photons $E_{\gamma,\text{max}}$ can thus be calculated as follows:

$$E_{\gamma,\text{max}} = h\nu_{\text{min}} = h\frac{c}{\lambda_{\text{max}}} = e \cdot U, \tag{2.11}$$

where h denotes the Planck constant, c the speed of light, ν_{min} corresponds to the minimal photon frequency and λ_{max} to the maximum photon wavelength. As the deceleration of electrons is a cascaded process, a continuous X-ray spectrum is generated. Example X-ray spectra are illustrated in Figure 2.2. In addition, characteristic peaks in the X-ray spectrum are observed, which originate from ionisations of atomic electrons in the inner shells. If for instance an electron is ejected from the K-shell (innermost shell), the resulting vacant position is filled by an electron from the outer shells. The difference in binding energy of both electron levels is emitted as X-rays of energies that are characteristic for the respective anode material.

2.2.2 Photon Interactions with Matter

The X-rays produced at the anode are subsequently traversing the imaged object. To understand the CT image formation process, the concepts of photon-matter interactions have to be studied. This subsection introduces the concepts involved in X-ray attenuation.

The macroscopic phenomenon of photon attenuation in matter can be described by the Lambert-Beer law. The difference in photon number ΔN between transmitted and initial number of photons N per traversed unit path length Δx in a medium is defined by the linear photon attenuation coefficient μ via

$$\mu = -\frac{\frac{\Delta N}{N}}{\Delta x}. \tag{2.12}$$

In principle, the comparison of the radiation intensity I before (I_0) and after traversing a medium of thickness d yields the linear photon attenuation coefficient μ of the traversed medium

$$I(d) = I_0(E) \cdot e^{-\mu d}. \tag{2.13}$$

However, the Lambert-Beer law does not fully capture the forward model for CT imaging as it does assume that the X-rays traverse only one material and does not consider a spatially varying attenuation coefficient $\mu(s)$, which can be done via:

$$I(d) = I_0 \cdot e^{-\int_o^d \mu(s)\,\mathrm{d}s}. \tag{2.14}$$

Additionally, it does assume a mono-energetic X-ray beam instead of a multi-energetic X-ray spectrum $I_0(E)$ (see Figure 2.2) and does not incorporate any knowledge on the energy dependence of the attenuation coefficient $\mu(E, s)$. In reality, the forward model for the intensity of X-rays is calculated as follows:

$$I(d) = \int_0^{E_{\text{max}}} I_0(E) \cdot e^{-\int_o^d \mu(E,s)\,\mathrm{d}s}\,\mathrm{d}E. \tag{2.15}$$

In CT images the photon attenuation coefficient of the tissue μ_t is represented by the CT number ξ given in Hounsfield unit (HU):

$$\xi = \frac{\mu_t - \mu_w}{\mu_w} \times 1000 = (\hat{\mu}_t - 1) \times 1000, \tag{2.16}$$

where the linear attenuation coefficient of pure water at standard pressure and temperature is denoted μ_w and the hat indicates a quantity that is relative to water.

Brief Derivation of Beer's Law of Attenuation

The intensity I of photon radiation after traversing a distance Δx through a medium with attenuation coefficient μ can be described by

$$I(x + \Delta x) = I(x) - \mu I(x) \Delta x. \tag{2.17}$$

Taking the limit of (Equation 2.17) as Δx approaches zero leads to

$$\lim_{\Delta x \to 0} \frac{I(x + \Delta x) - I(x)}{\Delta x} = \frac{dI}{dx} = -\mu I(x). \tag{2.18}$$

By reordering of (Equation 2.18) an ordinary linear and homogeneous, first order differential equation is obtained

$$\frac{dI}{I(x)} = -\mu \, dx, \tag{2.19}$$

which can be solved via separation of variables in combination with the initial condition of $I(0) = I_0$

$$I(x) = I_0 e^{-\mu x}. \tag{2.20}$$

This is known as Beer's law of attenuation.

The Three Interaction Mechanisms at Diagnostic Energies

The linear attenuation coefficient can be factorized into the electron density n of the medium and the cross section per electron σ_e:

$$\mu = n \cdot \sigma_e, \tag{2.21}$$

The three dominating interaction processes in the so-called diagnostic photon energy window (15 keV up to 150 keV) can be divided into their specific cross sections per electron, as all interactions involve atomic electrons:

- Photoelectric absorption: σ_e^{photo},
- Compton scattering: σ_e^{Com},
- Rayleigh scattering: σ_e^{Ray},

which add up to the total cross section per electron σ_e^{tot}

$$\sigma_e^{tot} = \sigma_e^{\text{photo}} + \sigma_e^{\text{Com}} + \sigma_e^{\text{Ray}}. \tag{2.22}$$

Their respective contribution to the total cross section is illustrated in Figure 2.3.
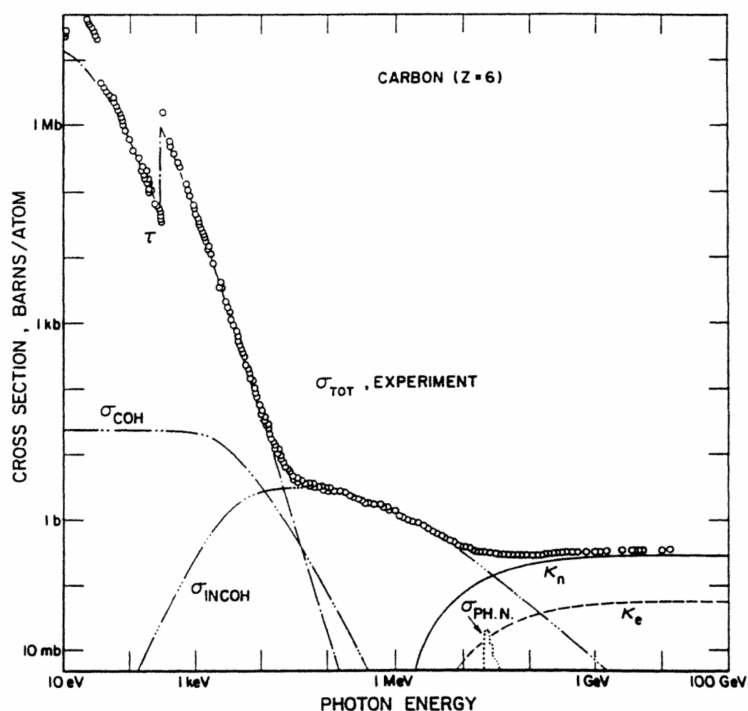
Figure 2.3: The contributions from Rayleigh-scattering σ_{coh}, photoelectric absorption τ and Compton scattering σ_{incoh} on the total measured cross section σ_{tot} (circles) in carbon are depicted. It is observable, that the dominant interaction process at low energies is the photoelectric effect and incoherent scattering (Compton) for higher energies of the diagnostic energy window (15-150 keV). ⓒInstitute of Physics and Engineering in Medicine. Reproduced by permission of IOP Publishing Ltd. All rights reserved [74].

Rayleigh scattering or Thomson scattering is a type of coherent, elastic scattering, meaning it does preserve the wavelength of the incident photon. This interaction is observed if they diameter of the interacting atomic nucleus is smaller than the photon's wavelength. The classical explanation for this phenomenon is that the electromagnetic field of the incident photon interacts with the electric field of the mostly inner-shell atomic electrons, setting them into oscillation. As discussed above, the corresponding acceleration and deceleration of charged particles create an electric dipole, which emits electromagnetic radiation of the same frequency as the incident photon. The interaction probability can be described by the Thomson cross-section [75]:

$$\sigma_e^{\text{Ray}} = \sigma_{\text{Thom}} = \frac{8\pi r_e^2}{3} \frac{\nu^4}{(\nu^2 - \nu_0^2)^2}, \tag{2.23}$$

where r_e denotes the classical electron radius, ν the photon frequency and ν_0 the natural frequency of the atomic electrons. Please note that the above descriptions and formulas only incorporate a classical scattering model and don't include quantum mechanical interaction principles. Rayleigh scattering is only relevant for low X-ray energies.

The interaction of photons with outer-shell electrons, *i.e.* weakly bound, quasi-free electrons, in which the incident photon imparts some kinetic energy to the electron - thereby ejecting it from the atom shell - is called **Compton scattering**. This process is incoherent, meaning the wavelength of the scattered photon λ_s does not match the wavelength of the incident photon λ_i. The change in wavelength $\Delta\lambda$ is dependent on the scattering angle Θ:

$$\lambda_s - \lambda_i = \Delta\lambda = \frac{h}{m_e c}(1 - cos(\Theta)). \tag{2.24}$$

Both the ejected electron and the scattered photon can partake in following ionization events. The probability of Compton scattering can be described by the Compton cross-section derived from the Klein-Nishina formula [76]:

$$\sigma_e^{\text{Com}} = 2\pi r_e^2 \left[\left(\frac{1 + E_{\gamma,s}}{E_{\gamma,s}^2} \right) \left(2\frac{(1 + E_{\gamma,s})}{1 + 2E_{\gamma,s}^2} - \frac{\ln(1 + 2E_{\gamma,s})}{E_{\gamma,s}} \right) \right.$$
$$\left. + \frac{\ln(1 + 2E_{\gamma,s})}{2E_{\gamma,s}} - \frac{1 + 3E_{\gamma,s}}{(1 + 2E_{\gamma,s})^2} \right], \tag{2.25}$$

where $E_{\gamma,s}$ denotes the energy of the scattered photon. Because the cross-section of Compton scattering does not involve any additional dependency on the material the photon interacts with, the probability of Compton events is only dependent on the electron density (Equation 2.21). Compton scattering only produces low-contrast information in the CT image formation process, as the difference in electron density of tissues is small. In this regard, Compton scattering is a rather disadvantageous interaction mechanism for conventional CT imaging, as it contributes to patient dose and scatter artifacts because the scattered photon is not locally absorbed and can still reach the detector (see Section 2.2.5). In addition, photons can be back-scattered from the back panel of the detector and thereby produce ghost images of

the panel further distorting the image.

In contrast to the partial energy transfer during Compton scattering, all of the photon energy is absorbed by the atom via **photoelectric absorption**. During this process, the atomic electron is ejected from the atom. Consequently, the energy of the photon is deposited locally to the lattice in the form of heat by cascaded processes of the ejected photoelectron. The interaction probability can be described by the absorption coefficient α and approximated via:

$$\alpha \propto \frac{\rho}{A} \frac{Z^4}{E_\gamma^3},$$

(2.26)

where ρ, A, Z denotes the density, atomic wight and atomic number of the absorber material, respectively, and E_γ the incident photon energy. The interaction probability decreases for higher X-ray energies. The Z-dependence of the photoelectric effect is the reason for the distinct contrast between bones (consisting of hydroxyapatite) and soft tissues as well as for the choice of contrast agents for X-ray imaging. Compared to Rayleigh and Compton scattering, photoelectric absorption does not contribute to scatter artifacts as the photon energy is locally absorbed.

2.2.3 Interventional Cone-Beam CT (CBCT)

Standard clinical CT machines consist of an X-ray tube and a multi-slice detector (usually up to 64 detector rows) rotating around the patient, those scanners are consequently called multi-detector computed tomography (MDCT). At the same time, the patient table is translated along the longitudinal axis, resulting in a helical motion of the X-ray source around the patient. Interventional imaging settings require highly mobile CT machines, that can acquire anatomical information in a flexible, fast way. For this purpose, X-ray source and detector can be installed on robotic C-arms. In contrast to MDCT scanners, flat-panel detectors are used, which consist of a 2D array of detectors with high spatial resolution. Because considerably more than 64 detector rows are used, usually up to 2000, those scanners are called cone-beam CT due to the corresponding shape of the X-ray beam, while MDCT is also referred to as fan-beam CT. CBCT machines conventionally use short-scan circular trajectories (180° + cone angle) around the imaged object to reconstruct 3D cross-sectional images of the patient. 360° acquisition is not necessary for full angular coverage due to data redundancies.

2.2.4 Image Noise

To understand the fundamentals of noise in CT images, one has to consider the stochastic nature of the image formation process inherent to electron-matter and photon-matter interactions. In the focus volume of the tube anode X-rays are generated from interactions of the incident electron beam with the anode material. Assuming that N electrons arrive at the anode in the time interval, [0, T], and partake in an X-ray producing interaction with probability p, the probability that the number of produced photons $\mathcal{N}$ is exactly n can be described by a Bernoulli distribution

$$P(\mathcal{N} = n) = \binom{N}{n} p^n (1-p)^{N-n},$$

(2.27)

with the binomial coefficient

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}. \tag{2.28}$$

The number of electrons arriving at the anode can be determined from the anode current I via $N = T \cdot I / e$, with e corresponding to the elementary charge. With the expectation value $E[\mathcal{N}] = n^*$ and the assumption $Np = n^*$ in the limit of $N \to \infty$, (Equation 2.27) simplifies to the Poisson distribution density function

$$P(\mathcal{N} = n) = \frac{(n^*)^n}{n!} e^{-n^*}. \tag{2.29}$$

The number of emerging photon quanta from the X-ray tube are thus Poisson distributed. For a Bernoulli detector irradiated with Poisson distributed quanta, the probability that n photons are detected $P_D(\mathcal{N} = n)$ can again be described by a Poisson distribution

$$P_D(\mathcal{N} = n) = \frac{(pn^*)^n}{n!} e^{-pn^*} \tag{2.30}$$

In combination with the attenuation of the generated photons in the scanned object, which also follows binomial statistics, this statistical chain is called a cascaded Poisson process. This leads to an expectation value of the random variable, which equals the variance of the random variable

$$E[\mathcal{N}] = E[(\mathcal{N} - n^*)^2]. \tag{2.31}$$

For the signal-to-noise ratio (SNR) of a Poisson-distributed signal this yields

$$\mathrm{SNR}(\mathcal{N}) = \frac{\mu}{\sigma} = \frac{n^*}{\sqrt{n^*}} = \sqrt{n^*}, \tag{2.32}$$

where μ corresponds to the mean (which is an estimate of the expectation value) of the signal and σ to the standard deviation. For a high number of X-ray quanta the Poisson distributed number of detected quanta approaches a Gaussian distribution

$$P_D(\mathcal{N} = n) \approx \frac{1}{\sqrt{2\pi n^*}} e^{-\frac{(n^*-n)^2}{2n^*}} \tag{2.33}$$

The scintillation detectors used in state-of-the-art CT systems are irradiated by multiple photons coincidentally and therefore work as energy-integrating detectors. This leads to deviations of the detected number of quanta from a Poisson distributed quantity.

2.2.5 Image Artifacts

This subsection briefly describes some common CBCT image artifacts.
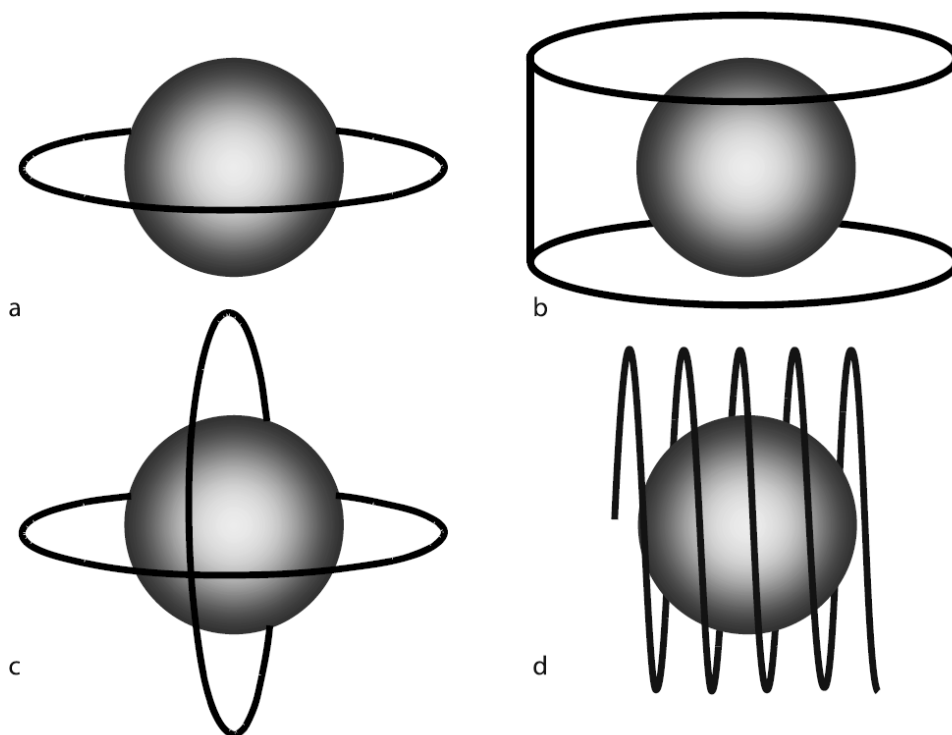
Figure 2.4: Four source orbits are illustrated by black lines together with the spherical FOV to be reconstructed. The standard circular CBCT orbit (a) does not fulfill the Tuy-Smith sufficiency condition, as not all planes through the FOV intersect the source trajectory at least once. In contrast, the two-circles plus line (the center of both circles are located outside the FOV) orbit (b), the double-circle (both circles are orthogonal) orbit (c) and the helical trajectory used in conventional MDCT all fullfill the Tuy-Smith condition. Note that all orbits share the common isocenter in the center of the spherical FOV. Reprinted with permission from [71]. Copyright ©2008, Springer-Verlag Berlin Heidelberg.

Cone-Beam Artifacts

The standard circular trajectory used for CBCT imaging does not capture sufficient projection data for an exact 3D reconstruction along the longitudinal axis. In contrast, the helical trajectory of conventional MDCT scanners does fulfill the Tuy-Smith-condition [77] for data completeness and enables an exact reconstruction of the axially covered FOV. An illustration of different source orbits is shown in Figure 2.4. The Tuy-Smith-condition states that for an exact reconstruction of a FOV, all planes through the FOV need to intersect the source trajectory at least once. For all planes through the FOV that do no intersect the source trajectory, undersampling artifacts in forms of streaks arise. In the case of the standard CBCT trajectory, such artifacts are called cone-beam artifacts.

Beam Hardening Artifacts

In the diagnostic energy range the total photon attenuation coefficient decreases for increasing energies. As a consequence, the mean energy of the photon spectrum is increased while traversing a medium, thereby decreasing the effective attenuation. For scanned media completely surrounded by photon absorbing matter, their attenuation coefficient is underestimated, if this knowledge is not incorporated into the backward model, (*i.e.* image reconstruction). Beam hardening artifacts basically arise from the difference between Equation 2.14 and Equation 2.15.

Photon Scattering Artifacts

Photon scattering artifacts arise when photons scattered either via Rayleigh or Compton scattering are detected in pixels and mistakenly assumed to have arrived from the position of the X-ray source. As a result, the measured signal on the detector is distorted. This phenomenon is more pronounced for CBCT in comparison to MDCT, as the opening angle of the photon beam is increased to irradiate the full flat-panel detector. As a result, image noise is substantially increased.

2.3 CT Image Simulation and Reconstruction

This section explains the theoretical background of the CBCT projection simulations and reconstruction algorithms used in this thesis.

2.3.1 Monte-Carlo (MC) Method for Medical Imaging

The MC method provides a numerical solution to address complex computational problems and is applied in many research fields, ranging from fluid dynamics to solid-state physics. The idea of using MC simulations for medical imaging is to model the process of image formation by simulating singular photon pathways. This is done via a random sampling of the interaction probability functions of the underlying physical effects, *i.e.* photoelectric absorption, Rayleigh and Compton scattering. First, a pseudo-random number is generated using a generator such as the Mersenne-Twister algorithm for instance [78]. Second, a technique to select the photon step size is applied, for instance by deriving the step size with the inverse distribution

method, which uses the inverse of the cumulative distribution function to sample a step size from each individual interaction probability. In the third step, the photon is moved to the point of lowest interaction length and either absorbed or scattered, depending on the interaction with lowest interaction length. These three steps are repeated until either the photon energy falls below a selected threshold or leaves the boundaries of the predefined volume. Due to the stochastic nature of photon-matter interactions, a very large number of photons have to be simulated to achieve sufficient accuracy.

2.3.2 Projection Operators

Another approach to simulate CBCT projections is the use of projection operators. The most simple and fast version is a voxel-ray intersection approach that is based on the Siddon ray-tracing algorithm [53],[54]. It computes the distance of two intersection points of orthogonal, uniformly spaced and parallel planes (as an alternative representation of pixels) of an infinitesimally narrow ray within a given voxel and multiplies it by the voxel intensity. Extensions of this algorithm include trilinear interpolation methods [55] or approximating the voxel footprint by two separable functions [56] to address geometrical inaccuracies such as discretization artifacts. Conversely, the backprojection operator is carried out by linking the inspected voxel with the source position via a ray that is prolonged until it reaches the detector. The corresponding pixel value is then assigned to the voxel via bilinear interpolation [79].

2.3.3 CT Image Reconstruction

This subsection briefly describes the iterative reconstruction techniques used in this thesis. All of the below algorithms do not require a standard circular trajectory. Instead, they can be used to reconstruct projection data acquired with arbitrary source orbits.

Simultaneous Algebraic Reconstruction Technique (SART) and Simulatenous Iterative Reconstruction Technique (SIRT)

In algebraic image reconstruction methods, the discrete inverse problem is considered. ART sets up linear equation systems in order to obtain pixel values. In contrast to Fourier based reconstruction these techniques enable the consideration of physical properties during reconstruction in the form of weighting factors a_{ij}, for instance. With the weighting factors

$$a_{ij} = \frac{\text{illuminated area of pixel j by ray i}}{\text{total area of pixel j}} \qquad (2.34)$$

the reconstruction can incorporate knowledge on the finite size of image voxels and detector pixels when before, infinitesimally small rays on point voxels and pixels were assumed. The linear equation system can be adjusted in order to account for the varying interaction lengths in each pixel as follows

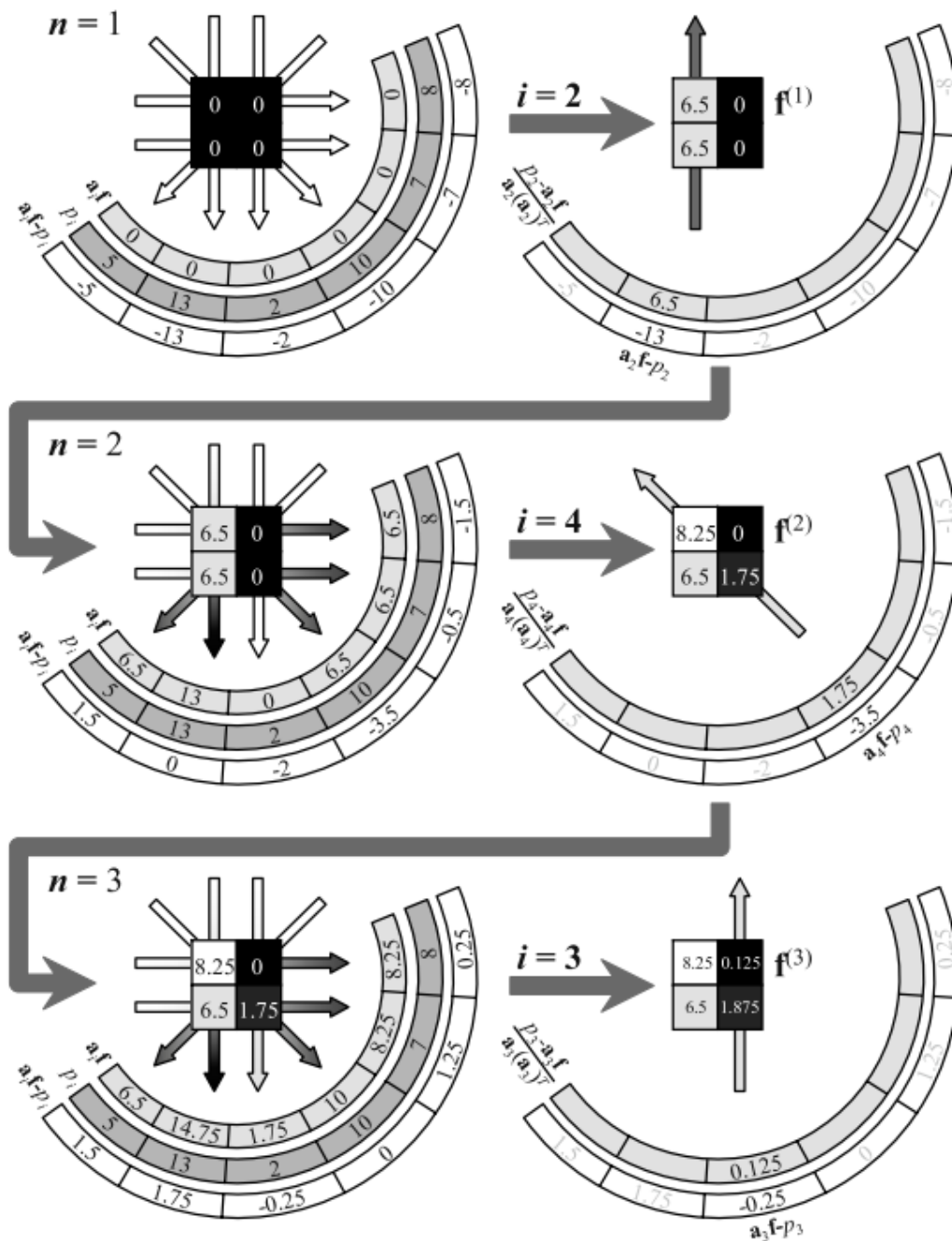$$\sum_{j=1}^{N} a_{ij} f_j = p_i. \qquad (2.35)$$

Figure 2.5: The first three Algebraic reconstruction technique (ART) iterations of the exemplary using $f^{(0)} = (0, 0, 0, 0)^T$ as an initial image. The correct image values are $\mathbf{f} = (8, 0, 5, 2)^T$. Reprinted with permission from [71]. Copyright ©2008, Springer-Verlag Berlin Heidelberg.

This can be rewritten with the $M \times N$ matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1N} \\ a_{21} & a_{22} & \ldots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1} & a_{M2} & \ldots & a_{MN} \end{bmatrix}, \tag{2.36}$$

where M corresponds to the number of obtained projections and N to the number of pixels, such that the system of equation becomes

$$\mathbf{p} = \mathbf{A}\mathbf{f}. \tag{2.37}$$

In CT the matrix $\mathbf{A}$ is called the system matrix. Thus $\mathbf{p}$ contains all values of the projection space and $\mathbf{f}$ contains all attenuation pixel values. So in order to obtain $\mathbf{f}$ one has to find the inverse of $\mathbf{A}$. One method is the *singular value decomposition*, where the $M \times N$ pseudo-inverse $\mathbf{A}^+$ is decomposed

$$\mathbf{A}^+ = \mathbf{V}\left(\mathbf{diag}\left(\frac{1}{\sigma_j}\right)\right)\mathbf{U}^T \tag{2.38}$$

into the $N \times N$ orthogonal matrix $\mathbf{V}$, the $M \times N$ orthogonal matrix $\mathbf{U}$ and the singular values σ_j. The solution can be found by

$$\widetilde{\mathbf{f}} = \mathbf{A}^+\mathbf{p}. \tag{2.39}$$

ART solves this problem iteratively. In a first step, a forward projection of the n-th image approximation $\mathbf{f}^{(n)}$ is performed and compared with the actual measured projection $\mathbf{p}$. This comparison yields a correction term which is applied to $\mathbf{f}^{(n)}$ via backprojection resulting in the $(n+1)$-th image approximation as follows

$$\mathbf{f}^{(n)} = \mathbf{f}^{(n-1)} - \lambda_n \frac{\left(\mathbf{f}^{(n-1)}\mathbf{a}_i - p_i\right)}{\mathbf{a}_i(\mathbf{a}_i)^{\mathrm{T}}}(\mathbf{a}_i)^{\mathrm{T}}, \tag{2.40}$$

where the heuristic relaxation parameter λ_n is introduced to speed up the convergence. A simple example of three ART iterations is shown in Figure 2.5. The simultaneous algebraic reconstruction technique (SART) [80] and the simultaneous iterative reconstruction technique (SIRT) are both of the ART family, but differ in the way the voxel values are updated. While SIRT updates the voxel values simultaneously from all projections in one iteration step, SART updates the voxel values projection by projection. This leads to some minor differences: SIRT performs slightly better in terms of overall computation time than SART, while SART provides slightly more accurate solutions.

Model-Based Iterative Reconstruction (MBIR)

MBIR algorithm incorporate more prior knowledge on the image formation process into the reconstruction routine. This is done to increase reconstruction accuracy. The MBIR method used in this thesis incorporates knowledge on the Poisson statistics of photon noise and penalizes image roughness $R(\mu)$ at the same time [81]. The available projection measurements (forward model) are assumed to be modeled via:

$$\bar{y}_i = g_i \exp\left(-\left[\mathbf{A}(\Omega)\mu\right]_i\right), \tag{2.41}$$

where the index i denotes measures related to the i-th measurement, $\hat{\mu}$ the expected mean photon count at an individual detector pixel, Ω incorporates knowledge on the underlying image geometry, μ the vector corresponding to the distribution of attenuation values in the imaged object and g incorporates knowledge on the pixel gain (bare beam fluence, detection efficiency, etc.). Deriving the log-likelihood function $L(y|\mu)$ via Equation 2.30 then yields the objective function to arrive at a best estimate $\hat{\mu}$:

$$\hat{\mu} = \text{argmax} L(y|\mu) - \beta R(\mu), \tag{2.42}$$

with the log-likelihood function

$$L(y|\mu) = \sum_i y_i \log\left(g_i \exp\left(-\left[\mathbf{A}(\Omega)\mu\right]_i\right)\right) - g_i \exp\left(-\left[\mathbf{A}(\Omega)\mu\right]_i\right), \tag{2.43}$$

and the Huber penalty $\text{R}(\mu)$[82]. Because $\log(y_i!)$ is only a constant, it does not matter for the maximization routine and can therefore be excluded from the penalized-likelihood estimator.

2.4 Convolutional Neural Networks (CNN) for Image Processing

This section gives a brief overview of some CNN concepts used for this thesis. For a more detailed introduction to machine learning with deep neuronal networks, the interested reader is referred to [83].

Generally speaking, artificial neuronal networks imitate the physiological function of an arrangement of interconnected neurons, each node is modeled as consisting of an input signal vector s_j (coming from an arbitrary number of nodes) that is modulated with a weighting factor vector w_j, with a bias vector b_j and eventually yields the output signal vector s_{j+1} via an activation function f_{act} to the next neuron (node) in the processing chain:

$$s_{j+1} = f_{\text{act}}(w_j \cdot s_j + b_j). \tag{2.44}$$

Most machine learning approaches for image processing are supervised, meaning that the output of the network of neurons for one specific input vector is connected to one specific input vector. During the network training procedure, the weight and bias parameters are updated iteratively according to a training scheme called backpropagation by comparing input and output vectors. The training procedure aims to minimize the loss function calculated between matched input and output vectors. CNNs are a specific type of neuronal networks, which mainly consist of convolutional layers. If data processing is performed with networks made up of more than one of such layers (the definition is vague), the technique is referred to as 'Deep Learning'.

2.4.1 Convolutional Layer

One powerful tool in machine learning for image processing is the connection of one vector (set of nodes) to the next via convolutional layers:

$$s_{j+1} = f_{\text{act}}(w_j * s_j + b_j). \tag{2.45}$$
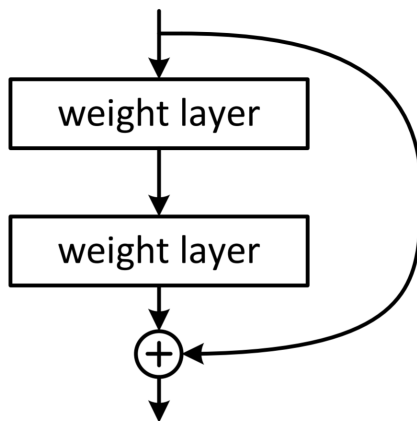
Figure 2.6: A residual block for image processing consists of multiple weight layers such as convolutional layers, in this case two. The input to the first layer in the block is stored and added to the output of the last layer.

A convolutional layer connects nodes of subsequent layers via a sliding window operation with specific window size and stride. The number of convolutions per convolutional layer determines the so-called channel dimension of the output vector. Because these convolution operations are highly parallelizable, the execution of CNNs on GPUs is really fast.

2.4.2 Max-Pooling Layer

Most CNN architectures for image processing incorporate upsampling and downsampling steps. A max-pooling layer is a form of downsampling operation. A sliding window operation is performed, similar to the convolutional layer above. Instead of summing up the weighted contributions from each vector element, the maximum value of all values in the window is assigned to the subsequent vector element.

2.4.3 Residual Blocks

Residual blocks have gained recent popularity in image processing. In principle, it's a series of convolutional layers, where the input to the first convolutional layer is temporarily stored and combined with the output of the last convolutional layer. A visualization of a residual block is shown in Figure 2.6.

2.4.4 Rectified Linear Unit (ReLU)

The rectified linear unit (ReLU) is a type of activation function that is often used in image processing to introduce non-linearities into CNN processing [84]. The ReLU activation function is applied after the weight and bias vectors:

$$f_{\mathrm{act,ReLU}} = \max(0, s_j) \tag{2.46}$$

3. Materials and Methods

This chapter is divided into six sections. Section 3.1 details the network configurations utilized in this thesis. Section 3.2.1 describes the different CT data - clinical CT scans and digital phantoms - used for network training, the simulation studies and the reconstruction experiments. Section 3.3 introduces the concept of the novel reconstruction scheme proposed in this thesis. Section 3.4 specifies how the used evaluation metrics are calculated. Section 3.5 outlines the conducted reconstruction and simulation experiments. Parts of this chapter have been published in [85][1], [86], [87], [88] and [89] and the corresponding figures and descriptions are adapted and partly replicated thereof.

3.1 Network Architectures

This section introduces the used network architectures, namely the U-Net (Section 3.1.1) and the ResNet (Section 3.1.2) architectures as well as the CycleGAN (Section 3.1.3) architecture.

3.1.1 U-Net

The U-Net is a type of encoder-decoder CNN, which can be divided into two halves. On one hand, the encoding branch consists of convolutional layers and down-sampling operations such as max-pooling layers, which are supposed to "fold" the information of the input image (n dimensions) into n+1 dimensions, effectively decreasing the image dimensions and increasing the number of the so-called side channels or feature maps. The decoding branch on the other hand consists of convolutional layers and transposed convolutional layers with a stride higher than one, thereby effectively up-sampling the information in the channels to eventually yield the size of the input images [90]. The U-net specifically incorporates skip-connections, which extract the feature maps before each down-sampling operation of the encoder branch and append it to the corresponding feature maps of same size in the decoding branch after each up-sampling step. Visual representations of U-nets can be seen in Figure 3.4 and in Figure 3.7.

3.1.2 ResNet

The ResNet utilized in this thesis is another form of encoder-decoder network like the U-net. In contrast to the U-net, however, the ResNet does not necessarily incorporate skip-connections but is rather characterized by its residual blocks [91], in this case the residual blocks are implemented between encoder and decoder branch. A visual representation of a ResNet can be seen in Figure 3.7.

---

[1]Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature International Journal of Computer Assisted Radiology and Surgery ©2019

3.1.3 CycleGAN

CycleGAN allow the supervised mapping between two domains X and Y given unpaired training samples [51]. The mapping functions $G : X \rightarrow Y$ and $F : Y \rightarrow X$ are called generators. Two additional adversarial networks D_X and D_Y, called discriminators, aim to distinguish between real and generated images (X and $F(Y)$ are distinguished by D_X and Y and $G(X)$ are distinguished by D_Y); both discriminator networks are trained on both real and generated images simultaneously. In this way, the generator networks learn to synthesize data which are more likely to be classified as real data by the discriminator networks while the discriminator is adversely trained to better discriminate between real and generated images. Specifically, the adversarial loss for generator network G and its discriminator network D_Y can be formulated as:

$$\mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) = \mathbb{E}_{x \sim p_{data}(x)}[\log(1 - D_Y(G(x)))] \\ + \mathbb{E}_{y \sim p_{data}(y)}[\log(D_Y(y)], \tag{3.1}$$

where $\mathbb{E}$ denotes the expectation value, $x, y \sim p_{data}(x, y)$ the data distributions. While the generator G tries to minimize this objective function, its adversary D_Y tries to maximize it. This is calculated in the same way for the mapping function F. Figure 3.6 shows the complete CycleGAN network architecture for the case of XCAT (X) and CT image (Y) domain. Both adversarial losses are then combined with the so-called cycle-consistency loss, which is calculated by comparing both pathway outputs of the consecutively applied generator networks, $F(G(X))$ and $G(F(Y))$, to their original input via:

$$\mathcal{L}_{\text{cyc}}(G, F, X, Y) = \mathbb{E}_{y \sim p_{data}(y)}[\|F(G(y)) - y\|_1] \\ + \mathbb{E}_{x \sim p_{data}(x)}[\|G(F(x)) - x\|_1], \tag{3.2}$$

where $\|\cdot\|_1$ denotes the L1 norm.

3.2 CT Data

This section details the CT images (Section 3.2.1) and digital phantoms (Section 3.2.2) utilized in this thesis.

3.2.1 CT Images

The clinical CT image data are obtained from four different sources. The first data set consists of in-house CT images, which are retrospectively utilized for this thesis. The remaining three sources are publicly available. One CT image per data set is exemplary illustrated in Figure 3.1.

In-House Scans

The in-house CT data set consists of 20 contrast-enhanced abdominal CT images from different patients. The scans were acquired at $90\,\text{kVp}$ with a dual source CT system (Somatom Force, Siemens Healthineers, Erlangen, Germany). A voxel spacing of $0.82 \times 0.82 \times 1.5\,\text{mm}^3$ and the reconstruction kernel Br36 (standard medium-smooth) were chosen. Patients were male and female of various ages. Most of those scans contained contrast enhanced blood vessels.

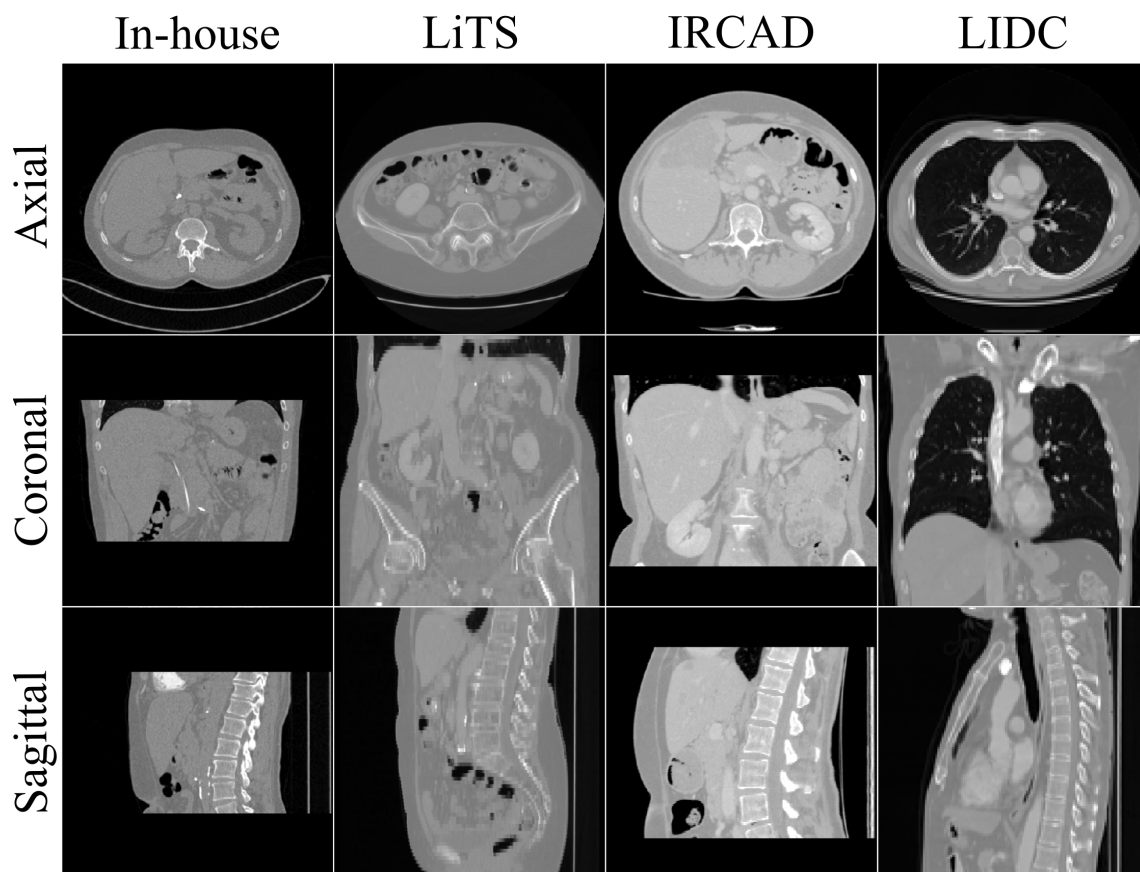**Figure 3.1:** Overview of the clinical CT image data sets used in this thesis, one image is shown for each data set. From top to bottom row: Axial, sagittal, coronal views; each column corresponds to another CT image data set: in-house CT, liver tumor segmentation challenge (LiTS), 3D image reconstruction for comparison of algorithm database (IRCAD), Lung image database consortium image collection (LIDC).

Liver Tumor Segmentation Challenge (LiTS)

The data set of the LiTS is comprised of 130 abdominal CT images of patients with liver lesions [92]. Imaging parameters and blood vessel contrast enhancement varied widely between scans.

3D Image Reconstruction for Comparison of Algorithm Database (IRCAD)

The first IRCAD database consists of 20 CT image, 10 of male patients and 10 of female patients [93]. 75 % of patients suffered from hepatic tumors. Imaging parameters as well as blood vessel contrast enhancement varied widely between scans.

Lung Image Database Consortium Image Collection (LIDC)

LIDC is a data set consisting of 1018 CT and CBCT scans acquired to screen the respective patient for lung cancer [94]. 800 thoracic CT and CBCT images are resampled to yield 128x128x128 cubic voxels with 0.5 mm edge length via nearest-neighbor interpolation. The utilized scan protocols varied widely between patients because the scans were conducted at various medical facilities with different scanner types.

3.2.2 Digital Phantoms

This subsection details the utilized digital phantoms. One exemplary XCAT phantom, the slanted wire phantom and one realization of the tetrahedral phantom are shown in Figure 3.2.

XCAT Phantom

The digital 4D XCAT phantom allows for the generation of cross-sectional photon attenuation images [66]. The anatomy of the phantom is based on the Visible Human Male and Female data sets of the National Library of Medicine [95], thereby creating a higher spatial resolution than that of standard CT protocols. The phantom is then realized using nonuniform rational B-splines to construct the organ shapes [96]. Multiple modifications can be made to the phantom including but not limited to: gender, different simulated X-ray tube energies, adjustment of the volume and position of various organs. The phantom also provides the segmentations of all included structures.

Slanted Wire Phantom

One experimental method to calculate the MTF of a CT imaging system consists of measuring a slanted wire phantom. This is why slanted metal wires are included in various image quality phantoms such as the widely used Catphan 500 (The Phantom Laboratory, Salem, NY, USA). To emulate such a phantom, a water cylinder of 70 mm radius, 140 mm length and 0.0215 1/mm linear attenuation coefficient - the attenuation of water for 70 keV photons - is generated using Matlab code (Math-Works, Natick, MA, USA). Inside the phantom, a tungsten wire of 50μm diameter (linear attenuation of tungsten: 49.25 1/mm) is placed at an 22.5°angle.
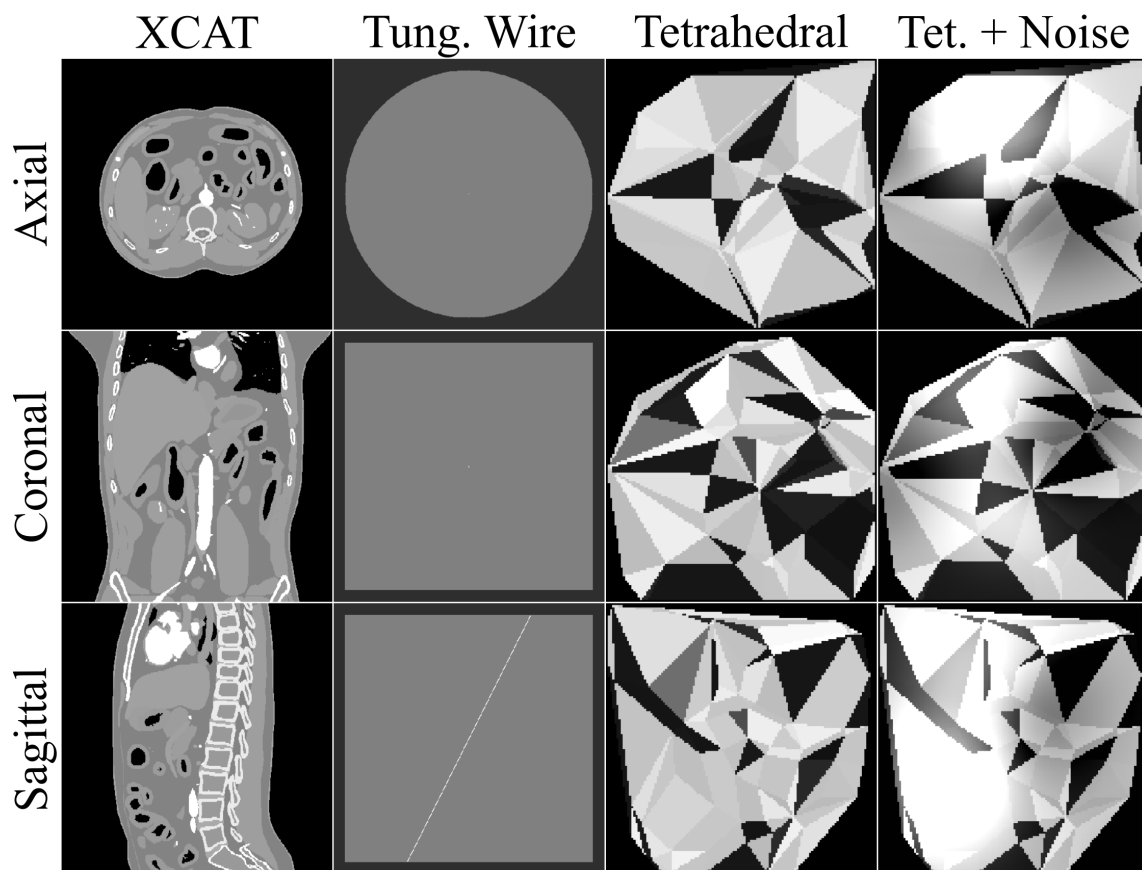
Figure 3.2: Overview of the digital phantoms used in this thesis. From top to bottom row: Axial, sagittal, coronal views; each column corresponds to another phantom: XCAT phantom, slanted tungsten wire phantom, tetrahedral phantom, tetrahedral phantom with low frequency noise.
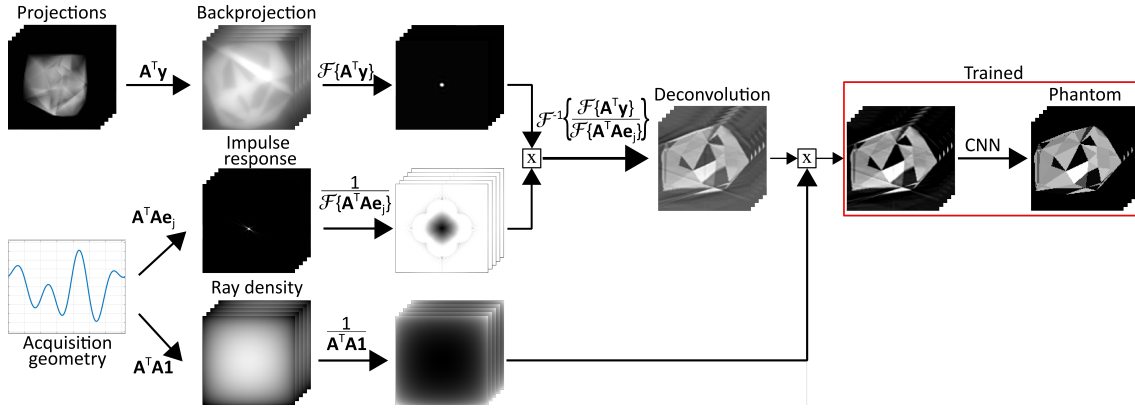
Figure 3.3: Flow chart illustrating the proposed reconstruction pipeline. First, an approximation of the system response is deconvolved, then a CNN is deployed to remove residual artifacts. Before CNN deployment, the deconvolution is divided by the ray density.

Tetrahedral Phantom

A scheme to procedurally generate tetrahedron phantoms is proposed. In an initial step, 40 vertex locations are sampled from a 3D normal distribution (ND) with a standard deviation of $\sigma = 80$ voxels around the center of the volume. After connecting the vertices using the Delaunay triangulation algorithm, the resulting mesh is voxelized on a 128x128x128 grid with 0.5x0.5x0.5 mm^3 voxel spacing. 3D Delaunay triangulation is used to connect a set of four discrete vertices to construct tetrahedrons in such a way, that no vertex is inside the circumcircle of any tetrahedron [97]. Within each voxelized tetrahedron, a uniform attenuation coefficient μ_P is randomly assigned based on the distribution of voxel values in a reference abdomen CT scan (*sans* background), which was acquired at 100kVp. To avoid the use of piecewise-constant phantom data, the option to add/subtract a random number (3D ND, mean $\mu = 30, \sigma = 10$) of Gaussian peaks with random center locations (3D ND, $\mu = 64, \sigma = 32$), widths given by a standard deviation uniformly distributed in the interval [256,1280] and random (signed) amplitude (ND, $\mu = 0, \sigma = \max(\mu_P)/4$), is implemented.

3.3   Proposed Reconstruction Pipeline

This section outlines the novel reconstruction scheme proposed in this thesis. The algorithm is visualized in Figure 3.3 and additional processing steps are detailed.

Recalling the basics from Section 2.3.3 and presuming log-transformed projection data, tomographic reconstruction seeks to solve the following inverse problem:

$$y = \mathbf{A}(\Omega)\mu, \tag{3.3}$$

where y denote the measured line integrals of attenuation (e.g., projections) and μ is the distribution of attenuation values in the object. Here, the dependence of
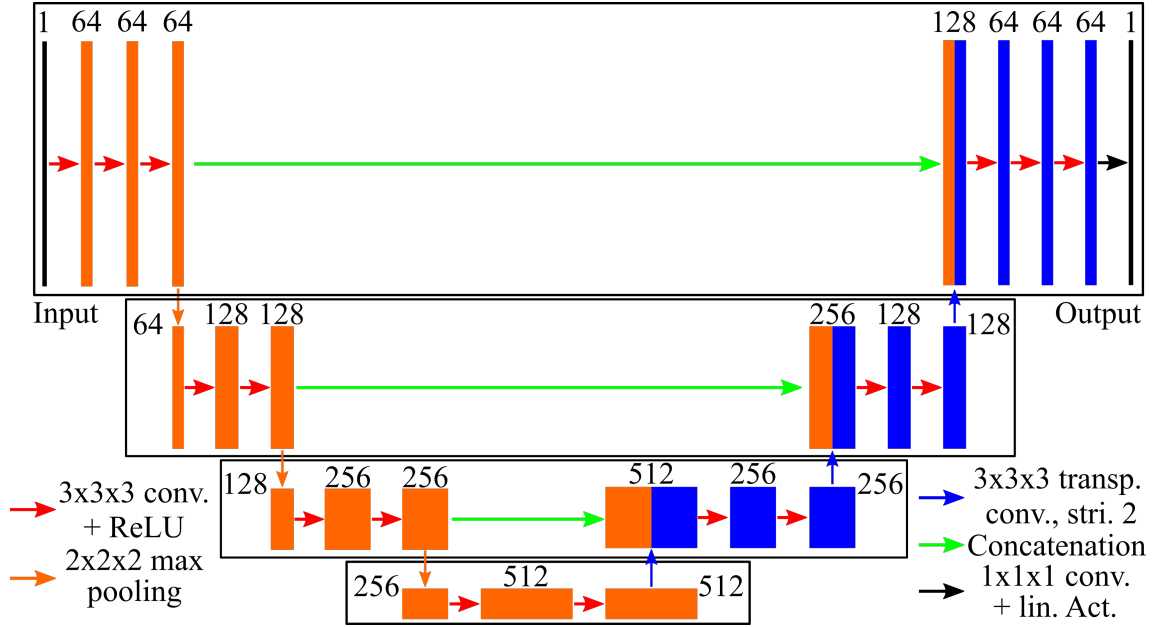
Figure 3.4: U-net architecture in the last step of the proposed pipeline. Numbers over feature channel blocks indicate the number of channels. Max pooling halves the size of each dimension, whereas transposed convolution with stride two doubles the size of each dimension. The ReLU is chosen as an activation function after the staple 3x3x3 convolutional layers.

the projection matrix, $\mathbf{A}$, on some parameterization of the acquisition orbit Ω is identified. Classic inversion approaches often seek to find the pseudo-inverse:

$$\mu = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{y}. \tag{3.4}$$

The pseudo-inverse has the advantage that solutions can be found for non-square and rank-deficient $\mathbf{A}$ that are possible for arbitrary trajectories.

It is noted that $\mathbf{A}^T$ represents a backprojection operation. Thus, the operator $(\mathbf{A}^T\mathbf{A})^{-1}$ represents a kind of generalized filtering operation. In fact, under idealized imaging conditions (parallel beam, sufficient sampling, etc.) and a circular acquisition geometry, $(\mathbf{A}^T\mathbf{A})$ represents the operator that applies the well-known intrinsic response of tomography - a $1/r$ blur function. Thus, in the ideal case, $(\mathbf{A}^T\mathbf{A})^{-1}$ is the inverse filter that removes $1/r$ blur. For non-circular orbits, divergent beams, etc., the blur induced by $(\mathbf{A}^T\mathbf{A})$ is not generally shift-invariant nor of the form $1/r$.

Motivated by the observations above, a novel reconstruction pipeline is proposed leveraging the speed and performance of CNNs and combining it with what is already well known about the required reconstruction process. Specifically, the backprojection step is maintained and the operator $(\mathbf{A}^T\mathbf{A})^{-1}$ is addressed. While one could develop a CNN to learn this inverse transformation directly, there is a way to provide a better network input. It is assumed that the geometry, and therefore $\mathbf{A}$ and $\mathbf{A}^T$ of the considered non-circular acquisitions, are known *a priori*. For this

reason, network inputs are devised to leverage this prior information. In the case of a shift-invariant imaging system, the system response, $\mathbf{A}^T\mathbf{A}$, can be deconvolved from the backprojection, $\mathbf{A}^T y$; with this deconvolution the dependence on the acquisition geometry can be effectively removed in the reconstruction process. There are, however, several deficiencies (deviations from ideal conditions) in the proposed reconstruction scheme, including but not limited to: Voxel sampling effects during backprojection, divergent beam geometry, redundancies in the projection data, shift-variant acquisition geometry. In addition, the proposed deconvolution procedure is noise amplifying and prone to artifacts. For these reasons, a post-deconvolution CNN is deployed to remove any residual artifacts.

To implement the deconvolution with the system response, it is assumed to be approximately shift-invariant, which is true for small objects and/or long geometries. The system response is approximated as $\mathbf{A}^T\mathbf{A}e_j$, where e_j denotes an impulse at the center of the image. In a first step, $\mathbf{A}^T\mathbf{A}e_j$ is deconvolved from $\mathbf{A}^T y$ via direct Fourier inversion, *i.e.*:

$$\mathcal{F}^{-1}\left\{\frac{\mathcal{F}\left\{\mathbf{A}^T y\right\}}{\lfloor\mathcal{F}\left\{\mathbf{A}^T\mathbf{A}e_j\right\}\rfloor}\right\}. \tag{3.5}$$

However, a practical difficulty of using this method originates in the fact that the ideal deconvolution kernel needs an unbounded support [98]. This requires to increase the backprojection volume of both system response and backprojected measurements onto a support considerably larger than the support of the true object, which makes the computation less efficient. For this reason, several techniques are implemented to mitigate artifacts associated with the deconvolution process, they are detailed in Section 3.5.3. First, a regularization operation is used in the denominator of Equation 3.5 to avoid division by zero and to suppress high frequency noise.

Second, the backprojection volume is expanded to approximately four times the reconstruction volume to mitigate spurious frequencies as a result of the DFT of signals with discontinuities at the boundaries; this helps to better satisfy the periodic boundary conditions necessary for fast Fourier transforms free of spurious frequencies. This is done in one of two ways: either by increasing the backprojection volume during the calculation of the impulse response by a factor of approximately 64 (approx. quadrupling number of voxels per dimension to 511x511x511), or by increasing the backprojection volume by a factor of 8 (doubling number of voxels per dimension to 256x256x256) and subsequently zero-padding the impulse response to yield another volume increase by factor of approximately eight (approx. doubling the number of voxels per dimension to 511x511x511). The uneven image output sizes are chosen to center the impulse response. Third, to mitigate artifacts in $\mathcal{F}\left\{\mathbf{A}^T\mathbf{A}e_j\right\}$ due to the combined effect of voxel sampling and ray-based projection operator, $\mathbf{A}^T\mathbf{A}e_j$ is computed at eight voxel locations around the central voxel of the image and the responses averaged.

After the deconvolution, the image is corrected for the sampling density by performing an element-wise division of the volume by $\mathbf{A}^T\mathbf{A}\mathbf{1}$, where $\mathbf{1}$ denotes a volume
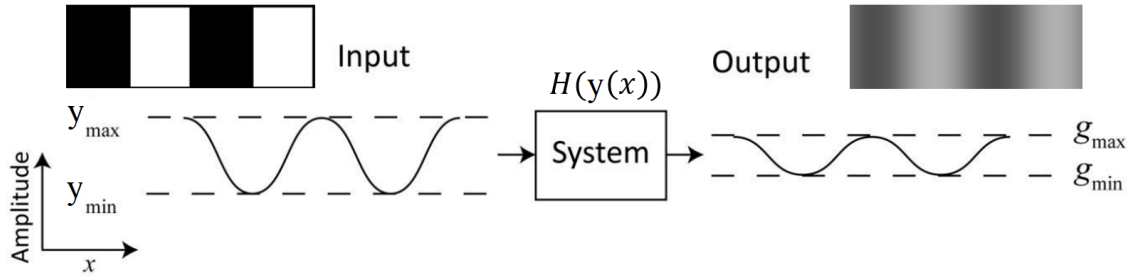
Figure 3.5: Visualization of the MTF, in this case the modulation of a periodic, one-dimensional signal composed of only one frequency. The input line pairs shown above the function can be understood as a combination of multiple of such periodic basis functions. After processing through the imaging system, understood as the mapping function $H : y(x) \rightarrow g(x)$, the amplitude of the periodic function is suppressed; the value of the suppression can be obtained exactly by a linear weighting with the value of the MTF at this frequency. As a result, the output line pairs are blurred. Figure adapted from [99].

of 1s. Finally, the image is truncated to the same size as the reconstruction image volume to save memory. The resulting image volume is used as input to the CNN. In summary, the input x to the CNN is represented mathematically as:

$$\mathbf{x} = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\{\mathbf{A}^T y\}}{\lfloor \mathcal{F}\{\mathbf{A}^T \mathbf{A} e_j\} \rfloor} \right\} \frac{1}{\mathbf{A}^T \mathbf{A} \mathbf{1}}. \tag{3.6}$$

For the CNN processing step, a U-Net architecture consisting of seven convolutional blocks, illustrated in Figure 3.4, is chosen. The U-net architecture is chosen due to its successful application in image deconvolution and CT reconstruction. The network is trained to predict the GT phantom images. The Adam optimizer with a learning rate of 0.001 is used and the root-mean-square error (RMSE) (Section 3.4.2) calculated between network output and GT phantom is chosen as a loss function.

3.4 Evaluation Metrics

This section details several image quality evaluation metrics used throughout the experiments of this thesis. The section is divided into two subsections: Section 3.4.1 details the metrics of absolute image quality, measures that do not need reference images; Section 3.4.2 introduces the relational metrics, which are calculated with respect to a reference image.

3.4.1 Absolute Metrics

The following metrics are calculated in CT images without the need for any reference image.

Modulation Transfer Function (MTF)

The MTF is commonly considered one of the most important measures to quantify spatial resolution of imaging systems. The MTF $T(f)$ is a measure of how an

imaging system (corresponding to a mapping function) H suppresses the incoming frequencies of a spatial signal $y(x)$ in the resulting measurement $g(x)$. When considering a periodic, 1D spatial signal defined by only one frequency f_1, such as $y(x) = A\sin(2\pi f_1 x) + B$ the so-called modulation in the original signal m_y and in the measurement m_g are defined as:

$$m_y = \frac{y_{max} - y_{min}}{y_{max} + y_{min}} = \frac{A}{B}$$
$$m_g = \frac{g_{max} - g_{min}}{g_{max} + g_{min}} = \frac{A}{B}\frac{H(f_1)}{H(0)} = \frac{A}{B} \cdot T(f_1), \tag{3.7}$$

where $y_{max} = B + A$ and $y_{min} = B - A$. In Fourier domain, the MTF can be understood as the linear scaling of each signal frequency component. This behavior is illustrated in Figure 3.5. The MTF is commonly understood as a universal image quality metric to characterize the performance of an imaging system, because it is calculated with respect to a specific imaging task [31]. On one hand, the MTF can be theoretically derived, for instance by performing a comprehensive cascaded systems analysis for CT imaging systems [100],[101]. On the other hand, the MTF can be experimentally determined by measuring either the point-spread function (PSF), LSF or edge-spread function (ESF) [102],[103],[104]. In most cases, oversampling the LSF or ESF is performed to prevent aliasing artifacts, which arise due to the finite sampling scheme in digital imaging systems [105].

In this thesis, the one-dimensional MTF in x-direction is obtained using the slanted tungsten wire phantom described in Section 3.2.2, in this case the wire is slanted in x-direction. In a first step, the center of mass (COM) location of the wire (*i.e.* voxel with maximum attenuation) in y-dimension is determined in every z-slice incorporating wire. This is done by cropping the image volume to yield a cuboid volume of interest (VOI) of minimum size, which still contains all of the wire. Afterwards, each attenuation voxel μ_i in the VOI is weighted with their respective y-position y_i and the COM in y-direction is calculated according to:

$$\text{COM}_y(\mu_i, y_i) = \frac{\sum_i \mu_i y_i}{\sum_i y_i}. \tag{3.8}$$

In the second step, the LSF in this z-slice is obtained by only considering the voxels in the row of the COM in y-dimension. Subsequently, a Gaussian peak is fitted to the LSF to obtain the COM in x-direction. An alternative way to calculate the COM in x-direction - instead of fitting a Gaussian peak - the weighting method used in y-dimension described above can be utilized for the x-dimension. In the third step, a linear regression based on the minimization of least-squares is performed on the COM positions in x dimension COM_x to obtain the linear equation $f_x(z)$ that describes the linear path of the tungsten wire. In addition, the quality of the linear regression is determined using the coefficient of determination R^2:

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}, \tag{3.9}$$

where the residual sum of squares is calculated as

$$SS_{\text{res}} = \sum_i (\text{COM}_{x,i} - f_{x,i})^2, \tag{3.10}$$

and the total sum of squares as

$$SS_{\text{tot}} = \sum_i (\text{COM}_{x,i} - \overline{\text{COM}}_x^2).$$ (3.11)

$\overline{\text{COM}}_x$ denotes the mean of all COM_x. In the fourth step, the LSFs of every slice are arranged in one vector based on their COM_x determined by the linear fit. In the last step, the oversampled LSF is Fourier transformed and divided by its integral to yield $\text{MTF}(f_x)(0) = 1$:

$$\text{MTF}(f_x) = \frac{\mathcal{F}\{\text{LSF}(x)\}}{\int \text{LSF}(x)\mathrm{dx}}.$$ (3.12)

Finally, the full width at half maximum (FWHM) in units of mm^{-1} is calculated to assign a scalar image quality metric to the calculated MTF. This is done by inferring the frequency at which the MTF has a value of 0.5. Because this point might not necessarily be covered by a value of the MTF, a cubic spline interpolation is performed on the MTF frequency.

Noise Power Spectrum (NPS)

Similarly to the MTF, the NPS is a Fourier metric and can either be theoretically derived or experimentally determined. To analyze any distortion (Poisson statistics of photon absorption, scattered photons, electronic jitter, etc.) of the original signal, *i.e.* noise, and its frequency content, the NPS is utilized. It is defined as the Fourier transform of the auto-covariance function, which is obtained by removing the signal mean from the auto-correlation [99]. In other words, the noise texture, which is characteristic for certain CT acquisition protocols (circular acquisition geometry, filtered backprojection based reconstruction), can be analyzed. It is considered to provide the most comprehensive approach to analyzing CT image noise. Typically, the 3D NPS is determined experimentally by Fourier transforming mean-detrended noise-only realizations, *i.e.* image regions of a homogeneous signal such as water or polymethyl methacrylate [104].

In this thesis, the radial NPS is calculated in patient MDCT scans, under the assumption of radial symmetry. A few adaptions to the usual approach have to be made. In the case of patient CT scans, the noise texture can be evaluated using an estimation of the NPS. For this purpose, a refined version of the approach from Walek *et al.* [106] is proposed here. In a first step, liver segmentation maps are obtained by applying a dedicated segmentation network [107]. The remaining non-stochastic information in the resulting axial slices I_i is removed by subtracting the respective Gaussian filtered slice I^G. The detrended slices are Fourier transformed using DFT and averaged afterwards to assess the axial NPS:

$$\text{NPS}_{\text{ax}} = b_x b_y \left\langle \frac{1}{L_i^G} |\text{DFT}(I_i - I_i^G)|^2 \right\rangle_i.$$ (3.13)

The NPS is normalized globally using the pixel spacing $b_{x,y}$ and slice-wise by the number of non-zero pixels in the Gaussian filtered slice L^G. Subsequently, the axial NPS is radially averaged to yield the radial NPS (Figure 4.2).

Noise Magnitude (NM)

The noise magnitude (NM) of CT images is usually quantified by calculating the standard deviation σ (in Hounsfield Units) in noise-only realizations. In this thesis, however, there is interest in quantifying the NM in patient CT scans. For this reason, the NM is calculated in the liver I_{liver}, which is segmented using a CNN:

$$\text{NM} = \sigma(I_{\text{liver}}). \tag{3.14}$$

NPS Maximum-Maximum Ratio (NMMR)

Since CNNs in general, and generative adversarial networks in particular, are prone to generate grid structures in the output images, a metric to quantify the extent of such grid structures is introduced. To achieve this, the extent of global maxima in the axial NPS (presumably caused by grid structures) are quantified using the NPS maximum-maximum ratio (NMMR). The maximum value of the axial NPS NPS_{ax} is divided by the maximum of the radial NPS NPS_{rad}:

$$\text{NMMR} = \frac{\max(\text{NPS}_{ax})}{\max(\text{NPS}_{rad})} \tag{3.15}$$

3.4.2   Relational Metrics

The following metrics are obtained by comparing the image to be evaluated, here abbreviated as the studied image s, to a reference image r.

Mean Absolute Error (MAE)

The mean absolute error (MAE) is calculated by voxel-wise subtraction of r from s and subsequent averaging:

$$\text{MAE}(s, r) = \frac{\sum\limits_{i=1}^{n} |r_i - s_i|}{n}, \tag{3.16}$$

where n indicates the overall voxel count.

Root-Mean-Square Error (RMSE)

Similarly to the MAE, the RMSE is calculated via:

$$\text{RMSE}(s, r) = \sqrt{\frac{\sum\limits_{i=1}^{n} (r_i - s_i)^2}{n}}, \tag{3.17}$$

Normalized Root Mean Squared Error (nRMSE)

The normalized root-mean-square error (nRMSE) is obtained by normalizing the RMSE with the mean of all attenuation values of the reference image $\bar{s}$:

$$\text{nRMSE} = \frac{\text{RMSE}}{\bar{s}} \tag{3.18}$$

Structural Similarity Index Measure (SSIM)

Wang *et al.* introduced the structural similarity index measure (SSIM) to quantify the structural similarity of two images s, r [108]. The SSIM considers luminance, contrast as well as structure and can be calculated as follows:

$$\text{SSIM}(s, r) = \frac{(2\mu_s\mu_r + T_1)(2\operatorname{cov}(s, r) + T_2)}{(\mu_s^2 + \mu_r^2 + T_1)(\sigma_s^2 + \sigma_r^2 + T_2)}, \tag{3.19}$$

where $\mu_{s,r}$, $\sigma_{s,r}$ and $\operatorname{cov}(s, r)$ correspond to the respective mean, variance and covariance values of the images. The variables T_i include the dynamical range of the pixel values.

Feature Similarity Index Measure (FSIM)

Zhang *et al.* introduced the feature similarity index measure (FSIM) in order to quantify the distortion of important low level features [109]. The metric is based on the identification of low level features using phase congruency (PC) and gradient magnitude (GM) maps. The subsequent calculation of their similarity $(S_{\text{PC}}, S_{\text{GM}})$, is followed by their combination reusing the PC as a weighting function:

$$\text{FSIM}(s, r) = \frac{\sum\limits_{s,r \in \Omega} S(S_{\text{PC}}, S_{\text{GM}}) \cdot \text{PC}_m(s, r)}{\sum\limits_{s,r \in \Omega} \text{PC}_m(s, r)}, \tag{3.20}$$

where Ω corresponds to the whole image spatial domain and $\text{PC}_m(s, r) = \max[\text{PC}(s), \text{PC}(r)]$.

Edge Generation Ratio (EGR)

Since the investigated generative adversarial networks are prone to randomly generate structures in the output CT images, the extent of additionally produced edges is explored by the introduction of the edge generation ratio (EGR). For this purpose, binary edge maps (EM) are extracted from the studied image and the reference image using the Canny algorithm [110]. Subsequently, the ratio of edge content is calculated:

$$\text{EGR}(s, r) = \frac{num(\text{EM}_s)}{num(\text{EM}_r)}. \tag{3.21}$$

Edge Preservation Ratio (EPR)

The edge preservation ratio (EPR) was proposed by Chen *et al.* to quantify the extent of edges preserved after deblurring [111]. It is obtained by calculating the intersection of the binary EM of both reference and studied image:

$$\text{EPR}(s, r) = \frac{num(\text{EM}_s \bigcap \text{EM}_r)}{(\text{EM}_r)} \tag{3.22}$$

Chen *et al.* also concluded, that the Canny algorithm is best suited to extract the binary EM in order to calculate the EPR.

Ratio of Corrupted Background Voxel (RCBV)

Since the investigated GAN are prone to randomly generate structures in the output CT images, a new metric to analyze the extent of artificially generated structures is introduced in this thesis. It is supposed to quantify the ratio of corrupted background voxel (RCBV), *i.e.* the number of voxels outside the patient contour that exhibit attenuation values substantially above background level. This is done by selecting an acceptance threshold according to 3σ of the background noise in the reference CTs (100 HU):

$$\text{RCBV}(s, r) = \frac{num(\mu_s > -900\,\text{HU})}{num(\mu_r = -1000\,\text{HU})}. \tag{3.23}$$

MTF Mean Percentage Error (MPE)

To compare the calculated MTF to a reference, the mean percentage error (MPE) is calculated by subtracting the cubic spline interpolation of the studied MTF $\text{MTF}_{s,\text{interp}}$ from the reference MTF $\text{MTF}_{r,\text{interp}}$ according to:

$$\text{MPE}(s, r) = 100 \cdot \frac{\text{MTF}_{s,\text{interp}} - \text{MTF}_{r,\text{interp}}}{\text{MTF}_{r,\text{interp}}} \tag{3.24}$$

NPS Correlation Coefficient (NCC)

In order to compare the NPS of a simulated CT image to a reference, the NPS correlation coefficient (NCC) is introduced here. It is supposed to evaluate if the noise texture in the simulated CT image follows the characteristic NPS in a reference data set. For this reason the radial NPS NPS_s (calculated as described in Section 3.4.1) is compared to the mean of the NPS array of the reference data set $\text{NPS}_{r,\text{mean}}$ in terms of monotonic Pearson correlation, further called NCC:

$$\text{NCC}(s, r) = \frac{\text{cov}(\text{NPS}_s, \text{NPS}_{r,\text{mean}})}{\sigma_{\text{NPS}_s}\sigma_{\text{NPS}_{r,\text{mean}}}}, \tag{3.25}$$

where $\text{cov}(\cdot)$ indicates the covariance

$$\text{cov}(s, r) = \mathbb{E}[(\text{NPS}_s - \mu_{\text{NPS}_s})(\text{NPS}_{r,\text{mean}} - \mu_{\text{NPS}_{r,\text{mean}}})] \tag{3.26}$$

and σ denotes the standard deviation, $\mathbb{E}$ the expectation value and μ the mean value.

NPS Mean Percentage Error (NMPE)

To compare the calculated NPS_s to the reference NPS data, the NPS mean percentage error (NMPE) is calculated according to:

$$\text{NMPE}(s, r) = \frac{100}{n \cdot \max(\text{NPS}_{r,\text{mean}})} \sum_{i=1}^{n} |\text{NPS}_s - \text{NPS}_{r,\text{mean}}|, \tag{3.27}$$

where n is the total number of NPS values.
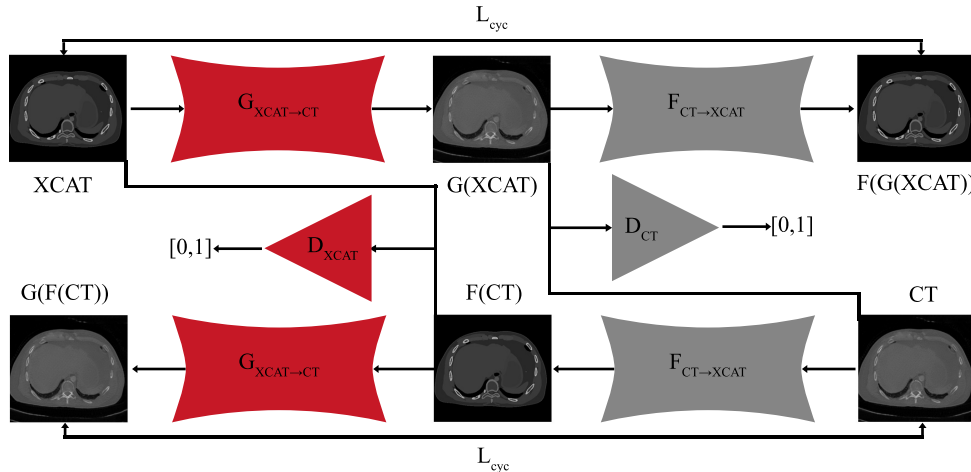
Figure 3.6: CycleGAN network architecture: The generators $G_{\text{XCAT}\rightarrow\text{CT}}$ and $F_{\text{CT}\rightarrow\text{XCAT}}$ map images from the XCAT domain to the CT domain and vice versa. Two additional adversarial networks D_{XCAT} and D_{CT}, called discriminators, are trained to distinguish between real and synthetic images: XCAT and $F(\text{CT})$ are distinguished by discriminator network D_{XCAT} and CT and $G(\text{XCAT})$ are distinguished by discriminator network D_{CT}. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature International Journal of Computer Assisted Radiology and Surgery ©2019.

Dice Similarity Coefficient (DSC)

The Dice similarity coefficient (DSC) is used to compute the accuracy of a predicted segmentation map segm_s by comparing it to a GT segmentation map segm_r. The DSC is defined as the ratio of the doubled number of intersection voxels divided by the number of the total segmentation voxels in both segmentations combined:

$$\text{DSC}(s, r) = \frac{2 \cdot num(\text{segm}_s \bigcap \text{segm}_r)}{num(\text{segm}_s) + num(\text{segm}_r)} \tag{3.28}$$

3.5 Experimental Design

This section is divided into three parts. Section 3.5.1 explains the experiments conducted to investigate CycleGAN-based CT Synthesis. Section 3.5.2 outlines the simulation frameworks, experimental setting and evaluation performed to prospectively optimize CBCT acquisition orbits. Section 3.5.3 details how the reconstruction pipeline proposed in Section 3.3 is applied to arbitrary CBCT orbit data and subsequently evaluated.

3.5.1 CT Synthesis Using CycleGAN

In this subsection, the CT synthesis experiments using CycleGAN are explained. First, the training routine, the utilized data as well as the applied evaluation metrics are described. The subsequent explanation of experiments is divided into three studies. In an initial, broader configuration study with reduced training and testing

data set size, the ideal CycleGAN network configuration for CT synthesis is investigated. The optimized configuration is subsequently utilized for a task-based study with an extended training data set as well as a task-based loss function. In the last step, the synthetic images from the task-based study are utilized to train blood vessel segmentation networks as a form of data augmentation in a proof-of-principle experiment. The last study is performed to investigate alternative uses for the generated CT images, besides the optimization of CBCT acquisition orbits.

Training and Evaluation

Image patches with 256×256 pixels are used for the training of the networks. In contrast to the original CycleGAN [51], the image translation from the XCAT to the CT domain is emphasized. Therefore, multiple adjacent slices in the XCAT domain are used, but the number of slices in the CT domain is reduced to one. For each axial slice, one random patch is extracted. The possible patch centers are sampled from a uniform distribution inside the body contour of the image to avoid too many background samples in the training data. Data augmentation is performed by randomly rotating the slices between -20° and +20° followed by a bi-linear interpolation. Most of the training configurations of the CycleGAN are adopted from Zhu *et al.* [51], see Section 3.1.3. Deviating from the original network, the deconvolution operations in the generator network are substituted by up-sampling operations in order to avoid checkerboard artifacts [112]. The Python-based TensorFlow library, setup on Keras, is used to implement the networks. For the implementation, an already developed software framework was adapted [113]. Training is carried out on an Nvidia Titan XP (Nvidia, Santa Clara, CA, USA).

All used data sets (in-house, LiTS, IRCAD, see Section 3.2.1) contain varying voxel spacing, spatial orientation and scan parameters. For this reason, all images are resampled to a voxel spacing of $0.75 \times 0.75 \times 1.5 \, \mathrm{mm}^3$. Aside from the variations in the voxel spacing, the magnitude of the contrast-enhancement fluctuates between patients and vessel branches. These fluctuations are due to the time dependent progression of the contrast agent through the blood vessels of the patient, which is not coherently distributed in every CT image. As input for the generative network, 120 XCAT abdomen phantoms with a tube voltage $U \in [70, 80, 90] \, \mathrm{kVp}$, a voxel spacing of $0.75 \times 0.75 \times 1.5 \, \mathrm{mm}^3$, and varying anatomical parameters are simulated. The intensity values of all input data are normalized from $[-1024, 2000]$ HU to $[-1, 1]$.

Each trained network is evaluated by comparing a set of evaluation metrics between the generated CT images and either the original XCAT input or real patient CT images. The RCBV, MAE, SSIM, FSIM, EPR and EGR are computed with respect to the corresponding XCAT volume. NCC, NMPE, NMMR are calculated with respect to the mean values of real CT images.

Configuration Study

The network architecture of the CycleGAN is visualized in Figure 3.6. Three parameter variations are investigated and the eight possible configurations are compared.
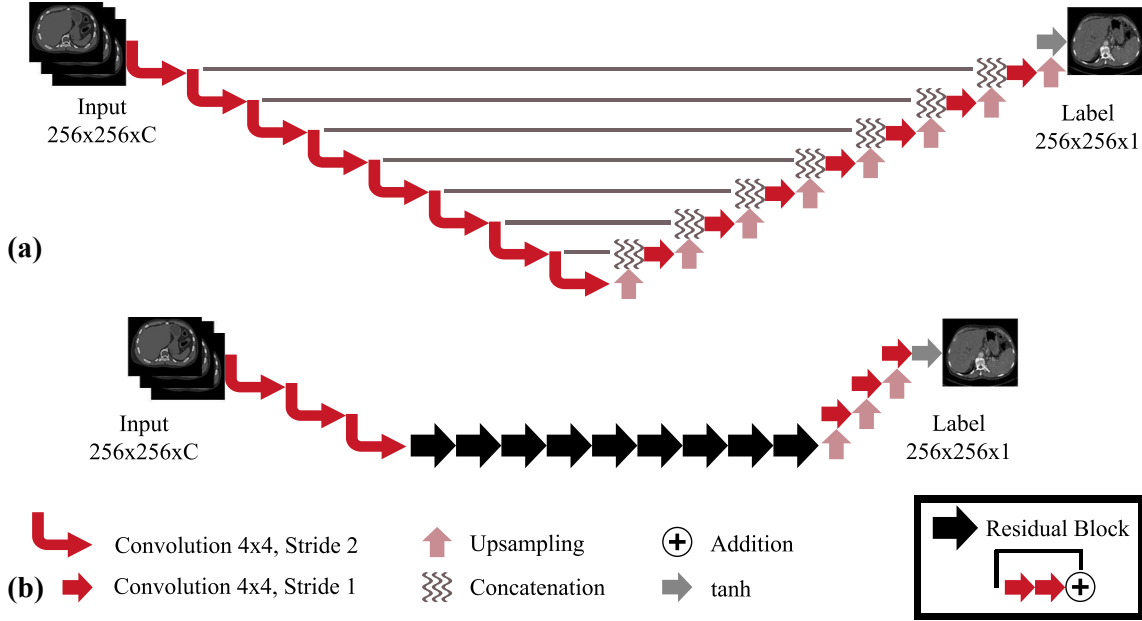
Figure 3.7: a U-Net and **b** Res-Net generator architectures. In this case, the U-Net architecture consists of 16 convolutional layers, eight encoding and eight decoding. The Res-Net consists of three encoding and three decoding convolutional layers and nine residual blocks. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature International Journal of Computer Assisted Radiology and Surgery ©2019

As a first variation, a Res-Net architecture is compared to a U-Net generator architecture (Figure 3.7). Secondly, the influence of the number of input slices $C \in 3, 5$ is explored. Finally, the effect of pre-processing the CT data with a body contour mask is tested. In all cases, the networks are trained using samples obtained from six XCAT abdomen phantoms (Section 3.2.2) and six patient CT scans originating from the in-house CT data (Section 3.2.1). The testing data set consisted of six additional XCAT phantoms.

Task-Based Study

For the task-based study, the training data set is increased to 60 XCAT phantoms and 60 patient CT images. The testing data set consists of 60 additional XCAT phantoms. The patient CT training data set is composed of in-house CT, IRCAD and LiTS data. Additionally, a regularization term to enforce HU-consistency in selected image regions is added to the standard loss function [114]. In particular, the adaption penalizes deviations between the original XCAT and the generated CT image in the contrast-enhanced vessels as well as deviations in both image domains in the background region. These deviations are determined using dedicated segmentation masks M_{ves} and M_{bg} for vessels and background, respectively. The additional regularization term is implemented as an HU-consistency loss using the L1-norm:

$$L_{HU} = \lambda_{bg}||(G(x) - x) \cdot M_{bg}||_1 + \lambda_{bg}||(F(y) - y) \cdot M_{bg}||_1$$
$$+ \lambda_{ves}||(G(x) - x) \cdot M_{ves}||_1. \tag{3.29}$$

The aim of this study is to investigate the influence of the proposed loss and the incorporation of public source CT images on the synthesis performance. In total, three

task-based networks are trained, one with the full training data set but without the HU-consistency loss (further called TASKNET1), the second with the full training data set and with the HU-consistency loss (TASKNET2), and the third with only in-house CT images and with the HU-consistency loss (TASKNET3).

Segmentation Study

For the vessel segmentation, three networks are trained on different data configurations: the first on 15 synthetic CTs from the best-performing network of the task-based study, the second using 15 real CT images acquired in-house, and the third on a combination of both sources adding up to 30 CT images. The data set size for the combined training is increased to incorporate this particular advantage of the generated data as a means for data augmentation. Each segmentation model is tested on the same five in-house CT images. The GT for the in-house CT images is determined semi-automatically by comparing various stages of contrast-enhancement followed by further processing steps. This is possible due to the availability of a time series for each in-house CT. For the synthetic CT images, the annotation masks of the corresponding XCAT phantom are used as GT. In particular, a 2.5D U-Net architecture with residual connections is deployed [107] and the Tversky loss function with $\alpha = 0.3$ is used. Image patches of 256×256 pixels are extracted from the axial plane for training and a three-slice input is used to predict the center slice. The patches are randomly rotated between -15° and +15° for data augmentation. Intensity values are normalized from $[-200, 1000]$ HU to $[-1, 1]$. The L1 regularization and batch normalization with a batch size of 16 is used. Each network is trained for 100 epochs. Eventually, the segmentation performance is evaluated using the DSC.

3.5.2 Orbit Optimization

This subsection describes the experiments performed for CBCT orbit optimization. The goal of these experiments is - under consideration of a specific imaging scenario in the intervention room - to find the better of two alternative source orbits to a standard circular trajectory: a tilted circular orbit consisting of a 360° rotation and a simultaneous 40° elevation angle change as well as a triple-arc geometry consisting of two 180° arcs and one orthogonal 40° arc. A visualization of the investigated orbit geometries is shown in Figure 3.8. In the specific imaging scenario considered here, the right-back-top octant is considered blocked space, i.e. a standard circular source orbit is not feasible. The experiment geometries such as size of the detector, magnification, etc., are chosen to imitate the imaging setting of a Siemens artis zeego device (Siemens Healthineers, Forchheim, Germany).

The subsection starts with a brief explanation of the GATE software used for MC projection simulation. Afterwards, the reconstruction and evaluation processes are detailed.

MC Simulations Using GATE

The following paragraphs are intended to give a brief overview on GATE and its code basis, GEANT4. For further information, the reader is referred to the associated publications [41],[43]. In its essence, GATE is a script language providing
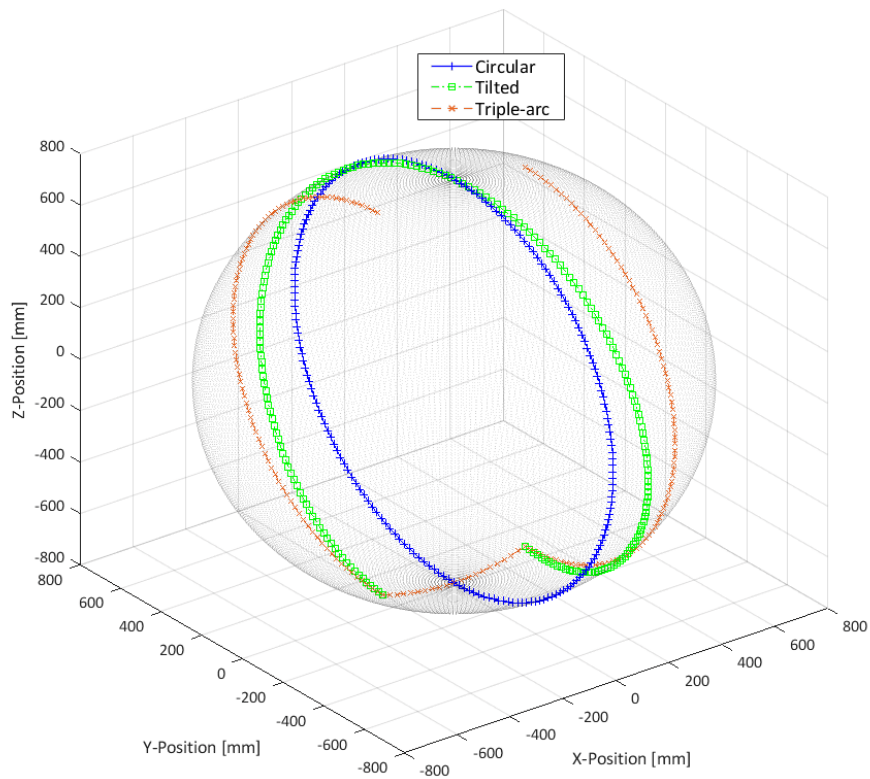
Figure 3.8: Visualization of the considered CBCT source orbit geometries. The image isocenter is at position (0,0,0). Alternative orbits to a standard circular geometry are investigated. In this case, the available actuation space is limited (due to personnel or interventional equipment). In this specific view, the right-back-top octant is considered to be blocked space, preventing the realization of a circular orbit. Instead, a tilted circular orbit with a maximum tilt angle of 20° relative to a circular orbit and a triple-arc orbit consisting of two 180° arcs and one orthogonal 40° arc on the bottom of the point sphere are evaluated as potential alternatives.
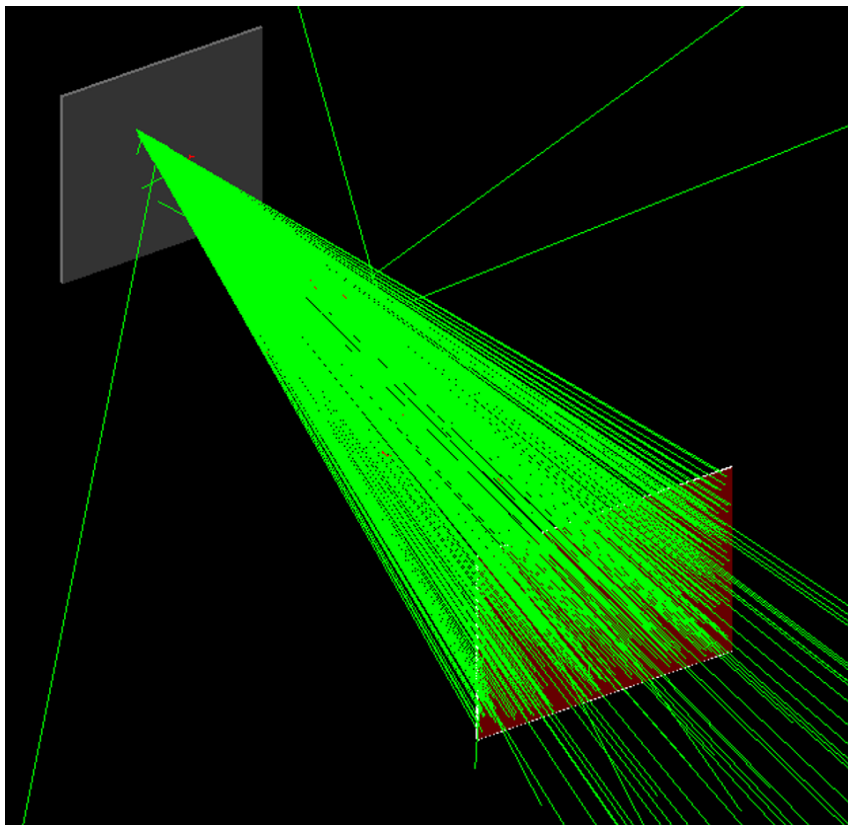
Figure 3.9: Simulation setup in GATE with the path of 1000 simulated photons visualized (green). The X-ray source produces a photon beam with circular base area. With the use of a lead collimator (grey), the beam shape is modulated into rectangular shape. The collimator has an opening, a trapezoid volume, which is designed in such a way, that the outermost photons hit the detector's (red) edge.

an easy-to-learn interface to create MC simulations (Section 2.3.1) suited for the needs of medical imaging. GATE is setup on the object-oriented C++ code of GEANT4, a toolkit that adapts the MC approach to interactions of particles with matter and provides particle detector models. GEANT4 also includes graphical drivers for visualization purposes. Due to the high computation time associated with GEANT4-based MC simulations, the Gatelab cluster is used for faster execution times [45],[46]. Gatelab is setup on the European Grid Infrastructure, which is a EU funded institution delivering computation services in the form of cloud computing and federated data centers.

Each GATE simulation follows a workflow consisting of multiple steps. First, a material database is setup, incorporating material-specific quantities such as the elemental composition of compounds, their density and state (solid, fluid, etc.). The second step consists of setting up the geometries, the world volume, the phantom, the collimator and the detector. For this, a detector panel consisting of 1240x960 Caesium Iodide (CsI) scintillator detectors of 0.5 mm thickness and 0.308 mm edge length is created. A source-axis distance (SAD) of 785 mm and a source-detector distance (SDD) of 1197 mm is chosen. The lead collimator of 5 mm thickness is placed within a 65 mm distance of the source. All parameters are chosen according to Siemens data on the artis zeego. A visualization of the simulation setup is illustrated in Figure 3.9. Besides the zeego model, a water cylinder phantom encompassing a slanted tungsten wire is created according to Section 3.2.2. Alternatively, an XCAT phantom is incorporated into the simulation framework to demonstrate the ability to simulate dose deposition maps depending on the source orbit geometry. In addition, a so-called dose-sensitive volume, a module provided by GATE, is attached to each XCAT phantom voxel; this enables the tracking of the imparted physical dose per voxel.

After the geometries are setup, the digitizer is initialized, a tool to mimic realistic photon detector properties. In this case, only an 'adder', 'thresholder' and a 'readout' module are used, which means that all simultaneous events in the detector are grouped together (given that they stem from the same original X-ray), and their deposited energy calculated and added up to produce a signal pulse, but only if the detected deposited energy is higher than the selected threshold of 10 keV. After the detector readout setup, the source position and spectrum are chosen. Here, an X-ray spectrum of 100kVp is simulated for a tungsten anode, an anode angle of 12° and 25 mm of aluminum pre-filtration, using the SpekCalc software [115],[116],[117]. Again, the parameters are chosen based on Siemens data on the artis zeego. The lower cutoff spectrum energy is set to 10 keV, which resembles exactly the value of the thresholder module in the detector model, and an energy bin size of 1 keV is selected. In the last step, the movement of the geometry according to the simulated source orbit is implemented by rotating the phantom.

Projection Simulation, Reconstruction and Evaluation

200 projections are simulated for each of the three investigated geometries. In total, 20 billion primary photons sampled from 100 keV spectrum are used per projection, amounting to 4 trillion primary photons in total. As an imaging phantom, the

slanted tungsten wire embedded in the water cylinder described in Section 3.2.2 is incorporated into the simulation in order to extract the MTF after reconstruction. The values of photons per projection are chosen based on the observed photon count number on the detector after traversing the maximum distance through the water cylinder; the photon counts are supposed to be of the order of several hundreds to guarantee sufficient contrast in the resulting CBCT image. Each simulated projection is normalized by dividing through a projection acquired with the same imaging settings but without the phantom in the beam line.

As an alternative method to MC simulations, an approach based on the projection operators described in Section 2.3.2 is investigated. The Astra toolbox is utilized to simulate CBCT projections acquired with arbitrary source orbits [67]. The Astra toolbox allows for the incorporation of the full projection matrix, which contains a set of vectors sufficient to unambiguously define the imaging geometry. The projection operator systematically sums up the attenuation values in each traversed image voxel of the ray incident on the detector pixels and returns the weighted sum as a result. In contrast to the MC simulations above, there is no need to normalize this data. The line attenuation values in the projections are translated into photon counts by using the Beer-Lambert law (Equation 2.13) and a bare beam fluence of 100000 photons per detector pixel. Subsequently, a noise model is applied by sampling each detector pixel value $N_{i,\text{noise}}$ from a Poisson distribution with mean λ equal to the respective photon count N_i and a probability mass function $f_{Poisson}$ given by:

$$f_{Poisson}(N_{i,\text{noise}} = n, N_i) = \frac{N_i^n e_i^{-N}}{n!} \tag{3.30}$$

and adding a random variable sampled from a normal distribution f_{Gauss} with mean μ and standard deviation σ equal to 0 and 7, respectively:

$$f_{Gauss}(n) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{n-\mu}{\sigma}\right)^2} \tag{3.31}$$

To compare raw fluence data simulated with the projection operators to the MC simulations, the fluence data simulated with the MC method has to be corrected with the factor $C(x,y)$ - for pixel counts in x and y direction with origin at the center of the detector - due the inverse square law according to:

$$C(x,y) = \frac{\text{SDD}^2 + \text{R}(x,y)^2}{\text{SDD}^2}, \tag{3.32}$$

where SDD denotes the source-detector distance and R the distance of the respective detector pixel from the detector origin calculated via:

$$\text{R}(x,y) = e_{det}\sqrt{x^2 + y^2}, \tag{3.33}$$

where e_{det} denotes the detector pixel edge length. After comparison of the fluence data, both data have to be log-transformed again, using their associated bare beam fluence, to carry out reconstruction.

After projection simulation, each set of normalized projections is reconstructed on a grid of 512x512x512 voxels with 0.3 mm cubic voxels. The SIRT algorithm (Section 2.3.3) is applied to the projection data for a total of 100 iterations using the CUDA supported implementation of the Astra toolbox. Finally, the evaluation of the reconstructed images is performed by calculating the FWHM and the MPE of the MTF, which is obtained using the LSF oversampling method described in Section 3.4.1.

3.5.3 Reconstruction Experiments

This subsection details the experiments conducted to train and validate the novel reconstruction pipeline proposed in Section 3.3. The subsection starts with a description of the parametrization leveraged to generate a general class of random orbits. Subsequently, the conducted experiments are divided into three parts. The first part investigates three different means of regularization to suppress artifacts in the deconvolution operation. The second part explains the initial series of experiments performed to study the performance of the proposed reconstruction scheme when trained with and applied to only one specific geometry at a time. The third part details the experiments conducted to research the pipeline's ability to reconstruct orbits it has not seen during the training procedure. Moreover, the complexity of the reconstruction task is substantially increased.

Orbit Parametrization

Reconstruction is exclusively performed on simulated data. The forward- and back-projection operators of the Astra toolbox are used with the option to incorporate the noise model described in Section 3.5.2. As an imaging geometry, a SAD of 1 m, a SDD of 1.5 m and a 256x256 pixel detector consisting of 0.75x0.75 mm pixel are chosen. The detector size is selected so that the projection data of the phantom, consisting of 128x128x128 cubic image voxels with 0.5 mm edge length, do not suffer from truncation. For the imaging geometry, 512 rotation angles, θ, are evenly distributed (in terms of rotation angle) between 0° and 360°. The elevation angles, ϕ, are parameterized as a linear combination of $K = 9$ sinusoidal basis functions:

$$\phi(\theta) = \sum_{j=1}^{K} \Omega_j b_j(\theta), b_j(\theta) = \begin{cases} \cos\left(\frac{j-1}{2}\theta\right), & \text{if } j \text{ is odd} \\ \sin\left(\frac{j}{2}\theta\right), & \text{if } j \text{ is even,} \end{cases} \tag{3.34}$$

with coefficients sampled from a uniform distribution $\Omega_j \in [-1, 1]$. The first four basis functions b are illustrated in Figure 3.10. An example of a random orbit generated using this scheme is shown in Figure 3.11. 25° are chosen as a maximum elevation angle according to the actuation space measured for a specific patient table position on a Siemens artis zeego, visualized in Figure 3.12.

Regularization Methods

Due to calculation issues inherent to the proposed reconstruction algorithm (explained in Section 3.3) three different means of regularization with the strength λ in Fourier domain and before the actual deconvolution operation are investigated. Parameter sweeps to determine the ideal λ are performed. In particular, the term

$$\lfloor \mathcal{F}\left\{\mathbf{A}^T \mathbf{A} e_j\right\} \rfloor \tag{3.35}$$
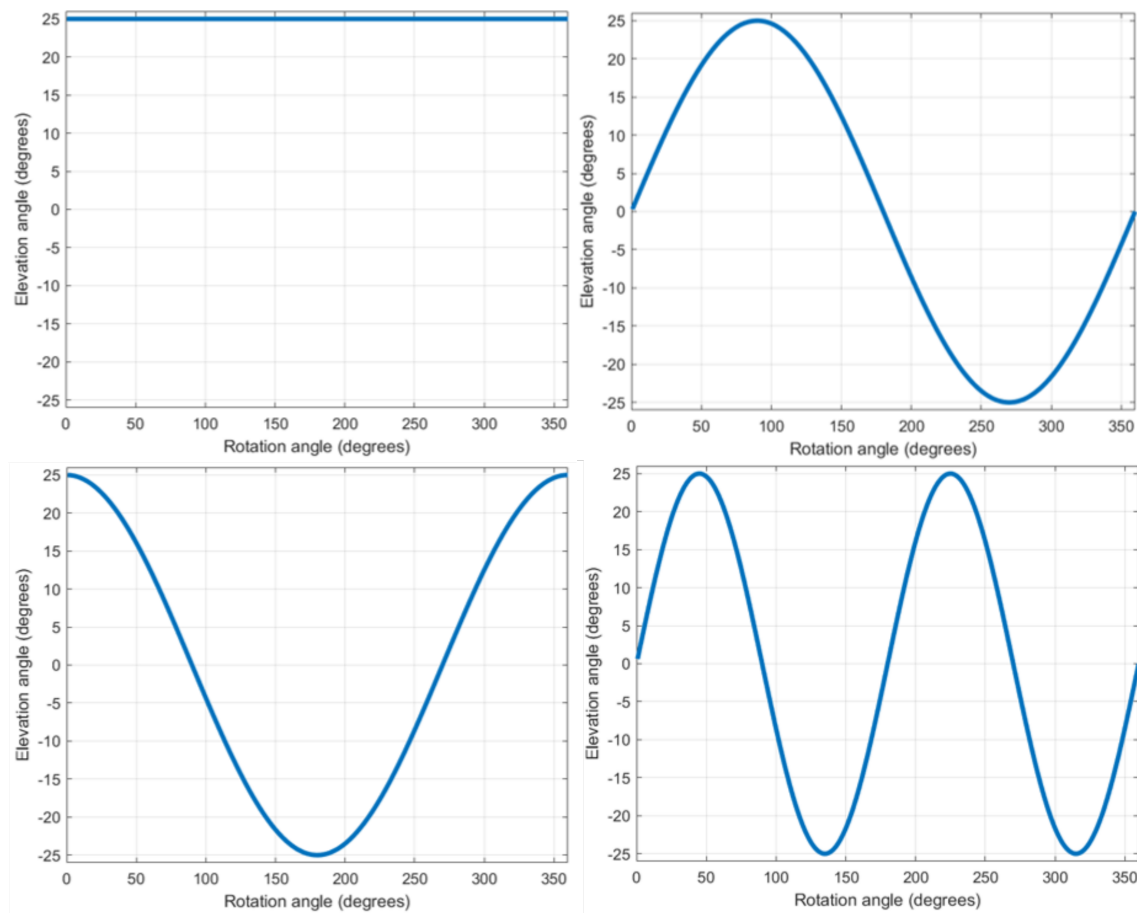
Figure 3.10: Periodic basis functions used for the parametrization of orbits. Top left shows the b_1, top right the b_2, bottom left b_3 and bottom right b_4 according to Equation 3.34.
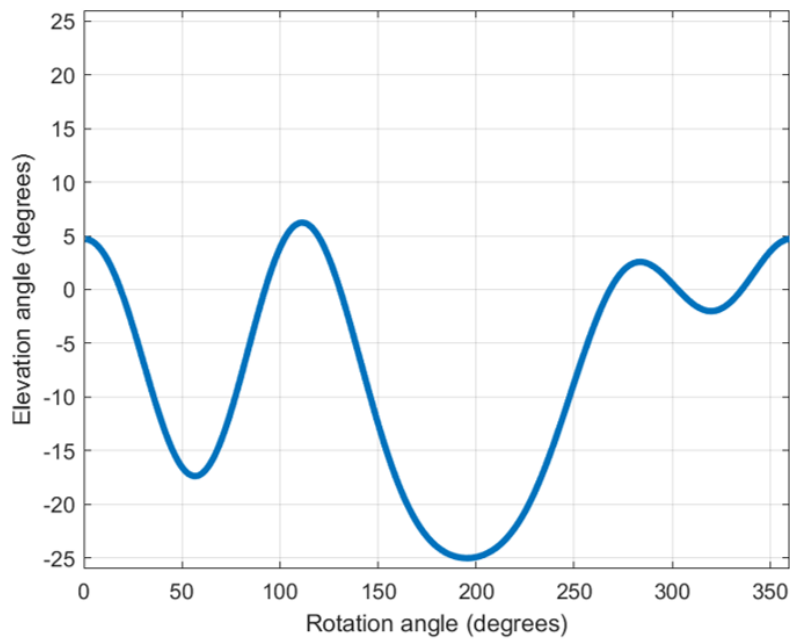
Figure 3.11: Random orbit geometry parametrized using Equation 3.34 and randomly generated coefficients: $\Omega_1 = -0.40$, $\Omega_2 = 0.02$, $\Omega_3 = 0.45$, $\Omega_4 = -0.42$, $\Omega_5 = -0.25$, $\Omega_6 = 0.07$, $\Omega_7 = 0.32$, $\Omega_8 = 0.15$, $\Omega_9 = 0.13$
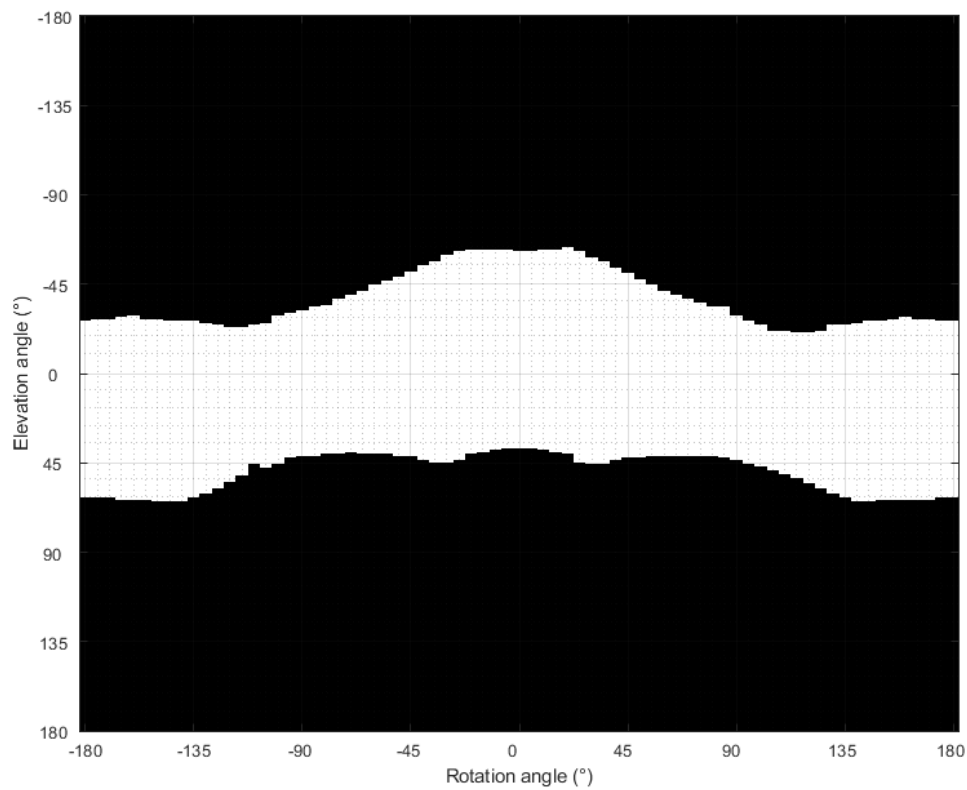
Figure 3.12: Experimentally determined actuation space available on the Siemens artis zeego. Black area indicates non-available, white area available actuation space.

is penalized. The first investigated methods adds a term that acts as a low pass filter and thereby penalizes high frequencies, as they are amplified by the deconvolution operation. The penalty term is obtained by generating a matrix the size of the impulse response (511x511x511) and calculating the quadratic penalty $R(\mu)$ via [21]:

$$R(\mu) = \lambda \mu^T \mathbf{R} \mu, \tag{3.36}$$

where μ denotes an image matrix and the matrix $\mathbf{R}$ denotes a matrix, which is defined to yield the penalization of the form:

$$R(\mu) = \frac{1}{2} \sum_i \sum_j w_{i,j}(\mu_i - \mu_j)^2, w_{i,j} = \begin{cases} 1, & \text{for the six nearest neighbors} \\ 0, & \text{otherwise.} \end{cases} \tag{3.37}$$

In the case of the impulse response, this penalization is formulated as:

$$\mathcal{F}\left\{\mathbf{A}^T \mathbf{A} e_j + \lambda \mathbf{R} e_j\right\}. \tag{3.38}$$

This can also be described by effectively setting the six nearest neighbor voxels to the center voxel to the value '-1' whilst setting the center voxel value to '+6'. The strength of the regularization is tuned with the parameter λ.

The second investigated method is a simple thresholding operation, where every absolute value (values possess an imaginary term) in Equation 3.35 that is lower than the threshold value λ is set to the value of λ, without the preservation of the imaginary part. The third investigated method is also composed of a thresholding operation, but instead of resetting the values with a modulus smaller than λ (μ_{res}), they are rescaled instead by the factor:

$$\frac{\lambda}{|\mu_{\text{res}}|} \tag{3.39}$$

to preserve the imaginary part.

Application to Single Geometries

For the application to single geometries, the thresholding regularization described above is performed. Initially, four networks are trained on data of only one orbit, while one network is trained on data of two different geometries in a common pool. This is done to investigate if our proposed approach is able to reconstruct data of more than one geometry. The five acquisition geometries are as follows:

- circular, $\phi = 0$ for all θ
- $\phi = \sin(2\theta)$,
- $\phi = \sin(3\theta)$,
- $\phi = \sin(2\theta)$ and $\phi = \sin(3\theta)$,
- one random orbit (generated as explained above).

For training, validation and testing, 1000 tetrahedral phantoms are procedurally generated without the addition of low frequency noise as described in Section 3.2.2. The projections are simulated without incorporating image noise. 800 of the resulting data is used for training, 100 for validation and 100 for testing. Each network is

trained for 100 epochs (*i.e.* 80000 steps). Training is carried out on an Nvidia V100 GPU (Nvidia, Santa Clara, CA, USA) provided by the bwForCluster computing cluster. The V100 has 16GB of dedicated GPU memory, which is necessary to carry out the training procedure due to its particularly high memory demand (3D model and resulting number of parameters take up GPU memory). The bwForCluster is a service provided by the bwHPC, a consortium of universities in the German state of Baden-Württemberg. The reconstruction pipeline is evaluated in terms of FSIM, SSIM and nRMSE (Section 3.4.2), calculated relative to GT. For comparison, an iterative reconstruction scheme, the SART described in Section 2.3.3, is also applied and quantitative performance measures (relative to truth) are computed.

Application to Random Geometries

In this paragraph, the reconstruction technique is improved by training the networks on a variety of randomized orbits to facilitate processing of new arbitrary orbits that are not used in CNN training. This is done by assigning a unique random orbit geometry to each generated phantom (1000 tetrahedron phantoms and 800 LIDC phantoms). For this study, the rescaling regularization of the Fourier term is chosen. Furthermore, the reconstruction pipeline is not only applied to tetrahedral phantoms, as in the previous experiments, but also to anthropomorphic data in training and evaluation. Further refinements of the simulated data are applied including low frequency textures (*i.e.*, Gaussian peaks) added to the tetrahedral phantoms as well as a realistic noise model for all simulated projections. Two networks are trained, distinguished by the ratio of epochs trained on LIDC data (see Section 3.2.1). While the first network is only trained and evaluated on the tetrahedron phantoms, a second network is trained on tetrahedral data (70 epochs and each epoch consists of 800 update steps), and subsequently fine-tuned (30 epochs/640 steps) and evaluated with anthropomorphic LIDC data.

In order to provide a baseline comparison with state-of-the-art algorithms, penalized-likelihood reconstruction (see Section 2.3.3) with a Huber penalty is utilized with 100 iterations, Huber penalty parameters of $\delta = 1$ and $\beta = 10^5$. For evaluation of the reconstruction performance, the nRMSE, the FSIM and the SSIM are calculated between the network output/MBIR reconstructions and the GT phantom.

4. Results

The experiments conducted below aim to find novel methods to implement interventional CBCT with non-circular acquisition orbits into clinical routine. The main focus is to investigate faster simulation and reconstruction schemes that still perform comparably to current standards. The investigation of different CT simulation methods aims to determine an appropriate method to prospectively optimize the acquisition geometry of arbitrary CBCT orbits depending on a specific interventional imaging tasks and restrictions of the robotic actuation range - all in a reasonable time frame. The reconstruction studies aim to provide insights on how the potential of CNNs can be exploited for faster reconstruction methods of arbitrary CBCT orbits. This chapter is divided into two parts: The first part starts with the results of the optimization studies consisting of three different methods of CT simulation: CycleGAN, MC approach, and projection operator; performed to investigate the prospective optimization of arbitrary CBCT acquisition orbits (Section 4.1). The second part represents the results of the novel CNN-based reconstruction method, consisting of a deconvolution operation with the shift-invariant system response and a posterior deployment of a CNN (Section 4.2). Parts of the results have been published in [85][1], [86], [87], [88] and [89] and the corresponding figures and descriptions are adapted and partly replicated thereof.

4.1 Simulation and Evaluation of CT Images

This section shows the results for the three methods investigated for CT simulation. In Section 4.1.1, a CycleGAN network architecture is utilized to perform a direct domain translation from images of digital body phantoms to CT images. In Section 4.1.2 two means of CBCT projection simulation are compared: MC method and projection operator. Subsequently, the simulated projections are used for reconstruction and evaluation.

4.1.1 CT Synthesis Using CycleGAN

This investigation is divided into two studies. First, a configuration study is performed to determine the ideal network configuration for CT simulation (Section 4.1.1).

---

[1]Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature International Journal of Computer Assisted Radiology and Surgery ©2019

Table 4.1: The averaged anatomical accuracy metrics (± standard deviation) using modified data as input to provide a baseline for the CT Synthesis evaluation framework

Metric	CT-CT	Rigid-Transfo.	Poisson
RCBV	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
MAE	228 ± 104	197 ± 139	30 ± 1
SSIM	0.57 ± 0.16	0.76 ± 0.07	0.92 ± 0.02
FSIM	0.55 ± 0.13	0.34 ± 0.05	0.98 ± 0.01
EPR	0.24 ± 0.21	0.09 ± 0.02	0.88 ± 0.02
EGR	1.05 ± 0.34	1.01 ± 0.13	1.87 ± 0.60

No Pre-Processing   Body Contour

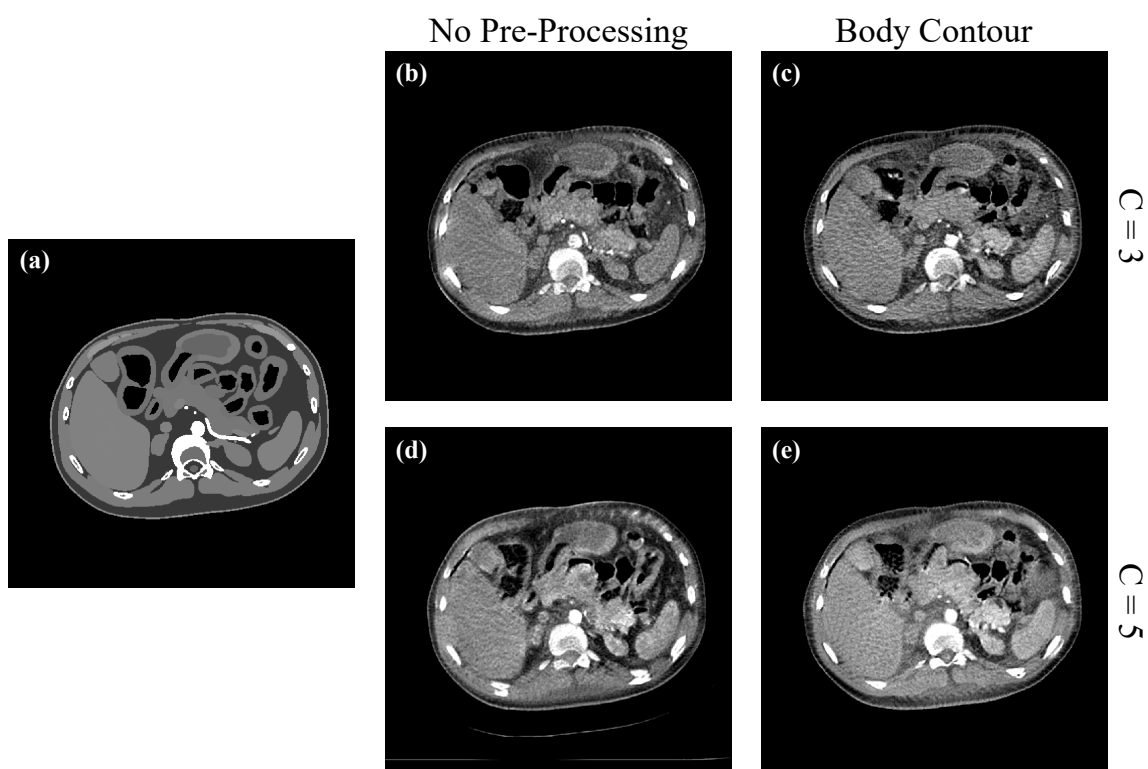**Figure 4.1:** **(b)**, **(c)**, **(d)** and **(e)** axial slices of the synthetic CT images generated with the Res-Nets trained in the configuration study (C $\widehat{=}$ number of input slices) and **(a)** the corresponding XCAT slice. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature International Journal of Computer Assisted Radiology and Surgery ©2019.

Figure 4.2: Noise power spectrum (NPS) estimation of a reference patient CT in the liver, **(a)** the axial NPS and **(b)** radially averaged NPS. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature International Journal of Computer Assisted Radiology and Surgery ©2019.
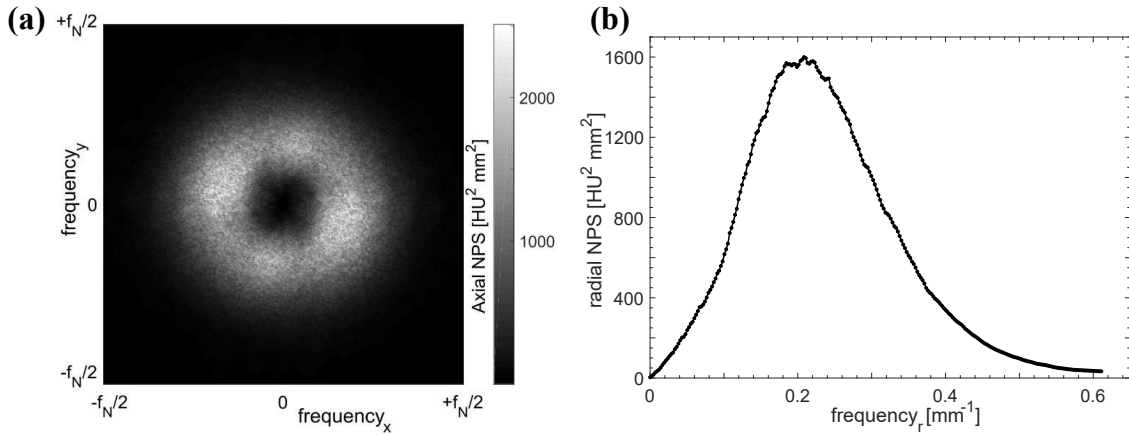
Secondly, the simulation approach is adapted to investigate task-based loss functions - specifically, for providing accurate HU in the blood vessels - for the training procedure and include public data sets to quantify changes in performance (Section 4.1.1). Evaluation is performed with several metrics determining anatomical accuracy and realistic noise properties.

In order to provide a baseline to compare the metrics of anatomical accuracy to, the associated evaluation framework is tested with modified data. A CT-CT evaluation is performed by randomly assigning in-house CT images to each other and subsequently calculating the associated anatomical accuracy metrics. In addition, XCAT phantoms are modified by (1) applying random rigid transformations and (2) adding Poisson noise. Afterwards, the anatomical accuracy metrics are calculated with respect to the original XCAT slices. The baseline metrics are shown in Table 4.1. The anatomical accuracy metrics calculated for synthetic CT images are supposed to range between the ones obtained by CT-CT comparison as the minimum, and up to the ones obtained by added Poisson noise as the maximum. This is true for every metric except the EGR, which is sensitive to noisy inputs, as no additional edges are generated for the baseline calculations.

Configuration Study

Several configurations are evaluated and compared to determine the best performing network configuration, an overview of evaluation metrics is shown in Table 4.2. First, the evaluation metrics of both network architectures are compared. Only in the EGR metric, the U-Net showed a performance improvement of approximately 10 %. Apart from this, the Res-Net performed similarly or better than the U-Net in all other metrics, ranging from 20 % for the NCC to up to 55 % for the MAE and even up to 95 % for the NMPE. As a consequence of these substantial deviations, the U-Net based networks are excluded from the following comparisons.
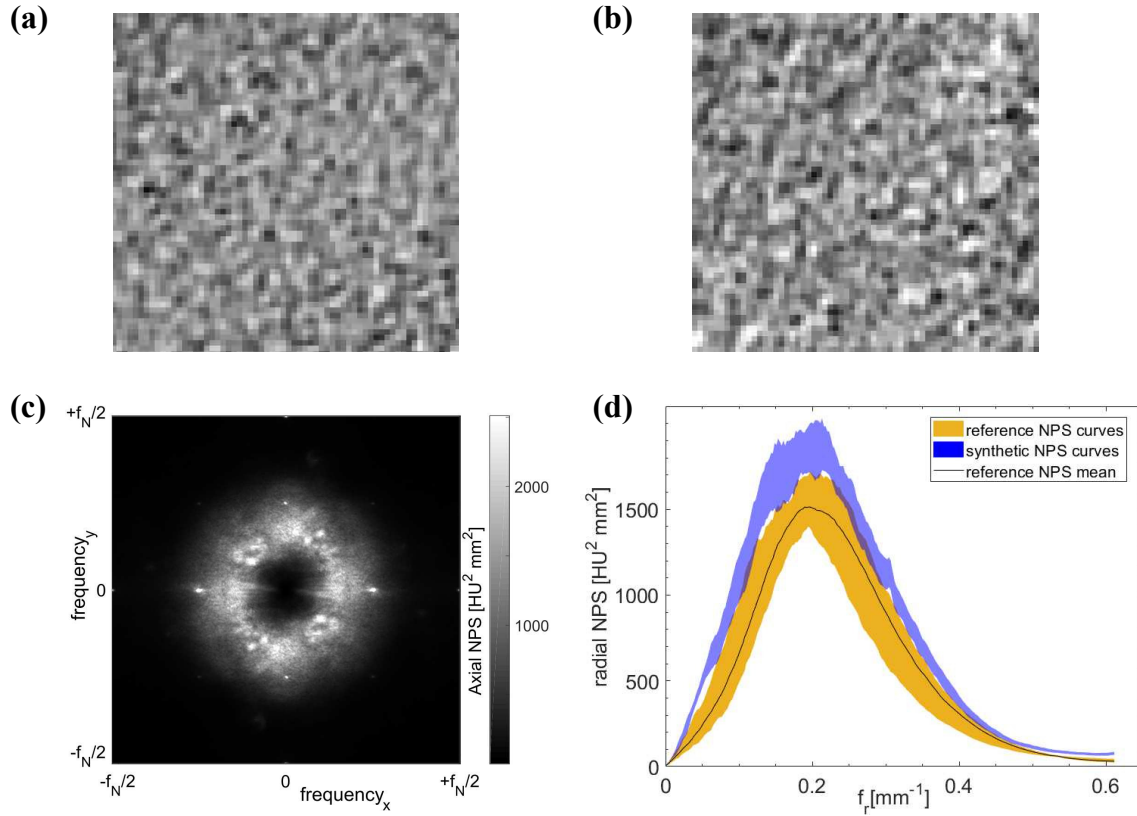
Figure 4.3: Noise realizations and NPS. **(a)** Synthetic, **(b)** real noise realization as well as **(c)** synthetic axial NPS (x-axis shows the frequency in x-direction, y-axis in y-direction, f_N denotes the Nyquist frequency) where and **(d)** comparison of synthetic and real radial NPS. The colored areas represent the convex hull of all calculated NPS, respectively. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature International Journal of Computer Assisted Radiology and Surgery ©2019.

Table 4.2: A comparison of the evaluation metrics from the configuration study in terms of mean values $\pm$ standard deviation. The number of network cases considered for each comparison is stated in brackets, it is reduced from 8 to 4 due to the exclusion of U-Nets. The NM in the reference patient CTs amounts to (33.26 ± 4.21) HU. Bold values indicate superior performance

Metric	Network architecture [8]		Nr. input slices [4]		Body contour [4]	
	Res-Net	U-Net	3	5	without	with
RCBV ↓	**0.30 ± 0.16**	0.80 ± 0.34	**0.22 ± 0.07**	0.38 ± 0.19	0.42 ± 0.16	**0.18 ± 0.03**
MAE ↓	**143 ± 116**	606 ± 294	**115 ± 26**	172 ± 158	**113 ± 25**	174 ± 157
SSIM ↑	**0.69 ± 0.05**	0.66 ± 0.05	**0.69 ± 0.05**	0.68 ± 0.05	**0.69 ± 0.05**	0.68 ± 0.05
FSIM ↑	**0.59 ± 0.05**	0.55 ± 0.10	**0.61 ± 0.04**	0.56 ± 0.06	**0.61 ± 0.03**	0.56 ± 0.06
EGR ↓	2.32 ± 0.43	**1.88 ± 0.35**	2.38 ± 0.39	**2.26 ± 0.46**	2.36 ± 0.43	**2.28 ± 0.43**
NCC ↑	**0.99 ± 0.01**	0.72 ± 0.14	0.99 ± 0.00	0.99 ± 0.01	0.98 ± 0.01	**0.99 ± 0.01**
NMPE ↓	**9.6 ± 5.3**	927 ± 686	**8.7 ± 4.5**	10.5 ± 6.1	**8.7 ± 4.4**	10.5 ± 6.1
NMMR ↓	6.03 ± 3.52	**5.68 ± 4.84**	7.38 ± 4.37	**4.69 ± 1.63**	**4.35 ± 1.50**	7.72 ± 4.16
NM	**32.0 ± 4.2**	61.9 ± 42.8	35.0 ± 4.0	28.9 ± 1.2	33.9 ± 4.8	30.0 ± 2.3

Figure 4.4: One axial slice of **(a)** XCAT phantom, **(b)** corresponding synthetic CT generated with TASKNET3 and **(c)** patient CT obtained from a similar body region. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature International Journal of Computer Assisted Radiology and Surgery ©2019.
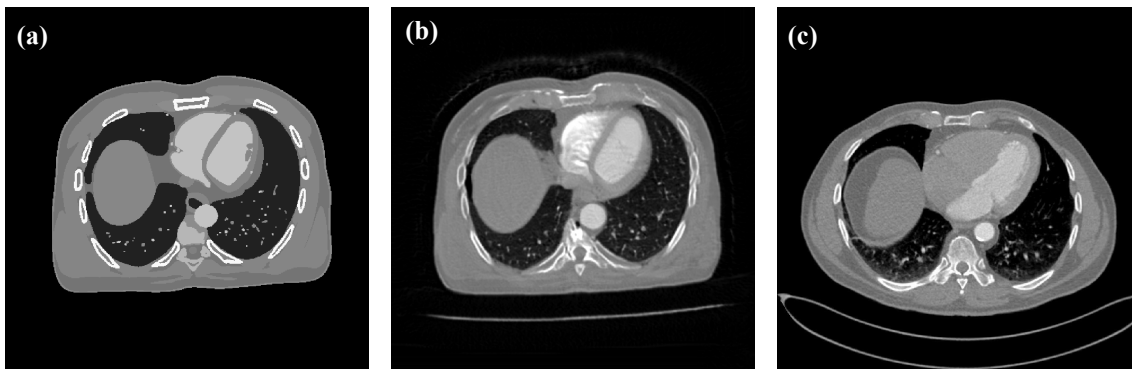
Secondly, the influence of the number of input slices on the Res-Nets is evaluated. Here, the NMMR showed a performance degradation of approximately 40 % for the training with three instead of five input slices. However, the configuration using three input slices yielded a performance improvement for both the EPR and the RCBV of about 15 % and 55 %, respectively. Both improvements can be observed in Figure 4.1. Here, **(b)** and **(c)** show a lower extent of structures outside the body contour as well as a better edge preservation in artery and bone regions when compared to **(d)** and **(e)**. Finally, the use of a body contour mask is evaluated. While this approach demonstrated an improvement in the RCBV of up to 65 %, performance reductions of 10 % for the MAE as well as 20 % for the SSIM are observed.

Based on these observations, the Res-Net architecture using three input slices and no additional pre-processing step is identified as the ideal network configuration. This is underlined by the corresponding accurate modeling of CT noise in the liver, which is illustrated in Figure 4.3. Here, a synthetic noise realization (**a**) is directly compared to a real one (**b**), demonstrating only minor deviations from radial symmetry (**c**) as well as similar noise spatial correlation and magnitude (**d**) when comparing the distributions of the radial NPS to reference data.

Task-Based Study

The synthetic CT images generated with the task-based networks are compared to study the impact of the dedicated loss function and the use of additional CT images from public sources as training data. The corresponding results are shown in Table 4.3. The computation time to simulate CT data from a digital body phantom is around one second per slice, amounting to around five minutes for a typical CT image with 300 to 500 slices.

When comparing the performance of TASKNET1 to TASKNET2, an improvement in the MAE_{ves} of 55 % in the contrast-enhanced vessels can be observed. Simultaneously, some metrics show a degradation in anatomical accuracy. However, these is-

Table 4.3: The averaged evaluation metrics (± standard deviation) of the task-based networks. **Bold** values indicate superior performance.

Metric	TASKNET1	TASKNET2	TASKNET3
RCBV ↓	0.07 ± 0.02	$\mathbf{0.02 \pm 0.00}$	0.05 ± 0.01
MAE ↓	172 ± 47	159 ± 49	$\mathbf{82 \pm 15}$
SSIM ↑	$\mathbf{0.80 \pm 0.10}$	0.61 ± 0.06	0.64 ± 0.08
FSIM ↑	0.64 ± 0.05	0.53 ± 0.05	$\mathbf{0.76 \pm 0.03}$
EPR ↑	0.22 ± 0.03	0.23 ± 0.03	$\mathbf{0.47 \pm 0.04}$
EGR ↓	$\mathbf{1.31 \pm 0.19}$	1.49 ± 0.27	2.29 ± 0.45
NCC ↑	0.45 ± 0.04	0.63 ± 0.10	$\mathbf{0.92 \pm 0.04}$
NMPE ↓	30.3 ± 1.2	24.1 ± 4.1	$\mathbf{20.8 \pm 3.5}$
NMMR ↓	$\mathbf{2.08 \pm 0.32}$	2.46 ± 0.77	2.38 ± 0.36
NM	24.4 ± 1.6	29.1 ± 2.1	$\mathbf{30.4 \pm 1.8}$
MAE_{ves} ↓	256 ± 144	119 ± 47	$\mathbf{71 \pm 34}$

sues are not observed in the performance of TASKNET3. For TASKNET3, anatomical accuracy metrics indicate superior performance closer to the Poisson baseline values, while at the same time improving the generation of realistic noise properties. Consequently, the output images of TASKNET3 are chosen for the following proof-of-principle experiment.

In Figure 4.4, a graphical comparison of (**b**) a synthetic CT slice to (**a**) the corresponding XCAT slice and (**c**) a real CT image obtained from a similar body region is shown. In particular, the synthetic slice accurately depicts the anatomy provided by the XCAT phantom, from high contrast structures such as bones and contrast-enhanced vessels, down to low contrast in soft tissue regions. In comparison to the real CT, however, major similarities regarding realistic CT noise as well as the CT characteristic imaging performance in the bone regions are observed. Moreover, the blurred appearance of the lung vessels are also reproduced. One apparent difference, however, can be seen in the left cardiac chamber, which is contrast-enhanced in the XCAT phantom, but not in the in-house CT images.

Apart from the similarities, it is also clearly observable, that certain features from the XCAT slice are in fact not preserved during CycleGAN processing. This is particularly noticeable in the bony regions, where for example the spine structure is substantially altered by the CycleGAN when comparing to the original XCAT input. Furthermore, bony structures are artificially generated, which is visible in the central right region, near the ribs. Here, bony structures are generated with no baseline in the corresponding XCAT slice. This behaviour is also observed via higher values in the EGR metric. Due to this behavior, the CycleGAN approach is not further investigated for the simulation of CBCT images. Instead, the simulated CT images are used to study the benefits of incorporating simulated CT images when training neural networks for segmentation (Section 4.1.1).

(a)	**(b)**
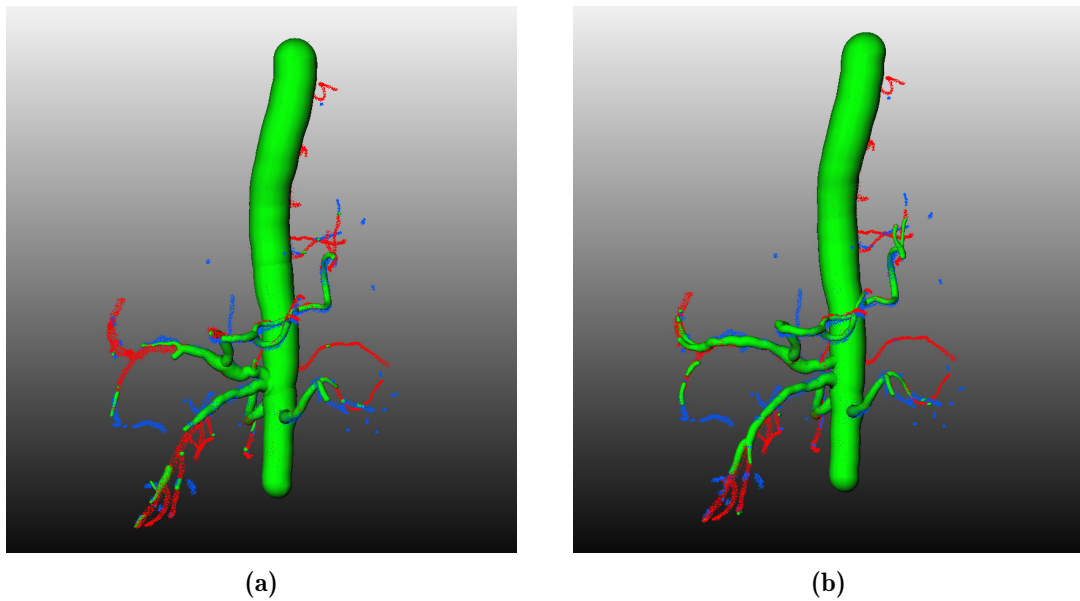
Figure 4.5: Segmentation by the U-Nets trained on **(a)** a real data set and **(b)** combined data set of both real and synthetic CT images. Green indicates where the segmentation matches the GT, red indicates false negatives and blue false positives. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature International Journal of Computer Assisted Radiology and Surgery ©2019.

(a)	**(b)**
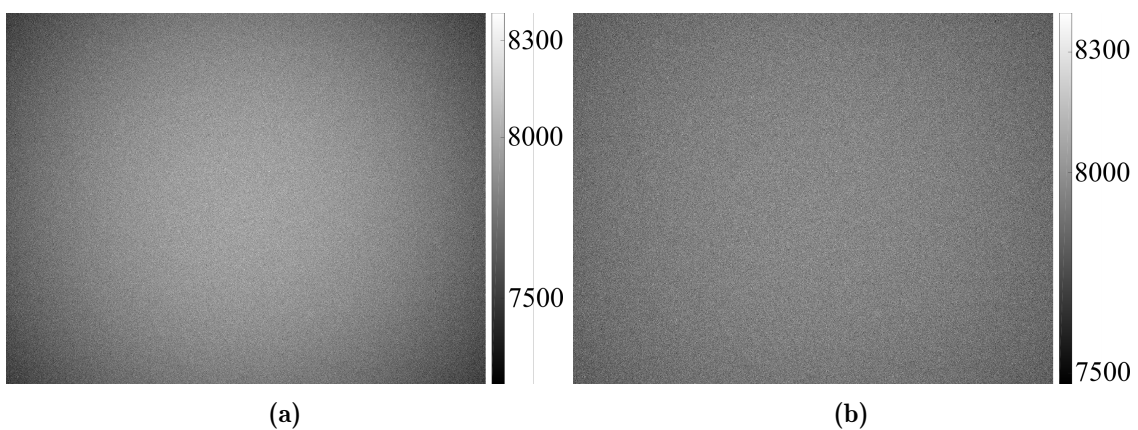
Figure 4.6: Bare beam fluence on the detector simulated with GATE using 20 billion photons emitted in conical shape from the source towards the detector. The bare beam fluence in **(a)** can be corrected for the inverse-square law to yield a stable noise level **(b)**.
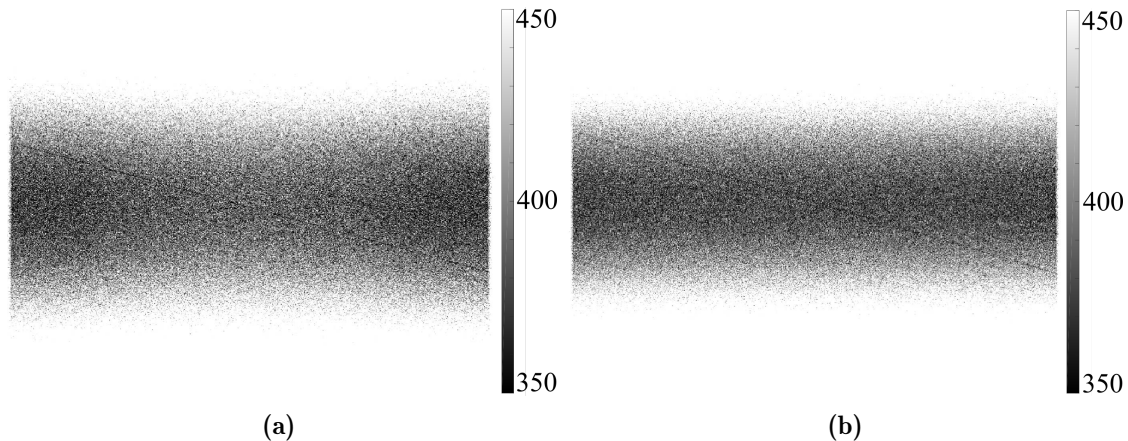
Figure 4.7: Fluence values on the detector simulated with the slanted tungsten wire phantom for **(a)** the MC approach and **(b)** the Astra projection operator plus noise model.

Table 4.4: Coefficient of determination (R^2) for the linear regression model applied to the center-of-mass positions, FWHM of the one-dimensional MTF and MPE of the alternative orbits in relation to the circular orbit are tabulated.

	GATE			Projection operator		
Metric	Circular	Tilted	Triple-arc	Circular	Tilted	Triple-arc
R^2 ↑	0.999	0.996	0.995	0.999	1.000	0.997
FWHM $[mm^{-1}]$ ↑	0.667	0.600	0.534	1.101	0.881	0.561
MPE [%] ↓	/	15.1	20.3	/	6.5	30.4

Segmentation Study

In the segmentation study, a DSC of 0.72 ± 0.05 is obtained with the U-Net trained on synthetic images (from TASKNET3), 0.78 ± 0.11 with the U-Net trained on in-house CT images and 0.83 ± 0.05 with the U-Net trained on a combined data set of both sources. Figure 4.5 shows the vessel graphs segmented by the U-Nets trained on **(a)** real data and **(b)** the combined data set compared to the GT. Here, the largest vessels are correctly segmented. However, all networks showed a high abundance of false negatives for smaller blood vessels.

4.1.2 Comparison of Monte-Carlo and Projection Operator Simulations

To investigate adequate simulation of CBCT projections two different methods are compared: MC simulations and a projection operator based approach. For the MC simulations, a bare beam fluence of 20 billion photons is simulated without any object between source and detector and the resulting detector counts are shown in Figure 4.6a. Upon visual inspection, the fluence level is observed to decrease with increasing distance from the image center, which coincides with the center of the detector. Compared to the bare beam fluence projection that is corrected for the inverse-square law, this behavior virtually disappears and a stable noise level is ob-

(a)   **(b)**
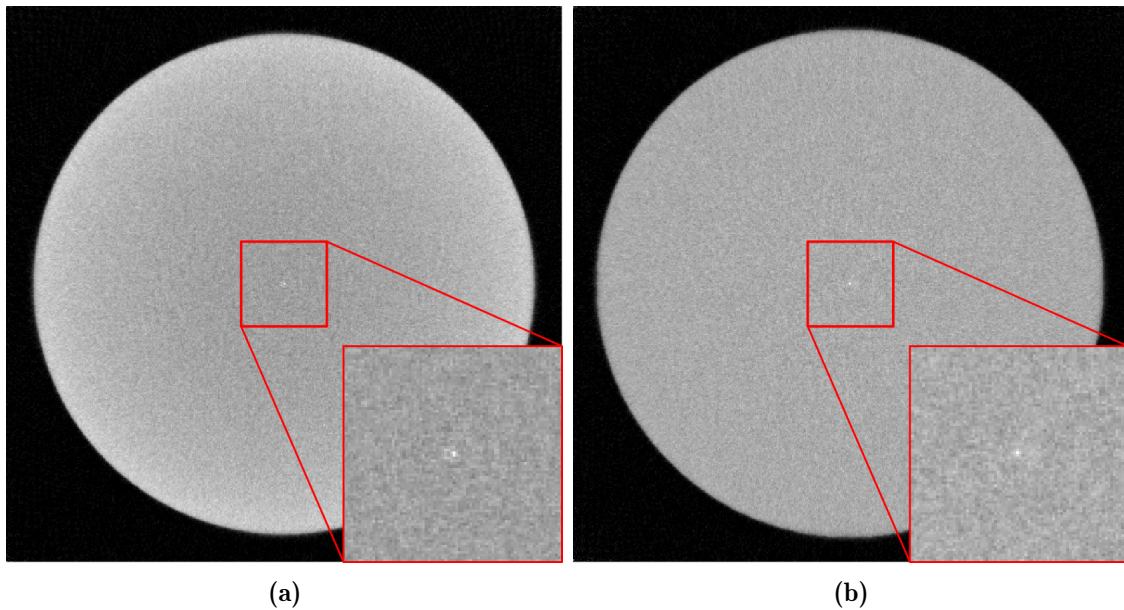
Figure 4.8: Axial slices and magnified central regions of the reconstructions calculated via SIRT of **(a)** the MC simulations and **(b)** of the projection operator simulations.

(a)   **(b)**
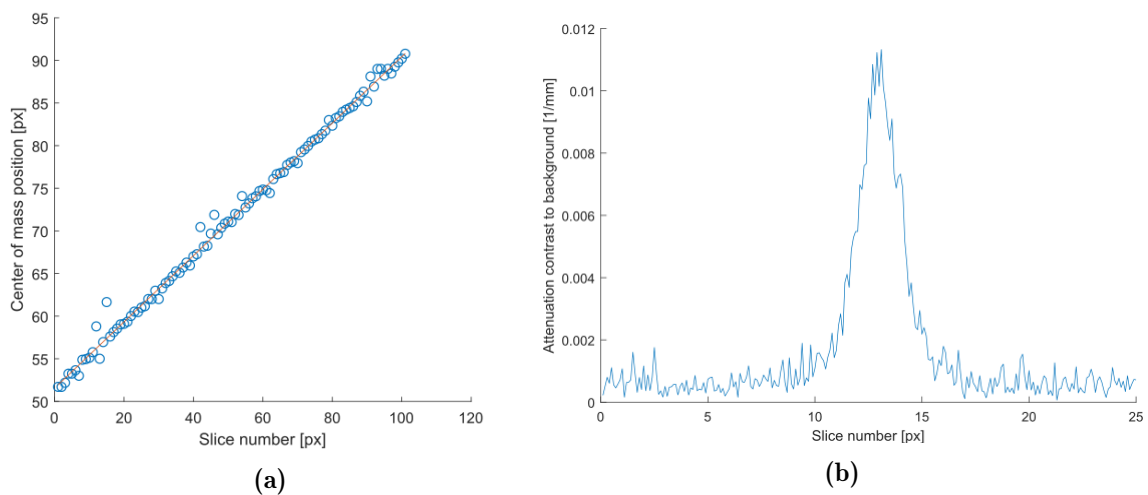
Figure 4.9: Linear regression model applied to the COM positions in each slice in z-dimension **(a)** and the resulting LSF in x-direction with sub-voxel precision **(b)** obtained using the projections simulated with GATE for a tilted circular orbit.
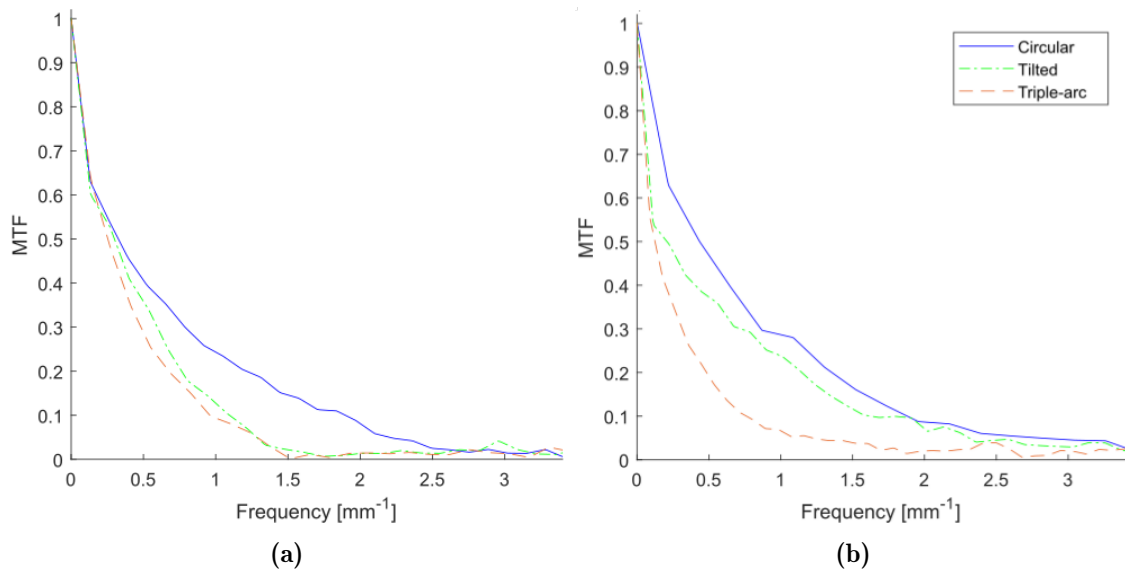
Figure 4.10: One dimensional MTF for **(a)** GATE simulations and **(b)** projection operator simulations of the circular, the tilted circular and the triple-arc orbit.

served in all regions of the projection.

By averaging 20 realizations of bare beam fluence simulations a baseline bare beam fluence I_0 is calculated, which is then used for the log-transformation of the simulated projections before reconstruction, yielding attenuation line integrals. In the next step a water cylinder containing a slanted tungsten wire is incorporated as the imaged object in the simulation. One exemplary projection of the circular orbit is shown in Figure 4.7a with a grayscale window adapted to underline the contrast in photon counts. The slanted tungsten wire is clearly visible with lower photon counts, *i.e.* more attenuation. A decrease in photon counts for increasing distance from the center can be observed in the profile along the longitudinal axis of the cylinder.

On a local, non-optimized machine with four CPU cores $4 * 10^5$ particles can be simulated per minute ($4 * 10^5$ particles per core) amounting to a calculation time of about five weeks per projection, consisting of 20 billion particles each. Processing time for 200 projections would therefore be around 20 years with this hardware. In comparison to the local machine, GateLab provides 500 cores per submitted job and allows ten simultaneous jobs per user, amounting to the availability of 5000 cores. In total, the calculation of 200 projections using GateLab takes about eight days.

To investigate a faster method for simulation, a projection operator based method with an added noise model is evaluated. In contrast to the MC method, the bare-beam fluence can be incorporated directly in the noise model and no bare beam fluence simulations are needed. An exemplary projection of the slanted tungsten wire phantom is depicted in Figure 4.7b. Reduced photon counts behind the slanted tungsten wire are clearly visible, while the water cylinder throws a clear rectangle 'shadow' with increasing counts towards the upper and lower edge of the cylinder.

In comparison to the GATE simulations, the photon count values behind the water cylinder are fluctuating in a similar range interval between 450 and 550 counts. In contrast to the MC projections, there is no inverse-square law to be corrected for. The simulation of a whole set of 200 projections including the noise model takes 7.5 minutes on the local machine, 0.7 ‰ of the GateLab calculation time. Ten seconds of those seven minutes are needed for the projections and the remainder for the noise simulation. A substantial reduction in calculation time can be achieved by decreasing the number of detector pixels.

The above described projections are simulated for three different orbits: standard circular, tilted circular and a triple-arc orbit, illustrated in Figure 3.8. The 100 iterations of SIRT reconstruction performed on the simulated projection data take around one hour per set of projections. Axial slices of the SIRT reconstruction obtained with the tilted circular orbit are shown in Figure 4.8, one each for the MC and for the projection operator simulations. While the reconstruction obtained from the MC simulations exhibits beam hardening artifacts in the form of underestimated attenuation coefficients towards the center of the cylinder, no such artifacts can be found in the volume reconstructed from the projection operator simulations. Furthermore, a minor ring artifact around the tungsten wire can be observed in the magnified region of the MC reconstruction. At the same time, the tungsten wire appears slightly more blurry in the reconstruction simulated with the projection operator.

From these reconstructions, the MTF in the direction longitudinal to the x-dimension (dimension perpendicular to the tilt of the tungsten wire) can be calculated using the oversampled LSF approach. The linear regression model applied to the COM positions for each reconstruction slice is shown in Figure 4.9 together with the resulting oversampled LSF, both exemplary depicting the evaluations of the tilted circular orbit. It can be observed, that there are some outliers in the COM data that did not fit the linear regression model. The obtained oversampled LSF exhibits a peak with sub-voxel precision (sampling rate of 10/voxel) overlayed with noise. The LSF is Fourier-transformed to yield the one-dimensional MTF, depicted in Figure 4.10. An overview of the evaluation metrics (R^2, FWHM and MPE) is provided in Table 4.4. For both methods, there is the most favourable values for the circular orbit, indicating the best performance. The tilted circular orbit yields the second-most favorable outcome, while in general the predicted MTF is higher for the projection operator based method. The least favorable results are predicted for the triple-arc orbit, while the MPE metric for this orbit is slightly worse for the projection operator based simulations. The implementation of the MTF evaluation takes around 15 seconds of calculation time per reconstructed image.

Summarizing computation time, the simulation of each set of projections takes 8 days for the MC method and 7.5 minutes for the projection operator approach. In addition, the SIRT reconstruction is computed in one hour and the final MTF evaluation in 15 seconds. In comparison, an implementation of a state-of-the-art detectability calculation (without Fourier plane optimization) [21] takes around three
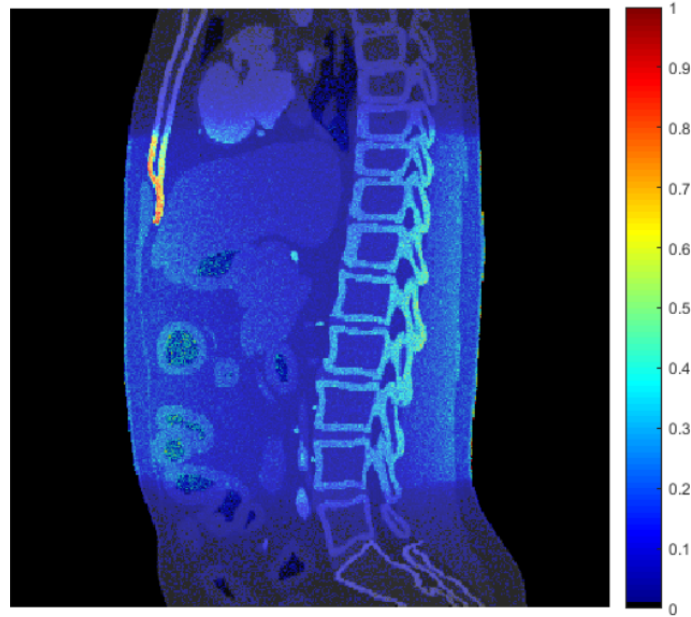
Figure 4.11: Sagittal slice of an XCAT phantom overlaid with a relative dose map simulated using GATE.

days on the same hardware.

Apart from simulating the signal on the detector, GATE simulations can be used to calculate the physical dose imparted in the imaged object. To demonstrate this capacity for anthropomorphic data, one realization of the digital XCAT phantom is incorporated and the relative dose in each image voxel is calculated. A sagittal slice of the phantom overlaid with a relative dose map is depicted in Figure 4.11. A distinctly conical dose distribution in the central torso region can be observed. bony structures and regions more closely located to the body surface show more pronounced dose values.

4.2 Fast Reconstruction of Arbitrary CBCT Orbits Using CNNs

This section shows the results for the experiments performed for a novel method to reconstruct CBCT images from arbitrary orbit geometries. The proposed algorithms consists of a deconvolution operation with the system response followed by CNN processing. First, the appropriate method for regularizing the deconvolution operation is investigated (Section 4.2.1), the algorithm is applied to single geometries (Section 4.2.2) and ultimately to arbitrary geometries (Section 4.2.3).

4.2.1 Regularization of Deconvolution

To evaluate different methods of regularization in Fourier domain, a parameter sweep is performed and the resulting deconvolved volumes are qualitatively assessed. Specifically, a quadratic penalty, a complex rescaling to threshold value and a fixating to threshold value approach are compared. A representative comparison of
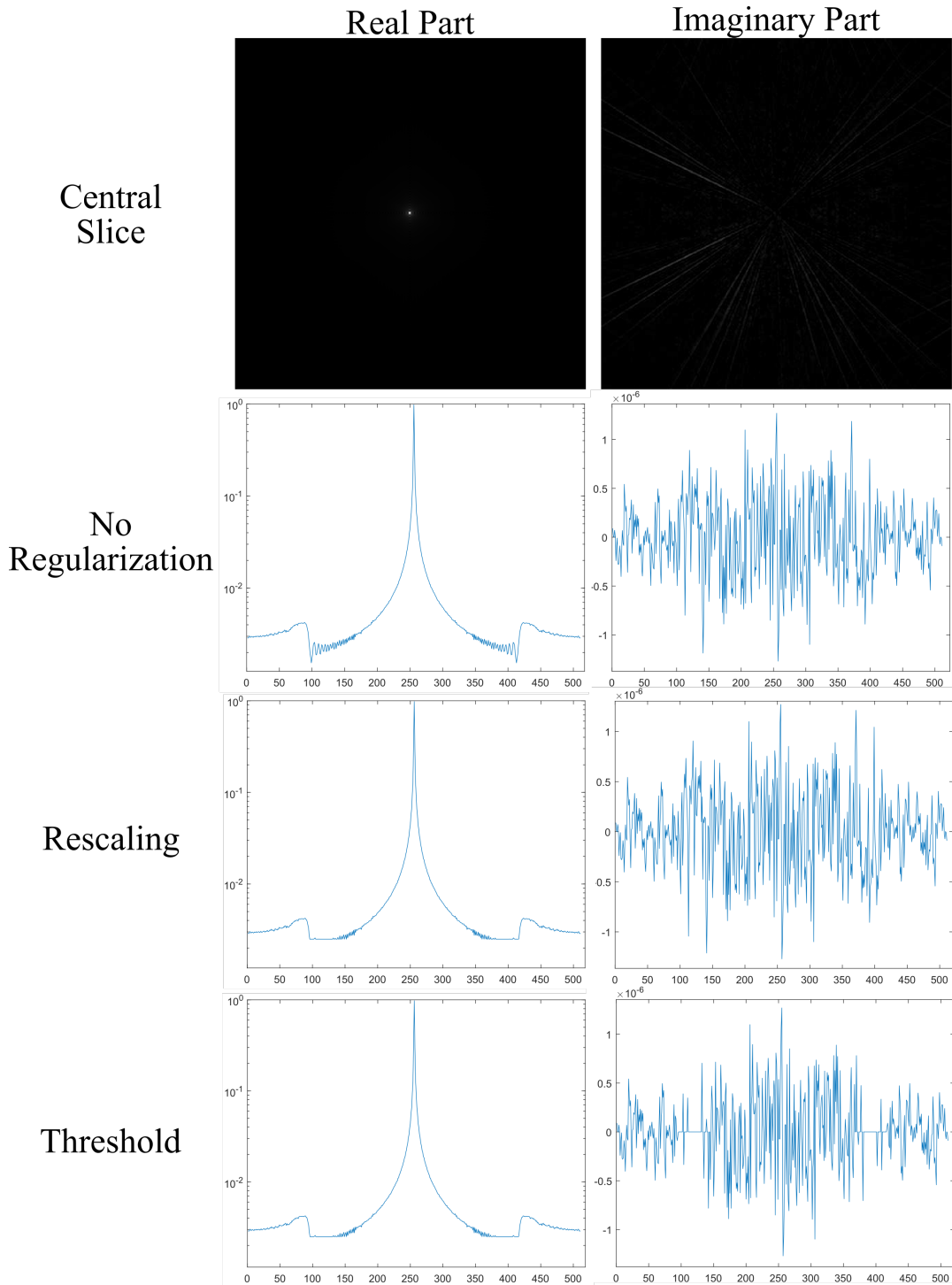
Figure 4.12: The first row shows the central slice of the Fourier transformed system response $\mathcal{F}\left\{\mathbf{A}^T\mathbf{A}e_j\right\}$ of a $\phi = \sin(3\theta)$ orbit - divided into real (left column) and imaginary part (right column). The second to fourth rows show the profile through the central row of the Fourier transformed system response for unregularized (second row), complex rescaled (third row) and thresholded regularization data. The x-axis in these 3 rows shows the column number. Note the regularized regions in columns 90-130 and 380-420, where modulus values exceed the selected threshold of 0.0015.
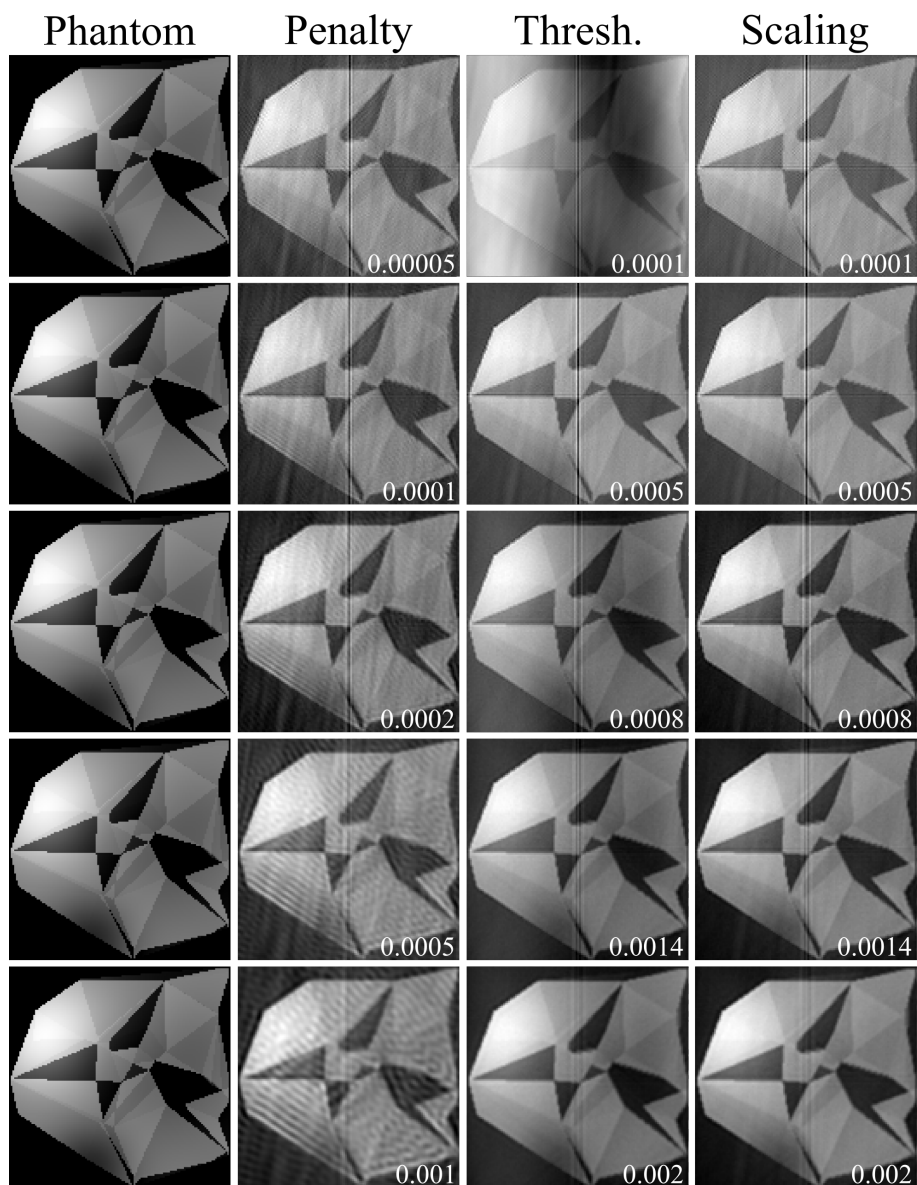
Figure 4.13: Axial slices of deconvolved image volumes obtained via different Fourier regularization methods for one random orbit. Each column corresponds to a specific method, from left to right: original phantom volume; quadratic penalty; fixed values at threshold; values rescaled to threshold. Rows from top to bottom increase in regularization strength λ, each image contains the utilized regularization value in white.

methods and regularization strengths is illustrated in Figure 4.13, obtained with one random orbit.

At the low end of the regularization strength, there is substantial amount of high-frequency noise distorting the images and thereby complicating the recovery of the original features by the CNN. This is particularly noticeable in an alternating grid structure at the upper left corner as well as stripe artifacts throughout the whole deconvolved volume of the deconvolution obtained with the quadratic penalty method. In addition, the center sagittal slices show a major decrease in attenuation values while the sagittal slices neighboring the central slice show an increase. This behavior is noticeable for all regularization methods and is consistently less visible for higher regularization strengths. When increasing regularization strength for the quadratic penalty method, the high-frequency grid starts transforming into a randomly directed composition of sinusoidal waves of decreasing frequency. At the same time, high-frequency features such as edges become less visible. The latter is true for all regularization methods when increasing regularization strength.

A point of ideal trade-off between lost high-frequency content, dominant high-frequency noise and sagittal artifacts is selected for each method; namely regularization strengths of $\lambda = 0.0002$, $\lambda = 0.0008$ and $\lambda = 0.0008$ for the quadratic penalty, complex rescaling and the thresholding method, respectively, are chosen based on qualitative inspection. In comparison with the rescaling and thresholding, the regularization strength for the quadratic penalty approach can not be satisfactory chosen in order to match the qualitative performance of both other methods. For the latter, a common ideal value of the regularization strength lead to comparably good performance.

One disparity of the rescaling method compared to the thresholding regularization is illustrated in Figure 4.12. Here, an exemplary Fourier transformed system response is depicted - divided into real and imaginary part. The profiles through the central row of the image are also shown. In the real part of the profile, the differences between regularized and non-regularized in the thresholded regions ($\lambda < 0.0015$ in this case) are noticeable, namely for column numbers around 90-130 and 380-420. While the non-regularized profile contains some oscillations in this region, the regularized profiles are fixated to the threshold value. In contrast to the real part, the imaginary part of the profile obtained with the rescaling approach is perfectly preserved. This is not the case for the thresholding approach, where all values in the above described region are set to zero.

For the following reconstructions, the thresholding regularization is chosen for our initial study where the proposed reconstruction pipeline is applied only to single geometries (Section 4.2.2), but eventually the rescaling method is favored for the application to arbitrary orbits (Section 4.2.3).
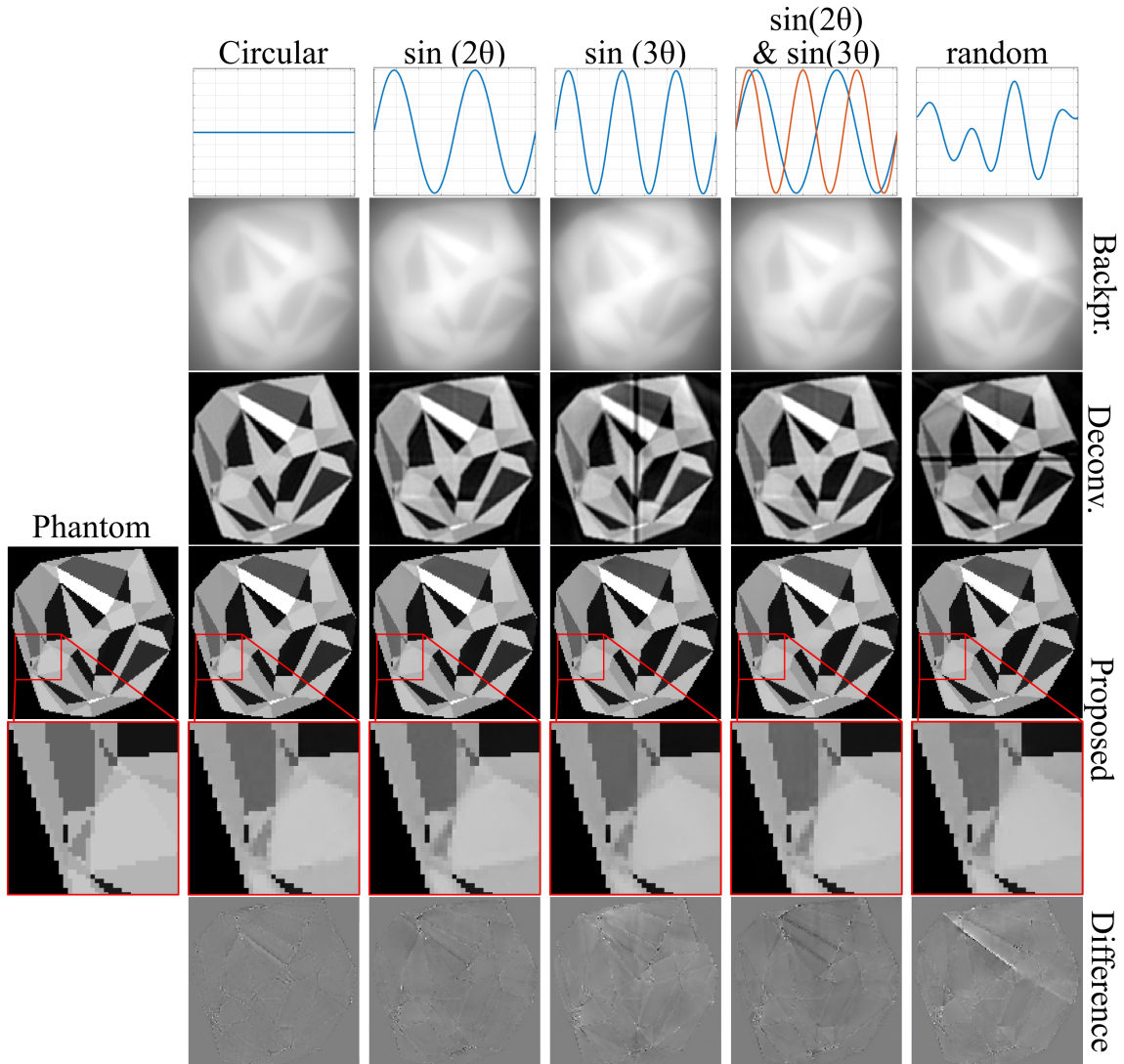
Figure 4.14: Intermediate image volumes and final reconstruction outputs from the reconstruction pipeline. Each column corresponds to a set of imaging geometries. Rows from top to bottom: elevation angle ϕ as a function of rotation angles θ; backprojected volume; volume after deconvolution; CNN reconstruction; zoomed in region of interest (ROI) within the slice; difference image between the reconstructions and GT phantom images (axial slices). The backprojection volumes underline the suppression of high-frequency content by the $\mathbf{A}^T\mathbf{A}$ operation. Note the central slice artifacts in the deconvolved volumes of the $\sin(3\theta)$ and the random orbit. They are not noticeable in the circular and the $\sin(2\theta)$ orbit volumes.

Table 4.5: Evaluation metrics for the proposed pipeline compared with SART. All metrics are evaluated between the reconstructions and GT phantom images. The better of two values is marked **bold**.

		circular	$\sin(2\theta)$	$\sin(3\theta)$
Prop.	nRMSE ↓	**0.033 ± 0.005**	**0.048 ± 0.007**	**0.060 ± 0.008**
	FSIM ↑	**0.991 ± 0.005**	**0.983 ± 0.010**	**0.979 ± 0.010**
	SSIM ↑	**0.994 ± 0.002**	**0.987 ± 0.004**	**0.984 ± 0.003**
SART	nRMSE ↓	0.116 ± 0.013	0.105 ± 0.016	0.109 ± 0.015
	FSIM ↑	0.937 ± 0.019	0.943 ± 0.015	0.940 ± 0.015
	SSIM ↑	0.941 ± 0.011	0.963 ± 0.011	0.956 ± 0.011
		$\sin(2\theta)\&\sin(3\theta)$	random	
Prop.	nRMSE ↓	**0.062 ± 0.007**	**0.061 ± 0.009**	
	FSIM ↑	**0.977 ± 0.013**	**0.979 ± 0.013**	
	SSIM ↑	0.944 ± 0.013	**0.985 ± 0.007**	
SART	nRMSE ↓	0.107 ± 0.016	0.108 ± 0.015	
	FSIM ↑	0.942 ± 0.015	0.941 ± 0.015	
	SSIM ↑	**0.960 ± 0.011**	0.958 ± 0.010	
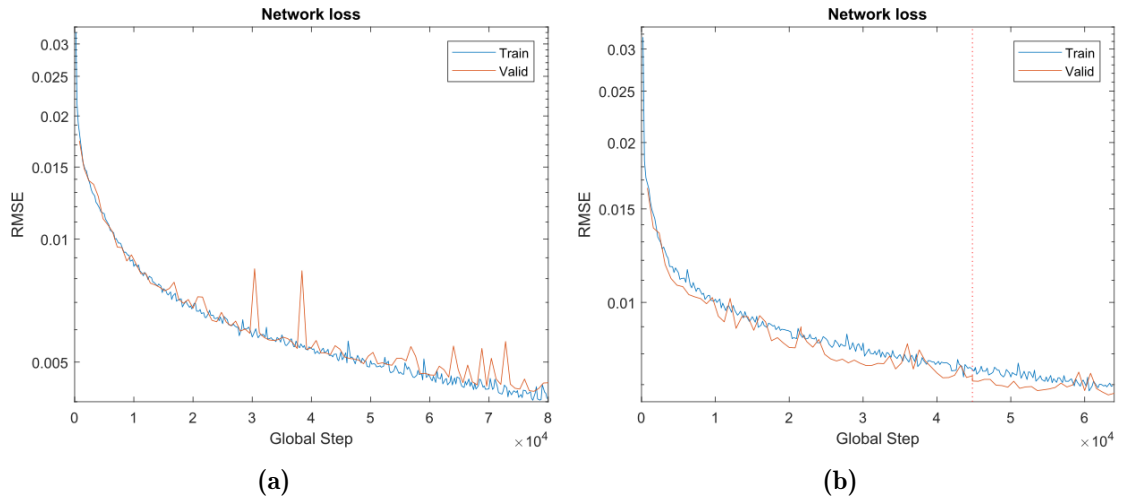
Figure 4.15: Plot of the loss functions for training and validation data sets of **(a)** the reconstruction CNN applied to both $\sin(2\theta)$ and $\sin(3\theta)$ orbits as well as **(b)** the reconstruction CNN trained on tetrahedral data for 70 epochs (red, vertical line indicates 70 epochs), fine-tuned on LIDC data for 30 epochs and applied to random orbits.

4.2.2   Application to Single/Dual Geometries

To investigate the feasibility of the proposed reconstruction pipeline, the initial focus is on its application to data from single orbits. Starting with circular orbits, this approach is applied to simple sinusoidal orbits and eventually a linear combination of sinusoidal basis functions with randomly generated coefficients. To investigate the ability of the proposed approach to generalize to more than one orbits, its application to two different sinusoidal orbits at the same time is also investigated. Intermediate images and final reconstruction outputs are illustrated in Figure 4.14.

The evaluation metrics are calculated for 10 test cases the networks have not seen before; they are compared in Table 4.5, for both the proposed reconstruction pipeline as well as SART. To illustrate training convergence and check for differences in the loss metric between training and validation data sets, the network loss metric of both data sets is plotted logarithmically in Figure 4.15a exemplary for the network trained on both $\sin(2\theta)$ and $\sin(3\theta)$ orbits. As the decrease of the loss function of both data sets is reaching a plateau, the metric values are simultaneously reaching convergence at an nearly identical loss level.

To mitigate artifacts originating from the boundary value problem, the size of the backprojected volumes (system response, ray density, backprojection of projection data) is increased to 512 voxels in every dimension. Note the residual artifacts in the deconvolved volumes of the $\sin(3\theta)$ and the random orbit, which are not noticeable in the circular and the $\sin(2\theta)$ data. Further investigations also show, that the deconvolved volumes obtained with a $\sin(4\theta)$ orbit exhibit none of these artifacts. This indicates that the proposed regularization methods and the increase of backprojected volumes are not sufficient to consistently mitigate artifacts.

The CNN-based approach consistently outperforms the SART reconstructions in terms of nRMSE and FSIM. This is also the case for SSIM except for the network trained on two sinusoidal geometries. After CNN processing, the tetrahedron outlines and attenuation values are well recovered for all orbit geometries. While SART performs comparably for all geometries with only slight deviations, the performance of the CNNs show noticeable differences for the different geometries. Specifically, reconstruction performance decreases with increasing orbital complexity. This is apparent in the slightly decreasing evaluation metrics, in the magnified areas and the difference images in Figure 4.14. The magnified regions contain fine-grain details, which every CNN struggles to reconstruct accurately, while the best performance is observed for the circular orbit and a decrease is observed for increasing orbital complexity. Compared to the networks trained on single geometries, visual inspection and the calculated metrics indicate only minor performance decrease of the network trained on two different geometries - $\sin(2\theta)$ & $\sin(3\theta)$.

To evaluate the performance consistency, the standard deviation of the calculated evaluation metrics can be examined. Compared to SART, the proposed reconstruction pipeline shows generally lower standard deviations, *i.e.* more consistent re-

Figure 4.16: Axial slices of intermediate image volumes and reconstruction outputs obtained with a random orbit and tetrahedral data. Columns from left to right: phantom input; volume after deconvolution; CNN reconstruction; MBIR reconstruction. Rows from top to bottom: Axial slice; zoomed in ROI within the slice; axial slice of the difference image between the respective volume and GT.

construction performance. While performance consistency seems to be independent of the type of reconstructed orbit geometry for the SART algorithm, the proposed approach performs less consistent with increasing orbit complexity.

The majority of the computation time for the proposed method is spent on the calculation of the system response (five minutes), the remainder for the backprojection (3.5 seconds), the calculation of the ray density (30 seconds), and ultimately for the deconvolution operation (20 seconds). The CNN prediction itself takes around 1 second, resulting in 6 minutes of overall computation time. In comparison, SART reconstructions take approximately 50 minutes for 50 iterations on a workstation with comparable specifications. Aside from the CNN processing, which has been performed on a computing cluster with optimized software, the mentioned implementations have not been optimized for runtime.

4.2.3 Application to Random Geometries

This section outlines the experiments performed to apply the proposed reconstruction pipeline to a general class of random acquisition orbits, with some minor adaptions compared to the single orbit approach. Here, the networks are trained on 1000 unique phantoms matched with the same number of unique orbits. This means that the orbits in training, validation and test data set are composed of data obtained
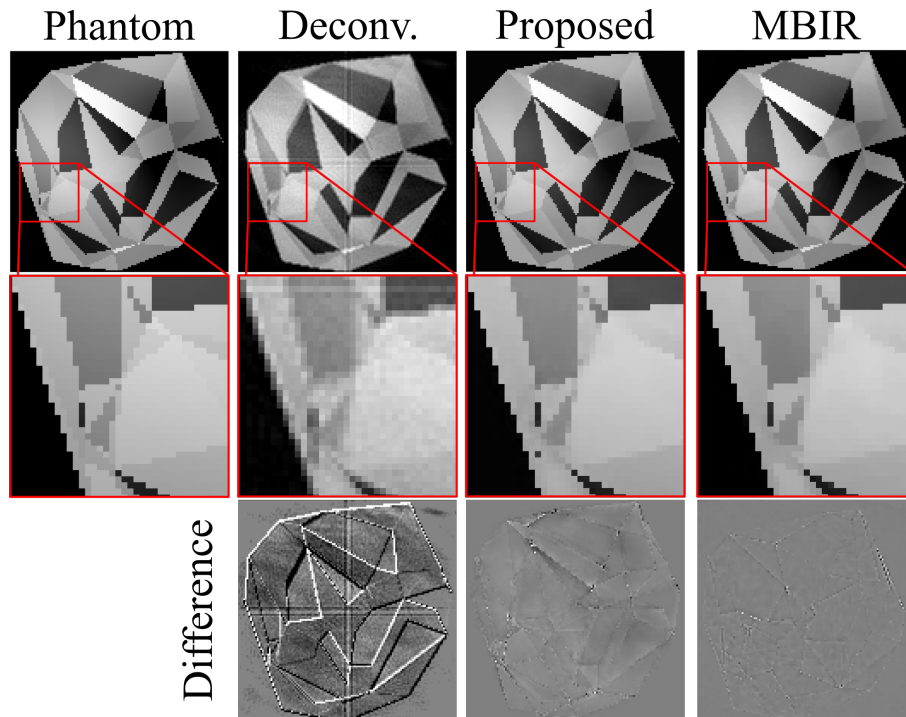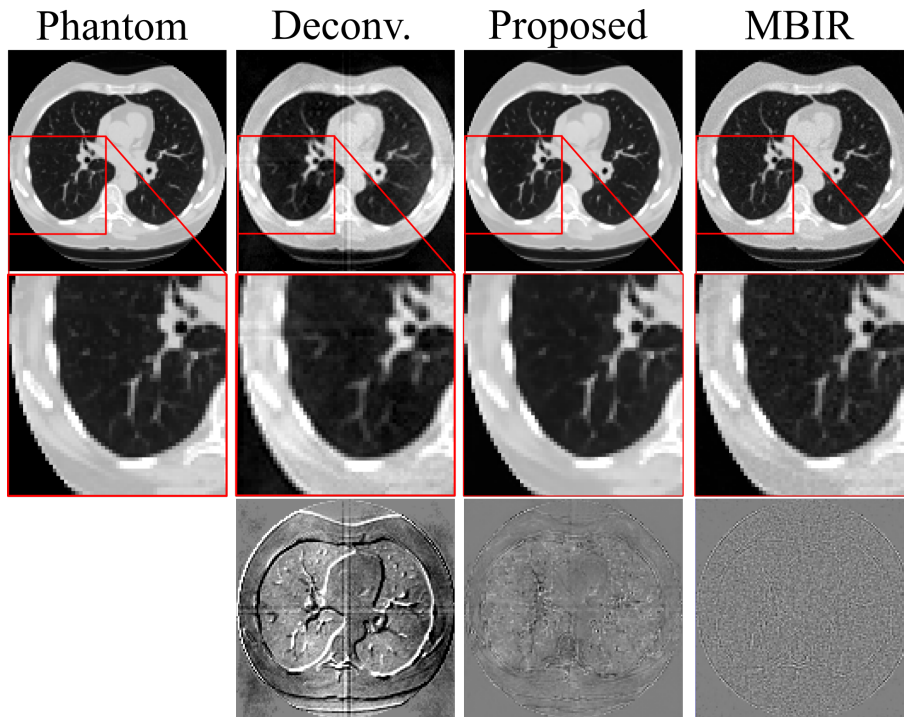
Figure 4.17: Axial slices of intermediate image volumes and reconstruction outputs obtained with a random orbit and LIDC data. Columns from left to right: phantom input; volume after deconvolution; CNN reconstruction; MBIR reconstruction. Rows from top to bottom: Axial slice; zoomed in ROI within the slice; axial slice of the difference image between the respective volume and GT.

Table 4.6: Evaluation metrics calculated between the reconstructions and GT. The results of the proposed reconstruction pipeline are reported for the tetrahedral data and for the LIDC data, where the network is pre-trained on tetrahedral data and then fine-tuned on anthropomorphic data. The better of two values is marked **bold**.

Proposed	Tetrahedrons	LIDC
nRMSE $\downarrow$	0.073 ± 0.016	0.060 ± 0.012
FSIM $\uparrow$	0.978 ± 0.017	0.988 ± 0.006
SSIM $\uparrow$	0.983 ± 0.007	$\mathbf{0.977 \pm 0.012}$

MBIR	Tetrahedrons	LIDC
nRMSE $\downarrow$	$\mathbf{0.033 \pm 0.006}$	$\mathbf{0.045 \pm 0.006}$
FSIM $\uparrow$	$\mathbf{0.992 \pm 0.005}$	$\mathbf{0.992 \pm 0.002}$
SSIM $\uparrow$	$\mathbf{0.986 \pm 0.006}$	0.969 ± 0.011

with different orbits. The network therefore needs to learn to handle data obtained with orbits it has not been previously trained on. In addition, the regularization in Fourier domain has been adapted to the complex rescaling method instead of the thresholding operation used for the application to data obtained with single orbits. To mitigate artifacts originating from the periodic boundary conditions - instead of increasing the size of the backprojected volumes during the calculation of the system response as performed above for single orbits - the impulse responses are zero-padded in Fourier domain, doubling the size of the volumes from 256 to 512 voxels before inverse Fourier transformation. Compared to the experiments for single orbits, the implementation of the proposed reconstruction routine was improved to yield faster computation times.

The intermediate (deconvolution only) images and the final reconstruction outputs are illustrated in Figure 4.16 for the reconstruction pipeline applied to tetrahedral data and random acquisition orbits. The same illustrations are depicted in Figure 4.17 for the application to anthropomorphic data and arbitrary geometries. Performance is quantified with the calculated evaluation metrics, shown in Table 4.6.

Notable differences between the deconvolved volumes and the phantoms can be observed, especially at the edges. After CNN processing, differences at the edges are substantially reduced. Similar observations can be made in the magnified regions, where the CNN manages to recover some of the fine detail from the blurry deconvolved volumes. Performance differences are also observed between the reconstruction of tetrahedral and anthropomorphic data. Specifically, the proposed reconstruction technique performed more accurately and consistently for anthropomorphic data.

In comparison, MBIR provides generally improved noise-suppression, better reconstructing fine details. While MBIR consistently outperforms the CNN-based approach in terms of nRMSE, FSIM and SSIM, performance differences are not particularly large with most differences appearing at the finest level of detail. In terms of consistency, MBIR performs comparably for all test cases, resulting in low standard deviations for the calculated evaluation metrics. Compared to the proposed reconstruction approach, MBIR performs more consistently. Also, MBIR performs notably superior on tetrahedral data than on anthropomorphic data.

There is substantial difference in computation time between the iterative reconstruction technique and the proposed reconstruction scheme. The MBIR needs twelve seconds of computation time per iteration, amounting to 20 minutes for 100 iterations. The proposed reconstruction approach took 20 seconds for the calculation of the system response. In combination with the backprojection (3 seconds), the calculation of the ray density (30 seconds), the deconvolution operation (20 seconds), and the CNN prediction (one second) this amounts to below 90 seconds of overall computation time.

To illustrate training convergence and check for differences in the loss metric between training and validation data sets, the network loss metric of both data sets is plotted logarithmically in Figure 4.15b for the network trained on tetrahedral data and anthropomorphic data. After 70 epochs, or 44800 steps, training is restarted with anthropomorphic data. There is no notable jump in the loss metric after 70 epochs, which indicates consistent performance regardless of the object to be reconstructed. As the decrease of the loss function of both data sets nearly reached a plateau after 70 epochs, there is only minor decrease in the loss metric after the switch to anthropomorphic data, indicating overall convergence.

5. Discussion

The main goal of this dissertation is to facilitate the implementation of arbitrary CBCT orbits in clinical applications. To better exploit the full potential of arbitrary acquisition orbits, two algorithmic steps are optimized in this work: 1) three different methods of simulating CT images are investigated to prospectively optimize imaging performance, and allow faster identification of viable alternatives to the standard circular trajectory; and 2) the investigation of novel methods for fast reconstruction of arbitrary CBCT orbits based on exploiting the speed of CNNs to reduce computation time.

In the past, it has been shown that the study of task-based detectability provides the most comprehensive approach to prospectively optimize the orbit geometry to achieve superior image quality, taking into account information about both the imaged object as well as the imaging task [21],[22]. This procedure can be applied to optimize performance for very general imaging scenarios, but is very time consuming as it exhausts a wide solution space. Alternatively, a smaller selection of orbits can be compared to explore some viable options for specific imaging scenarios in a shorter time. In addition, the detectability method has not yet been applied to cases, where interfering surgical equipment is blocking parts of the actuation/solution space. In combination with its computational complexity, these issues prevent the detectability approach from being implemented into routine clinical practice. Therefore, a faster approach to prospectively optimize acquisition geometries based on the available actuation space needs to be investigated and validated.

In addition to prospective optimization, time consuming image reconstruction techniques are another major limitation to the optimization and eventual implementation of arbitrary acquisition orbits in routine clinical practice. Analytically exact solutions are available for specific cases of geometries such as saddle trajectories and line-ellipse-line orbits [59],[118] but cannot be applied to a general class of orbits without modifications. Instead, time consuming MBIR is often performed to provide a best estimate based on the available data. Faster reconstruction techniques are needed to accommodate the time constraints in interventional imaging.

This chapter is comprised of two main sections. In the first section, three CT image simulation methods are evaluated and compared. First, a CNN-based approach in the form of a CycleGAN architecture is used to transform digital body phantoms directly into the CT image domain. A novel framework of already established and newly proposed evaluation metrics is developed to evaluate accurate anatomical features and realistic noise characteristics. Second, a MC approach is implemented and optimized for runtime to simulate arbitrary CBCT orbit projections using only publicly available algorithms and resources. Third, a projection operator algorithm is used in combination with an additive noise model to simulate projections of arbitrary CBCT orbit geometries. This is done to investigate a fast alternative while having a baseline state-of-the-art MC algorithm available for validation purposes.

In the second section, the results of the proposed fast reconstruction technique for non-circular CBCT geometries are discussed. For the first time - to the best of the author's knowledge - a deconvolution operation with the geometry-dependent system response in the backprojection domain is combined with a CNN post-processing step. Parts of the discussion were published in [85][1], [87], [86], [88] and [89] and the corresponding descriptions are adapted and partly replicated thereof.

5.1 CT Simulation and Evaluation

5.1.1 CT Synthesis Using CycleGAN

A CT synthesis framework based on the CycleGAN architecture is presented and evaluated using various already established as well as newly developed image quality metrics. The synthetic CT images are evaluated in terms of anatomical accuracy and realistic noise properties compared to the XCAT phantom input and reference CT images.

In the configuration study, the performance of eight CycleGANs are compared using the proposed image evaluation framework. Lower performance is found for networks using the U-Net architecture as generators, particularly for the MAE. Consequently, the U-Net architecture is considered inappropriate for the task of CT synthesis based on digital body phantoms. One possible explanation for the inferiority of the U-Net is that the skip connections of the U-Net do limit synthesis performance in this scenario. More research is needed to explore this hypothesis. The Res-Net architecture proves to be an ideal generator network, not only in the context of this work, but also during other studies carried out in our group on the synthesis of CT, CBCT and MRI images [119],[120],[121],[122].

The comparison in terms of the different number of input slices and the use of body contours yields favorable results for the networks trained with three input slices and without any pre-processing using the body contour. In particular, the realistic noise properties are accurately modeled. Apparently, it is not beneficial to show the network more than three input slices at a time to learn dependencies between slices. This is intuitive for the process of reproducing realistic noise properties, since there is no correlation of noise in the z-dimension for the MDCT images used in this study. However, more information from adjacent slices can potentially contribute to the recovery of anatomical features. In fact, the inclusion of additional information does not improve the synthesis performance in terms of anatomical accuracy. A possible explanation for this is that the network learns to synthesize anatomical features exclusively from the input XCAT phantom and not from the interdependences of anatomical structures in the training CT image, since the training samples are not matched. Based on these observations, the Res-Net architecture with three input slices and no body contour is considered for the subsequent task-based study, during which the training data set is extended and a task-based loss function is used for

---

[1]Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature International Journal of Computer Assisted Radiology and Surgery ©2019

training.

The comparison of three task-based networks shows substantial improvements in the form of more accurate vessel annotations, observed via the MAE_{ves} metric. These improvements are facilitated by the adaption of the loss function, which can be observed in a direct comparison of a network without (TASKNET1) and one with (TASKNET2) the use of this loss function. Moreover, the use of additional CT images from public sources as training data for the synthesis networks is investigated. The public source images contain a variety of noise textures due to varying scan properties such as the used reconstruction kernel. The results show that the scan-specific training with in-house CT images (TASKNET3), which also uses the modified loss function, performs better than the network trained with public CT images (TASKNET2). The characteristic noise is modeled most accurately and the XCAT anatomy is preserved most accurately by TASKNET3. This is due to the diversity of reconstruction kernels and thus noise texture in the training data of TASKNET2. Consequently, the training of scan specific networks and the adaption of the CycleGAN loss with a task-based loss function is recommended for future studies. The XCAT phantom provides segmentations for most anatomical structures, which allows the adaption of the loss function to many tasks, *i.e.* body regions or anatomical structures, such as liver contour preservation to generate training data for liver segmentation.

However, in all studies, the generation of arbitrary anatomical structures without any connection to the corresponding XCAT phantom is observed, which is clearly noticeable in the generated images as well as the EGR metric. Although CycleGAN is the fastest simulation method studied (one second per CT slice, *i.e.* around three to five minutes per CT image), this is a major shortcoming. Such arbitrary changes in anatomy can not be avoided by refining the network architecture or adjusting the loss function. Synthesis of MDCT images with this method is a first step before moving to standard CBCT images and eventually CBCT images obtained with non-circular acquisition orbits. However, the random generation of morphological structures in the simulated images substantially compromises their intended purpose, namely the evaluation of imaging performance, and is therefore considered to be impractical. Furthermore, to train CycleGANs for the synthesis of CBCT obtained with non-circular trajectories, training data in the form of images of patients and/or objects is required. A comprehensive data set is not yet available because the implementation of such orbits on clinical devices is an ongoing effort [24],[20],[35],[27]. Therefore, the synthesis of CBCT images acquired with arbitrary acquisition geometries using CycleGANs needs to be explored in future research.

Instead of using this method to simulate CBCT images to evaluate the image quality of different acquisition orbits, CT images are synthesized and subsequently used as training data for CT image segmentation tasks. In particular, a proof-of-principle experiment is performed using the synthetic CT images generated with TASKNET3 and real CT scans to train blood vessel segmentation networks. Performance evaluation with DSC revealed that training the segmentation network exclusively on

real data performed superior than training exclusively on synthetic data. However, combined training on both real and synthetic data showed the best performance, underlining the applicability of synthetic CTs as a form of data augmentation for segmentation networks.

5.1.2   CBCT Simulations using Monte-Carlo and Projection Operator Methods

Projection Simulations

The projection data are simulated and compared using two approaches that differ substantially in terms of physical precision and computational cost. At the high end of physical precision, an MC approach is used that is able to incorporate the relevant physical interaction processes in the image formation process, simulating each photon emitted from the X-ray source separately. In a first step, the bare beam fluence on the detector is simulated. Due to the flat shape of the detector and thus the different distances of the pixels to the source, the fluence decreases with increasing distance from the detector center. This effect can be accounted for by correcting with the inverse-square law. To this end, the counts in each pixel are multiplied by the ratio of the squared distance of the pixel from the source over the squared distance of the detector center from the source. For the following processing steps, each CBCT projection simulated with objects in the beam line is divided pixel-wise by the average value resulting from 20 realizations of the uncorrected bare-beam fluence.

In addition to MC simulations, a distance-driven projection operator is used to simulate CBCT projections. Here, each detector pixel value is the sum of the traversed image voxel values, weighted by a factor dependent on the distance of the virtual ray (from the source to the detector pixel) from the respective voxel center. The resulting line integrals are then converted to photon fluence values using the Lambert-Beer law. This allows the incorporation of a noise model based on the photon counts. Apart from this, no further refinements are made to imitate a realistic image formation process.

When comparing the two projection simulation approaches, there are noticeable beam hardening artifacts in the reconstructions obtained from the MC projection data, which are not present in the data obtained from the projection operator. In addition, the MC data exhibits minor ring artifacts around the tungsten wire that are not reproduced by the projection operator data. However, modeling this behavior is critical to ensure realistic prediction of the imaging performance for different orbits. Therefore, these are considered to be major constraints on the projection operator based CBCT simulations.

In the next step, 200 projections are simulated with each method for three different acquisition orbits, a standard circular orbit and two alternative trajectories, which would increase the flexibility of the CBCT system to maneuver the CBCT device

around potentially interfering surgical equipment. Compared with the MC projections, the projection operator simulations yielded a stable mean of attenuation values along the longitudinal axis of the cylinder phantom, as bare beam fluence is set to a constant value of 7800 for each detector pixel; this value is chosen to match the bare beam fluence values observed for the MC data. Furthermore, the fluence values behind the phantom range from 350 to 450 counts per pixel for both MC and projection operator simulations, indicating that the choice of phantoms and ultimately the simulation methods provide consistent results in terms of absolute numbers.

In the context of modeling a realistic image formation process during projection simulation, there are some limitations to the simulation methods studied. In the MC approach, the simulations can be closer to reality if accelerated electrons are selected as primary particles. In combination with the inclusion of the tungsten anode in the scenario, no assumptions need to be made on the number of photons or the expected X-ray spectrum. However, the use of SpekCalc data is considered to be sufficient for the intended purpose of the simulations since their data was validated using MC simulations. Additionally, the modeled X-ray spectrum does not include any pre-filtration that might be implemented in the simulated CBCT device. However, no detailed information about pre-filtration is publicly available for the modeled artis zeego system. Another limitation in the MC simulations is the lack of inclusion of the detector time resolution caused by the dead time of the detector crystals. Instead, the temporal independence of all incident photons on the detector is assumed. This leads to an overestimation of the number of detected photons, since sometimes multiple photons hit the sensitive area of the detector coincidentally, leading to a signal build-up and detection of only one event instead of the actual number of events. In reality, electronic noise introduces a Gaussian distribution around the expected number of photon counts in the detector, which is not modeled in the GATE simulations, but can be added to the simulated projection images in the future.

For the projection operator based simulations, a more realistic approach can incorporate the simulation of beam hardening artifacts. This can be done by considering the fluence distribution of the X-ray spectrum and setting up energy bins with their respective photon fluence. Due to the non-linear energy dependence of the attenuation coefficients of different materials at different energies, multiple realizations of the phantom need to be generated for different energy ranges. This approach can then merge mono-energetic projection simulations for different energy bins to obtain multi-energetic projections by using the fluence as weighting factors [123].

MTF Evaluation

The one-dimensional MTF is calculated from the reconstructed volumes using the oversampled LSF approach for all orbits and both simulation methods. In the case of circular orbits, the MTF calculated with this method should match in x- and y-direction due to the internal symmetry of the imaging system. However, in the case of non-circular orbits, the MTF in x-, y- and z-direction can deviate. For simplicity, the MTF is only considered and compared in the x-direction. A calculation of MTF

in y- and z-direction requires separate phantoms with wires slanted in orthogonal directions as well as separate time consuming GATE simulations.

To obtain the MTF, the first step is to calculate the COM for each slice, determining the point of maximum attenuation. A linear regression model is fitted to the COM data and the sub-voxel position of the tungsten wire is determined. However, pronounced noise levels in the reconstructed volumes around the wire resulted in some mislocalized COM positions. The goodness of fit indicated robust fitting performance and can confirm COM positions regardless of these outliers. Determining the position of the tungsten wire in each slice allowed to arrange all LSF curves in z-dimension. In this way, a combined LSF curve with a sub-voxel sampling rate of 10 samples per voxel is produced.

Finally, the oversampled LSFs are Fourier transformed to yield one-dimensional MTFs for all orbits and both simulation methods. This experiment is conducted to determine the most favorable alternative orbit for the case where a circular orbit is not feasible due to constraints on the available actuation range, for instance due to interfering surgical equipment. The predicted MTFs indicate the best imaging performance for a circular orbit with a FWHM of $0.667 \, \text{mm}^{-1}$ using GATE simulations and $1.101 \, \text{mm}^{-1}$ with the projection operator. Both simulation approaches predict the second-best performance for the tilted circular orbit (FWHM=$0.600 \, \text{mm}^{-1}$, MPE =$15.1 \, \%$ with GATE, and FWHM=$0.881 \, \text{mm}^{-1}$, MPE =$6.5 \, \%$ with projection operator), favoring it over the triple-arc trajectory (FWHM=0.534, MPE =$20.3 \, \%$ with GATE, and FWHM=0.561, MPE =$30.4 \, \%$ with projection operator). This difference in performance is more distinct in the projection operator based data, which generally predicted better performance than the GATE data. This is intuitive because the GATE approach is able to model beam hardening and scattering artifacts that result in poorer image quality in the reconstructed images. The MTF evaluations clearly show that the tilted circular orbit is a better alternative to the standard circular orbit when compared to the triple-arc trajectory.

The presented method of oversampling an LSF to calculate the one dimensional MTF for the prediction of imaging performance is subject to certain limitations. A major limitation is the high level of noise in the vicinity of the tungsten wire. This considerably hinders accurate calculation of the MTF. In future simulations, this can be avoided by substituting the water surrounding the tungsten wire with air. Furthermore, in CBCT systems, the divergent beam geometry, insufficient sampling, etc. lead to a system response that is not shift-invariant, i.e. imaging performance varies over different image regions. Spatially varying results are to be expected when evaluating MTFs. As a consequence, imaging performance must be evaluated at various locations in the image, and not only in the center regions, as is the case in the presented study.

Furthermore, the used approach of MTF calculation only yields one-dimensional data, while a 3-dimensional approach is necessary to capture the imaging performance more comprehensively. A possible solution to the mentioned limitations might

be the calculation of the 3D PSF, based on locally distributed impulses in a water cylinder. This allows the assessment of the extent of shift-variance as well as provide 3D information. However, this method is less accurate than the oversampled LSF approach due to the use of less densely sampled PSFs. The MTF calculation has recently been found to be insufficient for capturing task-specific imaging performance of non-linear imaging systems [124]. An imaging system is considered to be linear when the imaging output from a weighted sum of input signals is equal to the weighted sum of the output from each individual input. An example for a non-linear imaging system is a CT system incorporating an iterative reconstruction technique. In these non-linear cases, imaging performance depends on the imaged signal, for instance on its contrast. Therefore, a major step towards task-based imaging prediction for non-linear systems is the calculation of the task transfer function (TTF), which is obtained by computing the MTF from signals that resemble the actual imaging task. For example, to predict task-based imaging performance in the case of a biopsy needle localization task, high-contrast LSFs must be used (such as in the presented method), while for tumor localization, low-contrast spheres must be used instead to obtain the TTF.

By using GateLab to access the European Grid Infrastructure, the GATE simulation time can be substantially reduced by a factor of 1200. Since the Astra-based projection simulations take seven minutes of simulation time - about a thousandth of the GateLab calculation time - this method is considered more favorable to select adequate orbit designs. Reducing the computation time for prospective orbit optimization is particularly important given: 1) The abundance of recently proposed orbits and thus the extent of the search space; and 2) the application of this approach in an intra-interventional imaging context in the form of real time prediction of imaging performance, when time is of the essence. Compared to the time needed for projection simulation, however, the 60 minutes required for 100 iterations of SIRT reconstruction still make up most of the total computation time and underline the need for a faster reconstruction technique in this context. Further reduction in calculation time of the projection operator simulations can be achieved by omitting the noise model, which takes $98\,\%$ of the calculation time. However, the inclusion of this noise model in the projections is essential when comparing the imaging performance of various orbit designs, as it plays an important role when predicting imaging performance.

In summary, the presented method of prospective CBCT simulation and evaluation of a selection of orbit designs can be used to predict imaging performance and identify adequate alternatives to standard circular orbits. Comparing the outcome of both simulation methods, both lead to identical conclusions regarding orbit selection. This means that, notwithstanding the observed issues arising from inadequate modeling of the underlying physical processes, the projection operator-based method provides accurate simulation results in the cases studied. It is not only accurate, but also fast enough to provide real-time orbit solutions in the future. However, further validation in the form of more imaging scenarios and more orbits studied is required for the use of this optimization method in routine clinical practice. A state-of-the-

art MC approach is presented, which can be used for validation purposes in more comprehensive studies. In addition, recently proposed orbit optimization schemes in literature need to be included in such studies for comparison. At the same time, real measurements on clinical CBCT devices are needed to verify the simulation results. Due to their limited scope, the presented experiments can only be considered as proof-of-concept.

Dose Deposition

Aside from the simulation of projection data, the capability of the GATE framework to calculate the deposited dose is demonstrated. This is achieved by tracking individual photon capture and scattering events in the digital XCAT phantom, which is incorporated in the simulation scenario. The FOV of a CBCT resulting from a circular orbit is clearly noticeable as a distinct conical dose distribution inside the phantom, where most photon interactions are happening. Furthermore, deposited dose is observed to increase closer to the phantom surface as well as in bones. The former can be explained by the fact that the position of the maximum dose deposition of photons in matter is close to the object surface. The deposited dose is higher in bones because of the Z-dependency (Equation 2.26) of the photoelectric effect, which is the predominant interaction mechanism in the lower diagnostic photon energy range. Bones consist of a matrix of hydroxyapatite, which is a compound of calcium (Z=20) and phosphorus (Z=15), both high Z elements compared to elements most prevalent in body tissue such as oxygen (Z=8) or carbon (Z=6). The higher electron density of bones further increases photon cross-section for Rayleigh, Compton scattering as well as for the photoelectric effect.

In future research, this framework can be used to investigate orbit dependent dose distributions and identify favorable trajectories in terms of choosing more tolerable scenarios. When extending this approach to investigate dose deposited to the surroundings, orbits can be selected to spare medical personnel that might be present next to the patient during the imaging procedure. Avoiding dose exposure is particularly critical for medical personnel, as they are exposed repeatedly when conducting interventional procedures which - under certain circumstances - require them to be present in the room during imaging.

5.2   Fast Reconstruction of Arbitrary CBCT Orbits Using CNNs

In this part, the experimental results of the presented reconstruction pipeline are discussed. Fundamentally, the reconstruction scheme consists of an approximate deconvolution operation with the geometry-dependent system response followed by CNN processing to correct for any residual artifacts. Initially, different means of deconvolution regularization in Fourier domain are evaluated and compared in Section 5.2.1. In a first study in Section 5.2.2, the proposed pipeline is applied to single orbit geometries to investigate the approach's capability to reconstruct arbitrary CBCT orbits and to generalize to handle more than one orbit. During the second

study in Section 5.2.3, the reconstruction algorithm is applied to a general class of orbits to investigate its generalizability to a wide range of different orbits.

5.2.1 Deconvolution Regularization

It is well known, that image deconvolution operations performed with discrete Fourier transformations are sensitive to various forms of artifacts, which need to be adequately corrected [125],[126],[127],[128],[129]. This is mainly due to the periodic boundary conditions of the fast Fourier transform but also due to the overemphasis of high frequency noise when deconvolving with a function that approaches zero at its edges. To mitigate artifacts originating from the boundary value problem, the impulse responses in Fourier domain is zero-padded to yield volumes with double the edge length. This also helped to reduce computation time and further reduce artifacts during the random orbit experiments.

To address further issues during image reconstruction, three methods of regularization in Fourier domain are evaluated. To keep computational complexity as low as possible, the regularization approaches need to be simplistic. First, a quadratic penalty function, which is frequently applied to suppress high-frequency noise in CT image reconstruction [130], is investigated. During a parameter sweep of the regularization strength for this method, however, substantial artifacts persisted throughout the deconvolved image volume regardless of regularization strength. These artifacts appeared as stripes of varying frequency as well as grid artifacts, depending on the regularization strength. With increasing regularization strength, a decrease in the frequency of the stripe artifacts is observed. The extent of these artifacts does not reach acceptable levels, because the high frequency content such as edges is increasingly suppressed. Consequently, the quadratic penalty approach investigated here is considered inadequate because no ideal trade-off value for the regularization strength can be found.

In addition, two types of thresholding regularizations are investigated. The first approach consists of setting each value below the threshold to the threshold value. This procedure produced visually superior outcomes than the quadratic penalty regularization. A point of ideal trade-off between the sagittal stripe artifacts, which are increasingly prominent for lower regularization strength, and the loss of edge content when using higher regularization strength is found. For this reason the method is utilized as a regularization scheme for the application of the reconstruction approach to single orbit geometries. However, it is also observed that this technique set the imaginary part of the Fourier transform to zero for the thresholded values. As a consequence, phase information is lost. A method preserving phase information is expected to produce superior results.

To preserve phase information, a rescaling operation is investigated, which consisted of rescaling the real and imaginary part of all voxels with a modulus below the threshold to yield this threshold value. While this approach resulted in grid artifacts and sagittal stripes for low regularization strength, a parameter sweep is able to

reveal an ideal trade-off point when increasing regularization strength before too much edge content is lost. Consequently, this approach is utilized for the application to arbitrary orbit geometries.

5.2.2   Reconstruction of Single/Dual Geometries

The experiments performed with single orbit geometries symbolize an initial validation step for the proposed reconstruction scheme. Circular and sinusoidal orbits as well as one linear combination of sinusoidal basis functions with randomly generated coefficients are investigated. The calculated evaluation metrics show that the proposed reconstruction pipeline is able to perform robustly for each orbit geometry. Compared to a standard SART reconstruction, the proposed reconstruction pipeline performs superior, regardless of the orbit geometry. However, a slight decrease in performance for increasing orbit complexity is observed (order of orbit complexity: circular < sinusoidal (even frequency) < sinusoidal (odd frequency) < random). The consistency of the reconstruction performance of the proposed reconstruction approach is also higher than for SART, whilst also slightly decreasing with orbit complexity. At the same time, consistency remained on the same level for SART. This slight decrease in reconstruction performance and consistency of the proposed algorithm is due to the assumption of a shift-invariant system response in the deconvolution operation. This assumption becomes less and less accurate with higher orbit complexity as the system response becomes increasingly more shift-variant. This leads to an increasing extent of artifacts in the deconvolved volumes. The posterior CNN processing step is not able to fully correct these artifacts.

In terms of deconvolution artifacts, there are substantial deviations for different orbit geometries. In the case of the circular and the sinusoidal orbit with even-valued frequencies, there are no apparent artifacts in the central-most slices. This is because the symmetric system responses of even-numbered sinusoidal functions meet the requirements for periodic boundary conditions during Fourier transformations. The artifacts arising from asymmetric system responses can not be corrected by the chosen regularization methods and can be partly addressed by quadrupling the voxel number in each dimension of the system response. For this reason, a different form of regularization is chosen for the following studies with arbitrary geometries to suppress these artifacts more consistently.

Apart from deficiencies in the deconvolution operation, a limitation of the previous experiments is the use of piecewise-constant phantoms. A possible implication of exclusively using piecewise-constant phantoms is that the network, instead of learning to recover high-frequency content and correcting residual artifacts, might learn to segment the constant-valued tetrahedrons. This shortcoming is addressed in the following experiments with the inclusion of Gaussian peaks in the tetrahedral phantoms as well as the inclusion of anthropomorphic training data. Furthermore, the initial study does not include a noise model, which is added in the experiments for random orbits.

Summarizing the experiments with single/dual orbits, the proposed reconstruction pipeline outperforms an iterative reconstruction algorithm whilst reducing the calculation time by approximately 90 %. This is shown with networks that have been trained on specific cases of one or two orbits. Note that both algorithmic implementations are evaluated on a common hardware platform without fine-tuning or optimization for runtime. However, the relative computation times are expected to be consistent with additional optimization. In addition, the robust performance of the network trained on data from two different sinusoidal orbits underlined the potential of the algorithm to generalize to deconvolution data obtained from more than one orbit.

5.2.3   Reconstruction of Random Geometries

The following experiments extend the proposed reconstruction algorithm and address the issues observed during the investigations above. Specifically, the following improvements are made: 1) Inclusion of a noise model to the projection data; 2) Incorporation of non-piecewise constant tetrahedral as well as anthropomorphic training data; 3) Complex rescaling regularization in Fourier domain; 4) Zero padding at the edges of the system response volume; 5) Increase to 1800 unique random orbit geometries, one for each set of projections. Due to these changes, the following experiments are a major step towards applying the proposed reconstruction pipeline to real clinical data. In addition, the reconstructions are compared to a state-of-the-art MBIR algorithm instead of SART.

The deconvolved volumes show less artifacts than during the experiments with single orbit geometries due to the changes discussed above. However, the difference images of the deconvolved volume to GT phantom reveal distinct deviations, particularly at high contrast edge locations. This is due to a position mismatch because the deconvolution operation is performed on odd-valued matrices and the original phantom matrix is even-valued. This means that in contrast to the original phantom images, the deconvolved volumes are translated about half a voxel edge length from the original image center. This mismatch is easily corrected by the posterior CNN processing step. However, this limits the performance of the CNN - as some capacity has to go into the inverse translational operation - but might also lead to local minima in the loss function. Therefore, this mismatch needs to be corrected for in future studies. This can be done by having both GT and deconvolved volumes with odd-valued matrix sizes and then appending one zero slice in all dimensions prior to CNN processing.

When inspecting the loss function over the course of the training procedure, the accuracies of the training and validation data sets decrease at the same rate, which is an indication that the network did not overfit to the training data. If this were the case, the accuracy for the training data set would be considerably higher than for the validation data set. In such a case, the network would learn representations or characteristics specific for the training data and would not be able to generalize to new data. When fine-tuning the network, which has been trained with tetrahedral data for 70 epochs, with the anthropomorphic data, no accuracy deterioration at

the 70 epoch mark of the loss function is noticeable. This implies that the network has learned to correct the residual artifacts originating from tetrahedral data, but can also abstract what it had learned to new sorts of training data. It is therefore concluded that the types of phantoms as well as the splitting of epochs into pretraining and fine-tuning were adequately chosen.

The calculated evaluation metrics for 10 new test cases - data simulated for orbits the networks has not trained on - indicate slightly worse performance and consistency for the proposed method when compared to the MBIR reconstructions. This can also be observed in the reconstructed images, especially in the magnified regions, where the proposed routine is not able to recover high frequency content accurately. However, the proposed reconstruction pipeline performs robustly over all orbit geometries with an nRMSE of 0.073 and 0.060 for tetrahedral and anthropomorphic data, respectively, compared to 0.033 and 0.045 for MBIR. This slight decrease in performance in comparison to MBIR is acceptable especially when considering the reduction of calculation time from 20 minutes with MBIR down to less than 90 seconds with the proposed method. The same decrease in performance can also be observed in the calculated SSIM and FSIM metrics. The reduction in computational complexity allows the application of the proposed reconstruction pipeline for various intra-operational tasks. First, the fast availability of a reconstructed image for the surgeon, which would otherwise take substantially more time when using MBIR. Second, the proposed method can provide a prior image for the MBIR routine to speed up convergence of the iterative reconstruction procedure. Third, the fast reconstruction procedure can be used in combination with the above evaluated method for simulation and evaluation of arbitrary geometry orbits. When omitting noise simulation in the projections, one can obtain a prediction of image quality and a reconstructed image in less than 2 minutes (10 seconds simulation, 90 seconds reconstruction and 15 seconds evaluation) per orbit. The calculation time can be further reduced if the orbit geometry is known in advance. In such a case, the system response and the corresponding ray density can be pre-computed; this leads to a computation time below 40 seconds.

In this work, random orbits are generated as combinations of nine periodic basis functions with a frequency of up to four. This is somewhat limiting the generalizability of the trained reconstruction networks. An increase in the number of basis functions can increase orbit arbitrariness in future studies. Alternatively, arbitrary orbits can be parametrized by randomly sampling a set of vertex locations and subsequently connecting them using B-spline interpolation. Another possible extension of the presented reconstruction pipeline is the inclusion of more classes of orbits. Random combination of arcs or other forms of non-continuous orbits can be explored.

Further studies are needed to validate the presented reconstruction routine for clinical applications. A next step is the application to real projection data. In its current state, no truncation is present in the simulated projections, which needs to be included in the simulation data before moving to real data. The application to real

projections should start with the reconstruction of data obtained with circular orbits to reveal any issues before moving to more complex geometries. Another important step for evaluating non-linear imaging systems - like the ones using neural networks for reconstruction - is the analysis of image properties. In imaging systems where the linear approximation still holds true, conventional image contrast metrics such as the edge-spread-function can be used. However, for non-linear systems, new evaluation methods such as perturbation response analysis need to be applied to fully capture imaging performance in the future studies [131].

Future research should also investigate new methods to move to a deconvolution operation that captures the shift-variance of the system response of arbitrary orbits. While the performance of the proposed reconstruction pipeline is consistent for more complex geometries, a further increase in performance is expected when moving to shift-variant deconvolution kernels. This is expected to help removing more of the geometry-dependent artifacts in the deconvolved volumes. A possible implementation is to move from a deconvolution in Fourier domain to an image domain-based convolution with spatially varying kernels. Such kernels need to be calculated beforehand by forward- and backprojection of impulses localized in different regions of the image volume, for instance. Another advantage of this technique is the elimination of artifacts originating from the periodic boundary conditions during Fourier transformation. In this work, matrices consisting of 128 voxels in each dimension are utilized while in standard CBCT imaging, 512 voxel edge lengths are preferred. A lower extent of artifacts in the deconvolved volumes leads to lower expectations on the capabilities of the CNN. Therefore, more shallow networks can be deployed; this can help to save GPU memory to accommodate larger matrix sizes instead.

6. Conclusion and Outlook

The two main objectives of this thesis were (1) to investigate methods to simulate CBCT images acquired with arbitrary acquisition orbits, and prospectively optimize the orbit geometry by predicting the associated imaging performance; and (2) to develop a fast reconstruction routine for non-circular CBCT source orbits. These goals have been achieved in the two parts of this thesis.

In the first part, three different means of CT simulation were investigated with the goal of predicting imaging performance. The first investigated method uses CNNs in a CycleGAN architecture to synthesize CT images from digital XCAT phantoms. The network architecture was optimized by introducing an evaluation framework, which consisted of multiple metrics to determine anatomical accuracy and realistic noise properties. The CycleGAN was the fastest investigated simulation method, with around three to five minutes computation time per CT image. However, artificially generated anatomical structures were observed in the simulated CT images, which did not correspond to any morphology in the input XCAT phantom. For this reason, the CycleGAN method was considered to be impractical for CT simulation and was therefore excluded from the subsequent image quality prediction studies. Instead, the use of CT images simulated with CycleGAN as training data for artery segmentation networks was shown to improve the performance of these networks; this highlights the applicability of the CycleGAN approach as an additional method of data augmentation in cases where clinical training data are scarce. Future investigations of the CycleGAN method for simulating CBCT with arbitrary acquisition orbits will require training data in the form of measurements and clinical patient images.

The second investigated method uses GATE, which is based on GEANT4 code, to simulate CBCT projections acquired with arbitrary acquisition orbits via the MC method. When executed on the Gatelab cluster, a computation time of around eight days was achieved for 200 projections. Due to the computation time, this approach in its current form is not suitable for clinical practice. Since MC simulations are considered the current gold standard for modeling the physical interactions underlying the stochastic process of CBCT image formation, the MC projections served as validation images for the third studied simulation method. Here, a distance-driven projection operator is combined with a noise model to simulate CBCT projections. This approach achieved a substantially lower execution time of seven minutes per set of projections. However, several important imaging phenomena such as beam hardening and photon scattering - which are simulated by the MC method - are not reproduced.

Both projection simulation methods were used to determine the better of two alternative geometries, emulating a clinical imaging scenario in which a circular orbit is not possible. This was done by calculating the MTF in CBCT images reconstructed from projections, which were simulated for a tungsten wire phantom. Both simulation methods identified the circular orbit as the best-performing with a MTF FWHM

of $0.667\,\mathrm{mm}^{-1}$ using GATE simulations and $1.101\,\mathrm{mm}^{-1}$ with the projection operator and also agreed on the best alternative: a tilted circular instead of a triple-arc orbit. This agreement highlights the capability of the projection operator approach to predict the CBCT imaging performance based on the acquisition orbit. However, to further validate the robustness of the proposed simulation and evaluation framework, more orbits need to be considered. Moreover, the obtained results need to be verified by real measurements in the future. While only requiring seven minutes for the simulations of the projections and 15 seconds for the MTF evaluation, most computation time (60 minutes) was used for the reconstruction, underlining the need for a faster reconstruction approach. In summary, the presented method provides a fast but robust approach to determine suitable alternatives to the standard circular orbit.

In the second part, a fast reconstruction scheme for non-circular CBCT acquisition orbits was presented. The pipeline consists of a deconvolution operation of the backprojected data with an approximated system response and a posterior CNN processing step to remove residual artifacts. The proposed approach is first applied to single geometries and then trained and tested on projection data simulated with 1000 different orbits using the projection operator-based method evaluated in the first part of this thesis. Non-piecewise constant training and evaluation data are created via the addition of low frequency noise to the procedurally generated tetrahedral phantoms in the form of multiple Gaussian peaks. Furthermore, anthropomorphic data in the form of 800 CT and CBCT images from the LIDC is incorporated in training and testing.

While penalized-likelihood reconstructions perform slightly superior in terms of nRMSE, FSIM and SSIM, the proposed algorithm can reduce the computation time by $90\beta\,\%$, namely from 20 minutes for MBIR to under 90 seconds for the proposed method. In case the geometry is known in advance, pre-computations can reduce the reconstruction time to below 40 seconds. The proposed approach was found to perform better on the LIDC data than on the tetrahedral phantoms, which may be due to more structural similarities in these anthropomorphic data. Consequently, caution is advised when training such networks, and such a behavior underscores the need for diversity in the training data to obtain robust performance. Nonetheless, these results demonstrate the applicability of the reconstruction pipeline to a broad class of orbits and indicate that the network can robustly reconstruct data from orbits on which it has not been trained. Furthermore, these results represent an important step toward providing a fast alternative to state-of-the-art reconstruction techniques for interventional applications.

In future research, some limitations of the presented reconstruction algorithm need to be addressed. First, the proposed algorithm needs to be applied to physical projection measurements acquired in real imaging settings. In particular, projection data obtained on an experimental test-bench can provide further validation. Furthermore, image-based deconvolution strategies with shift-variant kernels for specific image regions will be investigated to potentially remove the artifacts resulting from the Fourier domain approach. By incorporating shift-variant kernels, this can help

to reduce the dependence of the deconvolved data on the acquisition orbit and improve reconstruction performance. Moreover, shift-variant processing using patch sub-volumes during the CNN step allows processing of larger matrix sizes than the limited volume sizes (128x128x128) used in this work.

The presented work provides a framework to predict the orbit-dependent CBCT image quality, consisting of the simulation, evaluation and reconstruction of CBCT images acquired with arbitrary acquisition trajectories. In the first part, a projection operator-based approach is used to determine the best alternative to a standard circular orbit; the approach is verified using MC simulations. In the second part, the speed of CNNs is exploited to reconstruct projection data from arbitrary source orbits within 90 seconds, while achieving image quality comparable to state-of-the-art reconstructions. All computation steps can be performed in less than ten minutes per CBCT image, when the algorithms developed in the two parts of this thesis are combined. Consequently, the presented work is a promising step toward the implementation of arbitrary CBCT orbits for interventional clinical practice.

List of Publications

Publications Used in this Thesis

- **T. Russ**, W. Wang, A.-K. Golla, D. F. Bauer, M. Tivnan, C. Tönnes, 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*, accepted for oral presentation, San Diego, USA, 2022.
- **T. Russ**, W. Wang, A.-K. Golla, D. F. Bauer, M. Tivnan, C. Tönnes, Y. W. Ma, T. Reynolds, S. Hatamikia, L. R. Schad, F. G. Zöllner, G. J. Gang, and J. W. Stayman, "Fast reconstruction of non-circular CBCT orbits using CNNs," *Proceedings of the 16th Virtual International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*, Virtual, 2021.
- **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," *European Congress of Radiology*, Virtual, 2020.
- **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," *4th Conference on Image-Guided Interventions*, Mannheim, DEU, 2019.
- **T. Russ**, S. Goerttler, A.-K. Schnurr, D. 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*, vol. 14, no. 10, pp. 1741-1750, 2019.

Journal Papers

- D. F. Bauer, J. Rosenkranz, A.-K. Golla, C. Tönnes, 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," *Medical Physics*, under review, 2021.
- 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 (2021), "An anthropomorphic pelvis phantom for MR-guided prostate interventions," *Magn Reson Med*, vol. 00, pp. 1– 8, 2021.
- A.-K. Golla, C. Tönnes, **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*, vol. 11, no. 11, p. 2131, 2021.

- S. Hatamikia , A. Biguri, G Kronreif, M. Figl, **T. Russ**, J. Kettenbach, M. Buschmann and W. Birkfellner, "Toward on-the-fly trajectory optimization for C-arm CBCT under strong kinematic constraints," *PLoS ONE*, vol. 16, no. 2, 2021.
- D. F. Bauer, **T. Russ**, B. I. Waldkirch, 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*, vol. 16, pp. 1277–1285, 2021.
- A.-K. Golla, D. F. Bauer, R. Schmidt, **T. Russ**, D. Nörenberg, K. Chung, C. Tönnes, 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*, vol. 68, no. 5, pp. 1518-1526, 2021.
- S. Hatamikia, A. Biguri, G. Kronreif, J. Kettenbach, **T. Russ**, H. Furtado, L. K. Shiyam Sundar, M. Buschmann, E. Unger, M. Figl, D. Georg, and W. Birkfellner, "Optimization for customized trajectories in cone beam computed tomography," *Medical Physics*, vol. 47, pp. 4786-4799, 2020.
- A.-K. Schnurr, K. Chung, **T. Russ**, L. R. Schad and F. G. Zöllner, "Simulation-based deep artifact correction with convolutional neural networks for limited angle artifacts," Zeitschrift für Medizinische Physik, vol. 29, no. 2, pp. 150-161, 2019.

Conference Contributions

- D. F. Bauer, C. Ulrich, A.-K. Golla, **T. Russ**, C. Tönnes, 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, DEU, 2021.
- S. Hatamikia, A. Biguri, G. Kronreif, **T. Russ**, J. Kettenbach, and W. Birkfellner, "Collision avoidance trajectories for on-line trajectory optimization in C-arm CBCT," *Proceedings of the 16th Virtual International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*, Virtual, 2021.
- Y. Q. Ma, T. Reynolds, G. J. Gang, O. Dillon, **T. Russ**, W. Wang, T. Ehtiati, C. Weiss, N. Theodore, J. Siewerdsen, R. O'Brien, and J. W. Stayman, "Non-circular orbits on a clinical robotic C-arm for reducing metal artifacts in orthopedic interventions," *AAPM 63rd Annual Meeting, John R. Cameron Early-Career Investigator Award Winner*, Virtual, 2021.
- T. Reynolds, Y. Q. Ma, G. J. Gang, O. Dillon, **T. Russ**, W. Wang, T. Ehtiati, C. Weiss, N. Theodore, J. Siewerdsen, R. O'Brien, J. W. Stayman, "Imaging from the cervical to the lumbar spine with a continuous multi-turn reverse helical 3D cone-beam CT scan," *AAPM 63rd Annual Meeting, Jack Fowler Early-Career Investigator Award Winner*, Virtual, 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," *Proceedings ISMRM Congress*, Virtual, 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," *Proceedings ESMRMB Congress*, Virtual, 2020.
- S. Hatamikia, A. Biguri, G. Kronreif, **T. Russ**, J. Kettenbach, and W. Birkfellner, "Short scan source-detector trajectories for target-based CBCT," *42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Virtual, 2020.
- D. 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," *Proceedings International Conference on Medical Imaging with Deep Learning*, London, GBR, 2019.
- D. 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," *4th Conference on Image-Guided Interventions*, Mannheim, DEU, 2019.
- G. Kabelitz, D. 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," *4th Conference on Image-Guided Interventions*, Mannheim, DEU, 2019.

Supervised Theses

- A. M. Abdelrehim, "A Monte Carlo simulation framework for modelling Artis Zeego CBCT using GATE," Master's thesis, 2019.

Bibliography

[1] M. J. Daly, J. H. Siewerdsen, D. J. Moseley, D. A. Jaffray, and J. C. Irish, "Intraoperative cone-beam CT for guidance of head and neck surgery: Assessment of dose and image quality using a C-arm prototype," *Medical Physics*, vol. 33, no. 10, pp. 3767–3780, 2006.   (cited on Page 1)

[2] C. Floridi, A. Radaelli, N. Abi-Jaoudeh, M. Grass, M. De Lin, M. Chiaradia, J. F. Geschwind, H. Kobeiter, E. Squillaci, G. Maleux, A. Giovagnoni, L. Brunese, B. Wood, G. Carrafiello, and A. Rotondo, "C-arm cone-beam computed tomography in interventional oncology: Technical aspects and clinical applications," *Radiologia Medica*, vol. 119, no. 7, pp. 521–532, 2014.   (cited on Page 1)

[3] C. Czerny, K. Eichler, Y. Croissant, B. Schulz, G. Kronreif, R. Schmidt, M. Von Roden, C. Schomerus, T. J. Vogl, I. Marzi, and S. Zangos, "Combining C-arm CT with a new remote operated positioning and guidance system for guidance of minimally invasive spine interventions," *Journal of NeuroInterventional Surgery*, vol. 7, no. 4, pp. 303–308, 2015.   (cited on Page 1)

[4] D. A. Jaffray, J. H. Siewerdsen, J. W. Wong, and A. A. Martinez, "Flat-panel cone-beam computed tomography for image-guided radiation therapy," *International Journal of Radiation Oncology Biology Physics*, vol. 53, no. 5, pp. 1337–1349, 2002.   (cited on Page 1)

[5] I. S. Grills, G. Hugo, L. L. Kestin, A. P. Galerani, K. K. Chao, J. Wloch, and D. Yan, "Image-guided radiotherapy via daily online cone-beam CT substantially reduces margin requirements for stereotactic lung radiotherapy," *International Journal of Radiation Oncology Biology Physics*, vol. 70, no. 4, pp. 1045–1056, 2008.   (cited on Page 1)

[6] W. Zbijewski, P. De Jean, P. Prakash, Y. Ding, J. W. Stayman, N. Packard, R. Senn, D. Yang, J. Yorkston, A. MacHado, J. A. Carrino, and J. H. Siewerdsen, "A dedicated cone-beam CT system for musculoskeletal extremities imaging: Design, optimization, and initial performance characterization," *Medical Physics*, vol. 38, no. 8, pp. 4700–4713, 2011.   (cited on Page 1)

[7] A. O'Connell, D. L. Conover, Y. Zhang, P. Seifert, W. Logan-Young, C. F. L. Lin, L. Sahler, and R. Ning, "Cone-beam CT for breast imaging: Radiation dose, breast coverage, and image quality," *American Journal of Roentgenology*, vol. 195, no. 2, pp. 496–509, 2010.   (cited on Page 1)

[8] A. C. Miracle and S. K. Mukherji, "Conebeam CT of the head and neck, part 2: Clinical applications," *American Journal of Neuroradiology*, vol. 30, no. 7, pp. 1285–1292, 2009.   (cited on Page 1)

[9] T. Kiljunen, T. Kaasalainen, A. Suomalainen, and M. Kortesniemi, "Dental cone beam CT: A review," *Physica Medica*, vol. 31, no. 8, pp. 844–860, 2015.   (cited on Page 1)

[10] Statistisches Bundesamt, "Operationen und Prozeduren der vollstationären Patientinnen und Patienten in Krankenhäusern (4-Steller) - 2013," 2014.   (cited on Page 1)

[11] Statistisches Bundesamt, "Operationen und Prozeduren der vollstationären Patientinnen und Patienten in Krankenhäusern (4-Steller) - 2020," 2021.   (cited on Page 1)

[12] M. Diana and J. Marescaux, "Robotic surgery," *British Journal of Surgery*, vol. 102, no. 2, pp. 15–28, 2015.   (cited on Page 1)

[13] A. Brodie and N. Vasdev, "The future of robotic surgery," *Annals of the Royal College of Surgeons of England*, vol. 100, pp. 4–13, 2018.   (cited on Page 1)

[14] F. Cornelis, H. Takaki, M. Laskhmanan, J. C. Durack, J. P. Erinjeri, G. I. Getrajdman, M. Maybody, C. T. Sofocleous, S. B. Solomon, and G. Srimathveeravalli, "Comparison of CT fluoroscopy-guided manual and CT-guided robotic positioning system for in vivo needle placements in swine liver," *CardioVascular and Interventional Radiology*, vol. 38, no. 5, pp. 1252–1260, 2015.   (cited on Page 1)

[15] V. Solomiichuk, J. Fleischhammer, G. Molliqaj, J. Warda, A. Alaid, K. von Eckardstein, K. Schaller, E. Tessitore, V. Rohde, and B. Schatlo, "Robotic versus fluoroscopy-guided pedicle screw insertion for metastatic spinal disease: A matched-cohort comparison," *Neurosurgical Focus*, vol. 42, no. 5, pp. 1–7, 2017.   (cited on Page 1)

[16] M. Unger, J. Berger, and A. Melzer, "Robot-assisted image-guided interventions," *Frontiers in Robotics and AI*, vol. 8, pp. 1–7, 2021.   (cited on Page 1)

[17] S. Hatamikia, A. Biguri, G. Kronreif, M. Figl, T. Russ, J. Kettenbach, M. Buschmann, and W. Birkfellner, "Toward on-the-fly trajectory optimization for C-arm CBCT under strong kinematic constraints," *PLoS ONE*, vol. 16, no. 2, pp. 1–17, 2021.   (cited on Page 1 and 2)

[18] P. Wu, N. Sheth, A. Sisniega, A. Uneri, R. Han, R. Vijayan, P. Vagdargi, B. Kreher, H. Kunze, G. Kleinszig, S. Vogt, S. F. Lo, N. Theodore, and J. H. Siewerdsen, "C-arm orbits for metal artifact avoidance (MAA) in cone-beam CT," *Physics in Medicine & Biology*, vol. 65, no. 16, p. 165012, 2020.   (cited on Page 1)

[19] F. E. Boas and D. Fleischmann, "CT artifacts: Causes and reduction techniques," *Imaging Med.*, vol. 4, no. 2, pp. 229–40, 2012.   (cited on Page 1)

[20] G. J. Gang, T. Russ, Y. Ma, C. Toennes, J. H. Siewerdsen, L. R. Schad, and J. W. Stayman, "Metal-tolerant noncircular orbit design and implementation on robotic C-arm systems," *International Conference on Image Formation in X-Ray Computed Tomography*, 2020.   (cited on Page 1, 2, and 77)

[21] J. W. Stayman, S. Capostagno, G. J. Gang, and J. H. Siewerdsen, "Task-driven source–detector trajectories in cone-beam computed tomography: I. Theory and methods," *Journal of Medical Imaging*, vol. 6, no. 2, p. 1, 2019.   (cited on Page 1, 2, 4, 50, 63, and 75)

[22] S. Capostagno, J. W. Stayman, M. Jacobson, T. Ehtiati, C. R. Weiss, and J. H. Siewerdsen, "Task-driven source–detector trajectories in cone-beam computed tomography: II. Application to neuroradiology," *Journal of Medical Imaging*, vol. 6, no. 02, p. 1, 2019.   (cited on Page 1, 2, and 75)

[23] S. Hatamikia, A. Biguri, G. Kronreif, J. Kettenbach, T. Russ, H. Furtado, L. K. Shiyam Sundar, M. Buschmann, E. Unger, M. Figl, D. Georg, and W. Birkfellner, "Optimization for customized trajectories in cone beam computed tomography," *Medical Physics*, vol. 47, no. 10, pp. 4786–4799, 2020.   (cited on Page 1)

[24] K. Chung, L. R. Schad, and F. G. Zöllner, "Tomosynthesis implementation with adaptive online calibration on clinical C-arm systems," *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, no. 10, pp. 1481–1495, 2018.   (cited on Page 1, 2, and 77)

[25] Z. Yu, A. Maier, G. Lauritsch, F. Vogt, M. Schönborn, C. Köhler, J. Hornegger, and F. Noo, "Axially extended-volume C-arm CT using a reverse helical trajectory in the interventional room," *IEEE Transactions on Medical Imaging*, vol. 34, no. 1, pp. 203–215, 2015.   (cited on Page 1)

[26] Z. Yu, G. Lauritsch, F. Dennerlein, Y. Mao, J. Hornegger, and F. Noo, "Extended ellipse-line-ellipse trajectory for long-object cone-beam imaging with a mounted C-arm system," *Physics in Medicine and Biology*, vol. 61, no. 4, pp. 1829–1851, 2016.   (cited on Page 1)

[27] T. Reynolds, Y. Q. Ma, G. J. Gang, O. Dillon, T. Russ, W. Wang, T. Ehtiati, C. Weiss, N. Theodore, J. Siewerdsen, R. O'Brien, and J. W. Stayman, "Imaging from the cervical to the lumbar spine with a continuous multi-turn reverse helical 3D cone-beam CT scan," in *Annual Meeting of the American Association of Physicists*, 2021.   (cited on Page 1, 2, and 77)

[28] G. J. Gang, D. J. Tward, J. Lee, and J. H. Siewerdsen, "Anatomical background and generalized detectability in tomosynthesis and cone-beam CT," *Medical Physics*, vol. 37, no. 5, pp. 1948–1965, 2010.   (cited on Page 2)

[29] G. J. Gang, J. Lee, J. W. Stayman, D. J. Tward, W. Zbijewski, J. L. Prince, and J. H. Siewerdsen, "Analysis of Fourier-domain task-based detectability index in tomosynthesis and cone-beam CT in relation to human observer performance," *Medical Physics*, vol. 38, no. 4, pp. 1754–1768, 2011. (cited on Page 2)

[30] G. J. Gang, J. W. Stayman, T. Ehtiati, and J. H. Siewerdsen, "Task-driven image acquisition and reconstruction in cone-beam CT," *Physics in Medicine and Biology*, vol. 60, no. 8, pp. 3129–3150, 2015. (cited on Page 2)

[31] P. Sharp, D. C. Barber, D. G. Brown, A. E. Burgess, C. E. Metz, K. J. Myers, C. J. Taylor, R. F. Wagner, R. Brooks, C. R. Hill, D. E. Kuhl, M. A. Smith, P. Wells, and B. Worthington, "Report 54," *Journal of the International Commission on Radiation Units and Measurements*, no. 1, 1996. (cited on Page 2 and 34)

[32] A. M. Davis, E. A. Pearson, X. Pan, C. A. Pelizzari, and H. Al-Hallaq, "Collision-avoiding imaging trajectories for linac mounted cone-beam CT," *Journal of X-Ray Science and Technology*, vol. 27, no. 1, pp. 1–16, 2019. (cited on Page 2)

[33] M. Thies, J. N. Zäch, C. Gao, R. Taylor, N. Navab, A. Maier, and M. Unberath, "A learning-based method for online adjustment of C-arm Cone-beam CT source trajectories for artifact avoidance," *International Journal of Computer Assisted Radiology and Surgery*, vol. 15, no. 11, pp. 1787–1796, 2020. (cited on Page 2)

[34] Z. Zhao, G. J. Gang, and J. H. Siewerdsen, "Noise, sampling, and the number of projections in cone-beam CT with a flat-panel detector," *Medical Physics*, vol. 41, no. 6, 2014. (cited on Page 2)

[35] Y. Q. Ma, T. Reynolds, G. J. Gang, O. Dillon, T. Russ, W. Wang, T. Ehtiati, C. Weiss, N. Theodore, J. Siewerdsen, R. O'Brien, and J. W. Stayman, "Non-circular orbits on a clinical robotic C-arm for reducing metal artifacts in orthopedic interventions," in *Annual Meeting of the American Association of Physicists*, 2021. (cited on Page 2 and 77)

[36] S. Ouadah, J. W. Stayman, G. J. Gang, T. Ehtiati, and J. H. Siewerdsen, "Self-calibration of cone-beam CT geometry using 3D-2D image registration," *Physics in Medicine and Biology*, vol. 61, no. 7, pp. 2613–2632, 2016. (cited on Page 2)

[37] E.-P. Rührnschopf and and K. Klingenbeck, "A general framework and review of scatter correction methods in cone beam ct. part 2: Scatter estimation approaches," *Medical physics*, vol. 38, no. 9, pp. 5186–5199, 2011. (cited on Page 2)

[38] M. R. Ay and H. Zaidi, "Development and validation of MCNP4C-based Monte Carlo simulator for fan- and cone-beam X-ray CT," *Physics in Medicine and Biology*, vol. 50, no. 20, pp. 4863–4885, 2005. (cited on Page 3)

[39] S. A. Pozzi, E. Padovani, and M. Marseguerra, "MCNP-PoliMi: a Monte-Carlo code for correlation measurements," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 513, no. 3, pp. 550–558, 2003. (cited on Page 3)

[40] M. Caon, G. Bibbo, and J. Pattison, "A comparison of radiation dose measured in CT dosimetry phantoms with calculations using EGS4 and voxel-based computational models," *Physics in Medicine and Biology*, vol. 42, no. 1, pp. 219–229, 1997. (cited on Page 3)

[41] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, F. Behner, L. Bellagamba, J. Boudreau, L. Broglia, A. Brunengo, H. Burkhardt, S. Chauvie, J. Chuma, R. Chytracek, G. Cooperman, G. Cosmo, P. Degtyarenko, A. Dell'Acqua, G. Depaola, D. Dietrich, R. Enami, A. Feliciello, C. Ferguson, H. Fesefeldt, G. Folger, F. Foppiano, A. Forti, S. Garelli, S. Giani, R. Giannitrapani, D. Gibin, J. J. Gomez Cadenas, I. Gonzalez, G. Gracia Abril, G. Greeniaus, W. Greiner, V. Grichine, A. Grossheim, S. Guatelli, P. Gumplinger, R. Hamatsu, K. Hashimoto, H. Hasui, A. Heikkinen, A. Howard, V. Ivanchenko, A. Johnson, F. W. Jones, J. Kallenbach, N. Kanaya, M. Kawabata, Y. Kawabata, M. Kawaguti, S. Kelner, P. Kent, A. Kimura, T. Kodama, R. Kokoulin, M. Kossov, H. Kurashige, E. Lamanna, T. Lampen, V. Lara, V. Lefebure, F. Lei, M. Liendl, W. Lockman, F. Longo, S. Magni, M. Maire, E. Medernach, K. Minamimoto, P. Mora de Freitas, Y. Morita, K. Murakami, M. Nagamatu, R. Nartallo, P. Nieminen, T. Nishimura, K. Ohtsubo, M. Okamura, S. O'Neale, Y. Oohata, K. Paech, J. Perl, A. Pfeiffer, M. G. Pia, F. Ranjard, A. Rybin, S. Sadilov, E. di Salvo, G. Santin, T. Sasaki, N. Savvas, Y. Sawada, S. Scherer, S. Sei, V. Sirotenko, D. Smith, N. Starkov, H. Stoecker, J. Sulkimo, M. Takahata, S. Tanaka, E. Tcherniaev, E. Safai Tehrani, M. Tropeano, P. Truscott, H. Uno, L. Urban, P. Urban, M. Verderi, A. Walkden, W. Wander, H. Weber, J. P. Wellisch, T. Wenaus, D. C. Williams, D. Wright, T. Yamada, H. Yoshida, and D. Zschiesche, "GEANT4 - A simulation toolkit," *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250–303, 2003. (cited on Page 3 and 42)

[42] S. Jan, D. Benoit, E. Becheva, H.-h. Lin, K.-s. Chuang, Y.-h. Lin, C. Pet, S. España, J. L. Herraiz, E. Vicente, D. Lazaro, I. Buvat, and G. Loudos, "GATE : a simulation toolkit for PET and SPECT," *Physics in Medicine and Biology*, vol. 49, pp. 4543–4561, 2004. (cited on Page 3)

[43] S. Jan, D. Benoit, E. Becheva, T. Carlier, F. Cassol, P. Descourt, T. Frisson, L. Grevillot, L. Guigues, L. Maigne, C. Morel, Y. Perrot, N. Rehfeld, D. Sarrut, D. R. Schaart, S. Stute, U. Pietrzyk, D. Visvikis, N. Zahra, and I. Buvat, "GATE V6: A major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy," *Physics in Medicine and Biology*, vol. 56, no. 4, pp. 881–901, 2011. (cited on Page 3 and 42)

[44] J. Bert, H. Perez-Ponce, Z. E. Bitar, S. Jan, Y. Boursier, D. Vintache, A. Bonissent, C. Morel, D. Brasse, and D. Visvikis, "Geant4-based Monte Carlo simulations on GPU for medical applications," *Physics in Medicine and Biology*, vol. 58, no. 16, pp. 5593–5611, 2013.   (cited on Page 3)

[45] S. Camarasu-Pop, T. Glatard, R. F. da Silva, P. Gueth, D. Sarrut, and H. Benoit-Cattin, "Monte Carlo simulation on heterogeneous distributed systems: A computing framework with parallel merging and checkpointing strategies," *Future Generation Computer Systems*, vol. 29, no. 3, pp. 728–738, 2013. (cited on Page 3 and 45)

[46] S. Camarasu-Pop, T. Glatard, and H. Benoit-Cattin, "Combining analytical modeling, realistic simulation and real experimentation for the optimization of Monte-Carlo applications on the European Grid Infrastructure," *Future Generation Computer Systems*, vol. 57, pp. 13–23, 2016.   (cited on Page 3 and 45)

[47] J. Maier, S. Sawall, M. Knaup, and M. Kachelrieß, "Deep Scatter Estimation (DSE): Accurate real-time scatter estimation for X-Ray CT using a deep convolutional neural network," *Journal of Nondestructive Evaluation*, vol. 37, no. 3, pp. 1–9, 2018.   (cited on Page 3)

[48] J. T. Guibas, T. S. Virdi, and P. S. Li, "Synthetic medical images from dual generative adversarial networks," *CoRR*, vol. https://arxiv.org/abs/1709.01872, 2017.   (cited on Page 3)

[49] C. Bermudez, A. J. Plassard, L. T. Davis, A. T. Newton, S. M. Resnick, and B. A. Landman, "Learning implicit brain MRI manifolds with deep learning," in *Proc. SPIE 10574, Medical Imaging 2018: Image Processing*, vol. 105741L, 2018.   (cited on Page 3)

[50] S. Olut, Y. H. Sahin, U. Demir, and G. Unal, "Generative adversarial training for MRA image synthesis using multi-contrast MRI," in *PRedictive Intelligence in MEdicine*, pp. 147–154, 2018.   (cited on Page 3)

[51] J.-y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2242–2251, 2017. (cited on Page 3, 26, and 40)

[52] 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 *Simulation and Synthesis in Medical Imaging*, pp. 14–23, 2017.   (cited on Page 3)

[53] R. L. Siddon, "Fast calculation of the exact radiological path for a three-dimensional CT array," *Medical Physics*, vol. 12, no. 2, pp. 252–255, 1985. (cited on Page 3 and 19)

[54] G. Han, Z. Liang, and J. You, "A fast ray-tracing technique for TCT and ECT studies," in *1999 IEEE Nuclear Science Symposium. Conference Record. 1999 Nuclear Science Symposium and Medical Imaging Conference (Cat. No.99CH37019)*, vol. 3, pp. 1515–1518, IEEE, 1999. (cited on Page 3 and 19)

[55] X. Jia, H. Yan, L. Cerviño, M. Folkerts, and S. B. Jiang, "A GPU tool for efficient, accurate, and realistic simulation of cone beam CT projections," *Medical Physics*, vol. 39, no. 12, pp. 7368–7378, 2012. (cited on Page 4 and 19)

[56] Y. Long, J. A. Fessler, and J. M. Balter, "3D forward and back-projection for X-ray CT using separable footprints," *IEEE Transactions on Medical Imaging*, vol. 29, no. 11, pp. 1839–1850, 2010. (cited on Page 4 and 19)

[57] A. Katsevich, "Theoretically exact filtered backprojection-type inversion algorithm for spiral CT," *SIAM Journal on Applied Mathematics*, vol. 62, no. 6, pp. 2012–2026, 2002. (cited on Page 4)

[58] C. Bontus, P. Koken, T. Köhler, and M. Grass, "CEnPiT: Helical cardiac CT reconstruction," *Medical Physics*, vol. 33, no. 8, pp. 2792–2799, 2006. (cited on Page 4)

[59] J. D. Pack, F. Noo, and R. Clackdoyle, "Cone-beam reconstruction using the backprojection of locally filtered projections," *IEEE Transactions on Medical Imaging*, vol. 24, no. 1, pp. 70–85, 2005. (cited on Page 4 and 75)

[60] A. Katsevich, "Image reconstruction for the circle and line trajectory," *Physics in Medicine and Biology*, vol. 49, no. 22, pp. 5059–5072, 2004. (cited on Page 4)

[61] Y. Lu, J. Yang, J. W. Emerson, H. Mao, T. Zhou, Y. Si, and M. Jiang, "Cone-beam reconstruction for the two-circles-plus-one-line trajectory," *Physics in Medicine and Biology*, vol. 57, no. 9, pp. 2689–2707, 2012. (cited on Page 4)

[62] J. Wang, T. Li, H. Lu, and Z. Liang, "Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose X-ray computed tomography," *IEEE Transactions on Medical Imaging*, vol. 25, no. 10, pp. 1272–1283, 2006. (cited on Page 4)

[63] H. Shan, A. Padole, F. Homayounieh, U. Kruger, R. D. Khera, C. Nitiwarangkul, M. K. Kalra, and G. Wang, "Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction," *Nature Machine Intelligence*, vol. 1, no. 6, pp. 269–276, 2019. (cited on Page 4)

[64] H. Chen, Y. Zhang, Y. Chen, J. Zhang, W. Zhang, H. Sun, Y. Lv, P. Liao, J. Zhou, and G. Wang, "LEARN: Learned Experts' Assessment-Based Reconstruction Network for sparse-data CT," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1333–1347, 2018. (cited on Page 4)

[65] I. Y. Chun, Z. Huang, H. Lim, and J. Fessler, "Momentum-Net: Fast and convergent iterative neural network for inverse problems," *IEEE Transactions*

on Pattern Analysis and Machine Intelligence, vol. 8828, pp. 1–17, 2020. (cited on Page 4)

[66] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui, "4D XCAT phantom for multimodality imaging research," *Medical Physics*, vol. 37, no. 9, pp. 4902–4915, 2010. (cited on Page 5 and 28)

[67] W. van Aarle, W. J. Palenstijn, J. Cant, E. Janssens, F. Bleichrodt, A. Dabravolski, J. De Beenhouwer, K. Joost Batenburg, and J. Sijbers, "Fast and flexible X-ray tomography using the ASTRA toolbox," *Optics Express*, vol. 24, no. 22, p. 25129, 2016. (cited on Page 5 and 46)

[68] W. J. Palenstijn, K. J. Batenburg, and J. Sijbers, "Performance improvements for iterative electron tomography reconstruction using graphics processing units (GPUs)," *Journal of Structural Biology*, vol. 176, no. 2, pp. 250–253, 2011. (cited on Page 5)

[69] W. C. Röntgen, "On a new kind of rays," *The British Journal of Radiology*, vol. 4, no. 37, pp. 32–33, 1931. (cited on Page 8)

[70] J. Radon, "On the determination of functions from their integral values along certain manifolds," *IEEE Transactions on Medical Imaging*, vol. 5, no. 4, pp. 170–176, 1986. (cited on Page 8)

[71] T. Buzug, *Computed Tomography: From Photon Statistics to Modern Cone-Beam CT*. Springer Berlin Heidelberg, 2008. (cited on Page 8, 10, 17, and 20)

[72] J. Hsieh, *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*. SPIE PRESS, 2015. (cited on Page 8)

[73] A. Maier, S. Steidl, V. Christlein, and J. Hornegger, eds., *Medical Imaging Systems*. Springer International Publishing, 2018. (cited on Page 8 and 9)

[74] J. H. Hubbell, "Review of photon interaction cross section data in the medical and biological context," *Physics in Medicine and Biology*, vol. 44, no. 1, pp. R1–R22, 1999. (cited on Page 13)

[75] R. P. Feynman and R. B. Leighton, eds., *The Feynman Lectures on Physics; Vol. 1*. Addison-Wesley Publishing, Massachusetts, 1963. (cited on Page 14)

[76] C. Leroy and P.-G. Rancoita, *Principles of radiation interaction in matter and detection*. World Scientific, 2004. (cited on Page 14)

[77] H. K. Tuy, "An inversion formula for cone-beam reconstruction," *SIAM Journal on Applied Mathematics*, vol. 43, no. 3, pp. 546–552, 1983. (cited on Page 18)

[78] M. Matsumoto and T. Nishimura, "Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator," *ACM Trans. Model. Comput. Simul.*, vol. 8, no. 1, p. 3–30, 1998. (cited on Page 18)

[79] X. Jia, B. Dong, Y. Lou, and S. B. Jiang, "GPU-based iterative cone-beam CT reconstruction using tight frame regularization," *Physics in Medicine and Biology*, vol. 56, no. 13, pp. 3787–3807, 2011.   (cited on Page 19)

[80] A. H. Andersen and A. C. Kak, "Simultaneous algebraic reconstruction technique (sart): A superior implementation of the art algorithm," *Ultrasonic Imaging*, vol. 6, no. 1, pp. 81–94, 1984.   (cited on Page 21)

[81] S. Tilley, M. Jacobson, Q. Cao, M. Brehler, A. Sisniega, W. Zbijewski, and J. W. Stayman, "Penalized-likelihood reconstruction with high-fidelity measurement models for high-resolution cone-beam imaging," *IEEE Transactions on Medical Imaging*, vol. 37, no. 4, pp. 988–999, 2018.   (cited on Page 21)

[82] P. Huber and E. Ronchetti, "Robust statistics," *Wiley Ser Probab Math Stat New York, NY, USA Wiley-IEEE*, vol. 52, p. 54, 1981.   (cited on Page 22)

[83] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning.* MIT press, 2016.   (cited on Page 22)

[84] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," 2017.   (cited on Page 23)

[85] 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*, vol. 14, no. 10, pp. 1741–1750, 2019.   (cited on Page 25, 53, and 76)

[86] 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," *4th Conference on Image-Guided Interventions*, 2019.   (cited on Page 25, 53, and 76)

[87] T. Russ, A. M. Abdelrehim, D. F. Bauer, S. Hatamikia, F. G. Zöllner, L. R. Schad, and K. Chung, "CBCT image quality and dose simulations for arbitrary source-detector trajectories with GATE," *European Congress of Radiology*, 2020.   (cited on Page 25, 53, and 76)

[88] T. Russ, W. Wang, A.-K. Golla, D. F. Bauer, M. Tivnan, C. Tönnes, Y. Q. Ma, T. Reynolds, S. Hatamikia, L. R. Schad, F. G. Zöllner, G. J. Gang, and J. W. Stayman, "Fast Reconstruction of non-circular CBCT orbits using CNNs," *16th International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*, 2021.   (cited on Page 25, 53, and 76)

[89] T. Russ, Y. Q. Ma, A.-K. Golla, D. F. Bauer, , C. Tönnes, T. Reynolds, S. Hatamikia, L. R. Schad, F. G. Zöllner, G. J. Gang, W. Wang, and J. W. Stayman, "Fast CBCT Reconstruction using Convolutional Neural Networks for Arbitrary Robotic C-arm Orbits," *Accepted for Oral presentation at SPIE Medical Imaging*, 2022.   (cited on Page 25, 53, and 76)

[90] 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 25)

[91] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.   (cited on Page 25)

[92] P. Christ, F. Ettlinger, J. Lipkova, and G. Kaissis, "Lits - liver tumor segmentation challenge," 2017.   (cited on Page 28)

[93] L. Soler, A. Hostettler, V. Agnus, A. Charnoz, J. Fasquel, J. Moreau, A. Osswald, M. Bouhadjar, and J. Marescaux, "3D Image Reconstruction for Comparison of Algorithm Database: A patient specific anatomical and medical image database," 2010.   (cited on Page 28)

[94] S. G. Armato III, G. McLennan, L. Bidaut, M. F. McNitt-Gray, C. R. Meyer, A. P. Reeves, B. Zhao, D. R. Aberle, C. I. Henschke, E. A. Hoffman, E. A. Kazerooni, H. MacMahon, E. J. R. Van Beek, D. Yankelevitz, A. M. Biancardi, P. H. Bland, M. S. Brown, R. M. Engelmann, G. E. Laderach, D. Max, R. C. Pais, D. P. Y. Qing, R. Y. Roberts, A. R. Smith, A. Starkey, P. Batra, P. Caligiuri, A. Farooqi, G. W. Gladish, C. M. Jude, R. F. Munden, I. Petkovska, L. E. Quint, L. H. Schwartz, B. Sundaram, L. E. Dodd, C. Fenimore, D. Gur, N. Petrick, J. Freymann, J. Kirby, B. Hughes, A. V. Casteele, S. Gupte, M. Sallam, M. D. Heath, M. H. Kuhn, E. Dharaiya, R. Burns, D. S. Fryd, M. Salganicoff, V. Anand, U. Shreter, S. Vastagh, B. Y. Croft, and L. P. Clarke, "Data from LIDC-IDRI," *The Cancer Imaging Archive*, 2015.   (cited on Page 28)

[95] V. Spitzer, M. J. Ackerman, A. L. Scherzinger, and D. Whitlock, "The visible human male: A technical report," *Journal of the American Medical Informatics Association*, vol. 3, no. 2, pp. 118–130, 1996.   (cited on Page 28)

[96] L. Piegl, "On NURBS: A survey," *IEEE Computer Graphics and Applications*, vol. 11, no. 1, pp. 55–71, 1991.   (cited on Page 28)

[97] B. Delaunay, "Sur la sphère vide. A la mémoire de Georges Voronoï," *Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et na*, vol. 6, pp. 793–800, 1934.   (cited on Page 30)

[98] P. R. Smith, T. M. Peters, and R. H. Bates, "Image reconstruction from finite numbers of projections," *Journal of Physics A: General Physics*, vol. 6, no. 3, pp. 361–382, 1973.   (cited on Page 32)

[99] J. G. Gang, "Task-based imaging performance in 3D X-Ray tomography: Noise, detectability, and implications for system design," *PhD Thesis*, 2014.   (cited on Page 33 and 35)

[100] M. Rabbani, R. Shaw, and R. V. Metter, "Detective quantum efficiency of imaging systems with amplifying and scattering mechanisms," *J. Opt. Soc. Am. A*, vol. 4, no. 5, pp. 895–901, 1987.   (cited on Page 34)

[101] D. J. Tward and J. H. Siewerdsen, "Cascaded systems analysis of the 3D noise transfer characteristics of flat-panel cone-beam CT," *Medical Physics*, vol. 35, no. 12, pp. 5510–5529, 2008.   (cited on Page 34)

[102] J. Baek and N. J. Pelc, "Use of sphere phantoms to measure the 3D MTF of FDK reconstructions," in *Physics* (N. J. Pelc, E. Samei, and R. M. Nishikawa, eds.), vol. 7961, p. 79610D, 2011.   (cited on Page 34)

[103] A. L. C. Kwan, J. M. Boone, K. Yang, and S.-Y. Huang, "Evaluation of the spatial resolution characteristics of a cone-beam breast CT scanner," *Medical Physics*, vol. 34, no. 1, pp. 275–281, 2006.   (cited on Page 34)

[104] S. N. Friedman, G. S. Fung, J. H. Siewerdsen, and B. M. Tsui, "A simple approach to measure computed tomography (CT) modulation transfer function (MTF) and noise-power spectrum (NPS) using the American College of Radiology (ACR) accreditation phantom," *Medical Physics*, vol. 40, no. 5, pp. 1–9, 2013.   (cited on Page 34 and 35)

[105] J. M. Boone, "Determination of the presampled MTF in computed tomography," *Medical Physics*, vol. 28, no. 3, pp. 356–360, 2001.   (cited on Page 34)

[106] P. Walek, J. Jan, P. Ourednicek, J. Skotakova, and I. Jira, "Methodology for estimation of tissue noise power spectra in iteratively reconstructed MDCT data," *21st International Conference on Computer Graphics, Visualization and Computer Vision*, pp. 243–252, 2013.   (cited on Page 35)

[107] 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 35 and 42)

[108] Z. Wang, A. C. Bovik, and H. R. Sheikh, "Image quality assessment: From error measurement to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600 – 612, 2004.   (cited on Page 37)

[109] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A Feature Similarity Index for image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.   (cited on Page 37)

[110] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986. (cited on Page 37)

[111] 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, 2016.   (cited on Page 37)

[112] A. Odena, V. Dumoulin, and C. Olah, "Deconvolution and checkerboard arti-
      facts," *Distill*, 2016.   (cited on Page 40)

[113] A.-K. Golla, "Optimized training pipeline for deep learning applications in
      medical image processing," *PhD Thesis*, 2021.   (cited on Page 40)

[114] X. Jin, Y. Qi, and S. Wu, "Cyclegan face-off," *CoRR*,
      vol. https://arxiv.org/abs/1712.03451, 2017.   (cited on Page 41)

[115] G. G. Poludniowski and P. M. Evans, "Calculation of X-ray spectra emerg-
      ing from an X-ray tube. Part I. Electron penetration characteristics in X-ray
      targets," *Medical Physics*, vol. 34, pp. 2164–2174, 2007.   (cited on Page 45)

[116] G. G. Poludniowski, "Calculation of X-ray spectra emerging from an X-ray
      tube. Part II. X-ray production and filtration in X-ray targets," *Medical
      Physics*, vol. 34, pp. 2175–2186, 2007.   (cited on Page 45)

[117] G. G. Poludniowski, G. Landry, F. DeBlois, P. M. Evans, and F. Verhaegen,
      "SpekCalc: a program to calculate photon spectra from tungsten anode X-ray
      tubes," *Physics in Medicine and Biology*, vol. 54, no. 19, pp. N433–N438, 2009.
      (cited on Page 45)

[118] Z. Guo, G. Lauritsch, A. Maier, and F. Noo, "Theoretically-exact filtered-
      backprojection reconstruction from real data on the line-ellipse-line trajectory,"
      in *15th International Meeting on Fully Three-Dimensional Image Reconstruc-
      tion in Radiology and Nuclear Medicine*, 2019.   (cited on Page 75)

[119] 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," in *International Conference on Medical Imaging with
      Deep Learning - Extended Abstract Track*, 2019.   (cited on Page 76)

[120] 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," in *4th Conference on Image-Guided Interventions*, 2019.
      (cited on Page 76)

[121] D. F. Bauer, A.-K. Russ, T. Schnurr, L. R. Schad, and F. G. Zöllner, "Gen-
      eration of fully annotated abdominal T1-weighted MR ground truth data for
      image segmentation and registration using CycleGANs and the XCAT phan-
      tom," in *Proceedings ESMRMB Congress*, 2020.   (cited on Page 76)

[122] D. F. Bauer, T. Russ, B. Waldkirch, C. Tönnes, W. Segars, L. R. Schad,
      F. G. Zöllner, and A.-K. Golla, "Generation of annotated multimodal ground
      truth datasets for abdominal medical image registration," *Int. J. Comput. Ass.
      Radiol. Surg.*, vol. 16, no. 8, pp. 1277–1285, 2021.   (cited on Page 76)

[123] 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 79)

[124] S. Richard, D. B. Husarik, G. Yadava, S. N. Murphy, and E. Samei, "Towards task-based assessment of CT performance: System and object MTF across different reconstruction algorithms," *Medical Physics*, vol. 39, no. 7, pp. 4115–4122, 2012. (cited on Page 81)

[125] F. Aghdasi and R. K. Ward, "Reduction of boundary artifacts in image restoration," *IEEE Transactions on Image Processing*, vol. 5, no. 4, pp. 611–618, 1996. (cited on Page 83)

[126] R. Liu and J. Jia, "Reducing boundary artifacts in image deconvolution," *Proceedings - International Conference on Image Processing, ICIP*, pp. 505–508, 2008. (cited on Page 83)

[127] B. Münch, P. Trtik, F. Marone, and M. Stampanoni, "Stripe and ring artifact removal with combined wavelet-Fourier filtering," *EMPA Activities*, vol. 17, no. 2009-2010 EMPA ACTIVITIES, pp. 34–35, 2009. (cited on Page 83)

[128] M. Šorel, "Removing boundary artifacts for real-time iterated shrinkage deconvolution," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2329–2334, 2012. (cited on Page 83)

[129] A. Mosleh, J. M. Langlois, and P. Green, "Image deconvolution ringing artifact detection and removal via PSF frequency analysis," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8692 LNCS, no. 4, pp. 247–262, 2014. (cited on Page 83)

[130] J. D. Evans, D. G. Politte, B. R. Whiting, J. A. O'Sullivan, and J. F. Williamson, "Noise-resolution tradeoffs in X-ray CT imaging: A comparison of penalized alternating minimization and filtered backprojection algorithms," *Medical Physics*, vol. 38, no. 3, pp. 1444–1458, 2011. (cited on Page 83)

[131] G. J. Gang, X. Guo, and J. W. Stayman, "Generalized image quality analysis for nonlinear algorithms in CT," in *Annual Meeting of the American Association of Physicists*, 2019. (cited on Page 87)

List of Figures

List of Tables

List of Acronyms

ART	Algebraic reconstruction technique
CBCT	Cone-beam computed tomography
CNN	convolutional neural networks
COM	center of mass
CPU	central processing unit
CT	computed tomography
CycleGAN	Cycle-consistent generative-adversarial networks
DFT	discrete Fourier transformation
DSC	Dice similarity coefficient
EGR	edge generation ratio
EM	edge maps
EPR	edge preservation ratio
ESF	edge-spread function
FOV	field-of-view
FSIM	feature similarity index measure
FWHM	full width at half maximum
GAN	generative adversarial networks
GATE	GEANT4 application for tomographic tmission
GM	gradient magnitude
GPU	graphical processing units
GT	ground truth
HU	Hounsfield unit
IRCAD	3D image reconstruction for comparison of algorithm database
LIDC	Lung image database consortium image collection
LiTS	liver tumor segmentation challenge
LSF	line-spread function
MAE	mean absolute error

MBIR	model-based iterative reconstruction
MC	Monte Carlo
MDCT	multi-detector computed tomography
MPE	mean percentage error
MRI	magnetic resonance imaging
MTF	modulation transfer function
NCC	NPS correlation coefficient
ND	normal distribution
NM	noise magnitude
NMMR	NPS maximum-maximum ratio
NMPE	NPS mean percentage error
NPS	noise power spectrum
nRMSE	normalized root-mean-square error
PC	phase congruency
PSF	point-spread function
RCBV	ratio of corrupted background voxel
ReLU	rectified linear unit
RMSE	root-mean-square error
ROI	region of interest
SAD	source-axis distance
SART	simultaneous algebraic reconstruction technique
SDD	source-detector distance
SIRT	simultaneous iterative reconstruction technique
SSIM	structural similarity index measure
TTF	task transfer function
VOI	volume of interest
XCAT	extended cardiac-torso phantom

Declaration

This thesis is the result of my independent investigation under supervision. Where my work is indebted to the work or ideas of others, for example from the literature or the internet, I have acknowledged this within the thesis.

I declare that this study has not already been accepted for any degree, nor is it currently being submitted in candidature for any other degree.

I am aware that a false declaration could have legal implications.

Erklärung:

Ich versichere, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Mannheim, December 6th 2021

_____

Tom Russ